US 20040014093A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2004/0014093 A1**

Duclos et al. (43) **Pub. Date: Jan. 22, 2004**

(54) **POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN SUPERFAMILY, PROTEASE-42**

(76) Inventors: **Franck Duclos**, Washington Crossing, PA (US); **Jian Chen**, Princeton, NJ (US); **John N. Feder**, Belle Mead, NJ (US); **Akbar Nayeem**, Newtown, PA (US); **Thomas C. Nelson**, Lawrenceville, NJ (US)

Correspondence Address:
STEPHEN B. DAVIS
BRISTOL-MYERS SQUIBB COMPANY
PATENT DEPARTMENT
P O BOX 4000
PRINCETON, NJ 08543-4000 (US)

(21) Appl. No.: **10/390,585**

(22) Filed: **Mar. 14, 2003**

Related U.S. Application Data

(60) Provisional application No. 60/364,941, filed on Mar. 14, 2002.

Publication Classification

(51) **Int. Cl.**[7] **C12Q 1/68**; G06F 19/00; G01N 33/48; G01N 33/50; C07H 21/04; C12N 9/64; C12P 21/02; C12N 5/06

(52) **U.S. Cl.** **435/6**; 435/69.1; 435/226; 435/320.1; 435/325; 536/23.2; 702/19

(57) **ABSTRACT**

The present invention provides novel polynucleotides encoding Protease-42 polypeptides, fragments and homologues thereof. Also provided are vectors, host cells, antibodies, and recombinant and synthetic methods for producing said polypeptides. The invention further relates to diagnostic and therapeutic methods for applying these novel Protease-42 polypeptides to the diagnosis, treatment, and/or prevention of various diseases and/or disorders related to these polypeptides. The invention further relates to screening methods for identifying agonists and antagonists of the polynucleotides and polypeptides of the present invention.

FIG. 1A

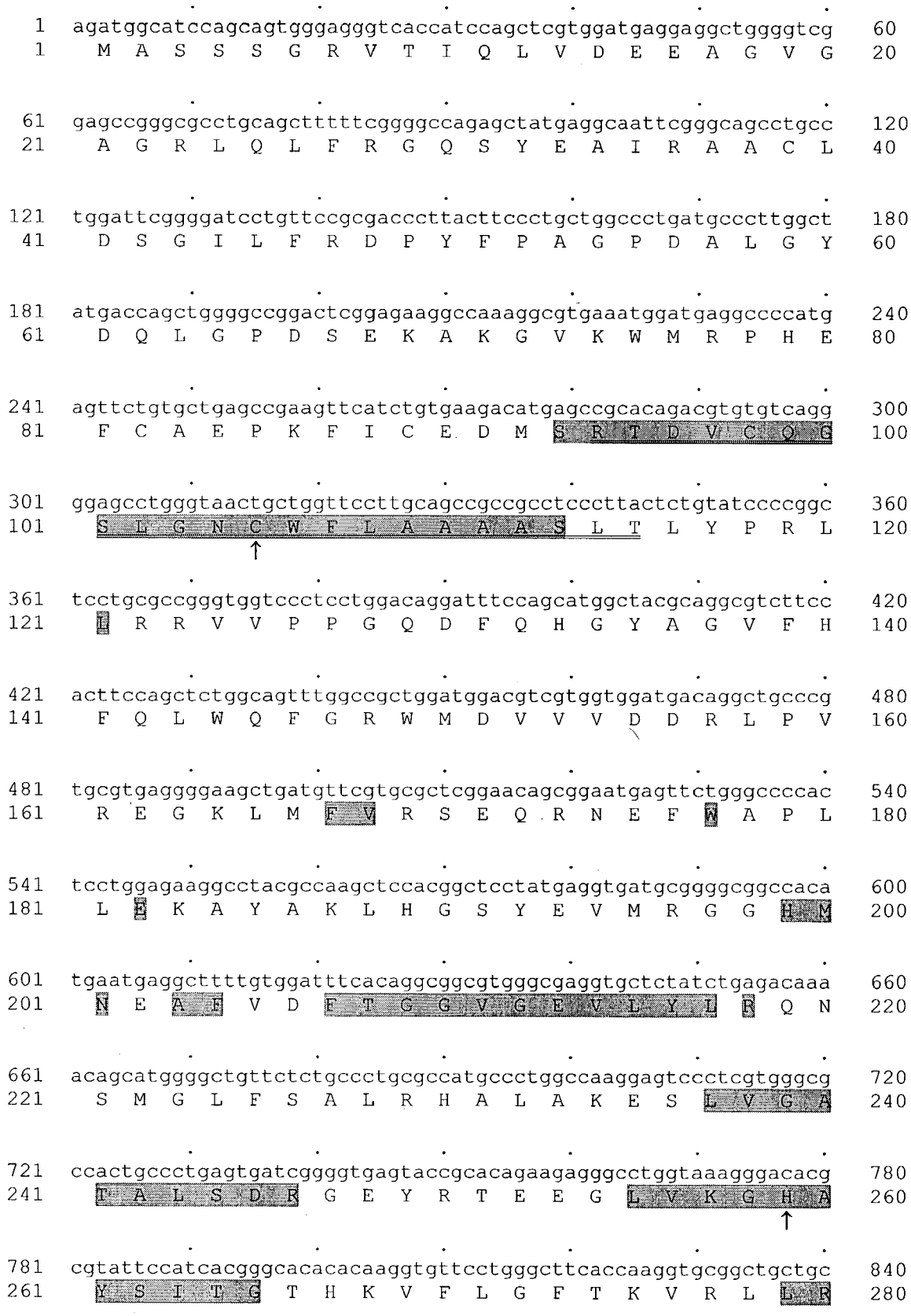

1 agatggcatccagcagtgggagggtcaccatccagctcgtggatgaggaggctggggtcg 60
1 M A S S S G R V T I Q L V D E E A G V G 20
61 gagccgggcgcctgcagcttttcggggccagagctatgaggcaattcgggcagcctgcc 120
21 A G R L Q L F R G Q S Y E A I R A A C L 40
121 tggattcggggatcctgttccgcgacccttacttccctgctggccctgatgcccttggct 180
41 D S G I L F R D P Y F P A G P D A L G Y 60
181 atgaccagctggggccggactcggagaaggccaaaggcgtgaaatggatgaggccccatg 240
61 D Q L G P D S E K A K G V K W M R P H E 80
241 agttctgtgctgagccgaagttcatctgtgaagacatgagccgcacagacgtgtgtcagg 300
81 F C A E P K F I C E D M S R T D V C Q G 100
301 ggagcctgggtaactgctggttccttgcagccgccgcctcccttactctgtatccccggc 360
101 S L G N C W F L A A A A S L T L Y P R L 120
361 tcctgcgccgggtggtccctcctggacaggatttccagcatggctacgcaggcgtcttcc 420
121 L R R V V P P G Q D F Q H G Y A G V F H 140
421 acttccagctctggcagtttggccgctggatggacgtcgtggtggatgacaggctgcccg 480
141 F Q L W Q F G R W M D V V V D D R L P V 160
481 tgcgtgaggggaagctgatgttcgtgcgctcggaacagcggaatgagttctgggccccac 540
161 R E G K L M F V R S E Q R N E F W A P L 180
541 tcctggagaaggcctacgccaagctccacggctcctatgaggtgatgcggggcggccaca 600
181 L E K A Y A K L H G S Y E V M R G G H M 200
601 tgaatgaggcttttgtggatttcacaggcggcgtgggcgaggtgctctatctgagacaaa 660
201 N E A F V D F T G G V G E V L Y L R Q N 220
661 acagcatggggctgttctctgccctgcgccatgccctggccaaggagtccctcgtgggcg 720
221 S M G L F S A L R H A L A K E S L V G A 240
721 ccactgccctgagtgatcggggtgagtaccgcacagaagagggcctggtaaagggacacg 780
241 T A L S D R G E Y R T E E G L V K G H A 260
781 cgtattccatcacgggcacacacaaggtgttcctgggcttcaccaaggtgcggctgctgc 840
261 Y S I T G T H K V F L G F T K V R L L R 280

FIG. 1B

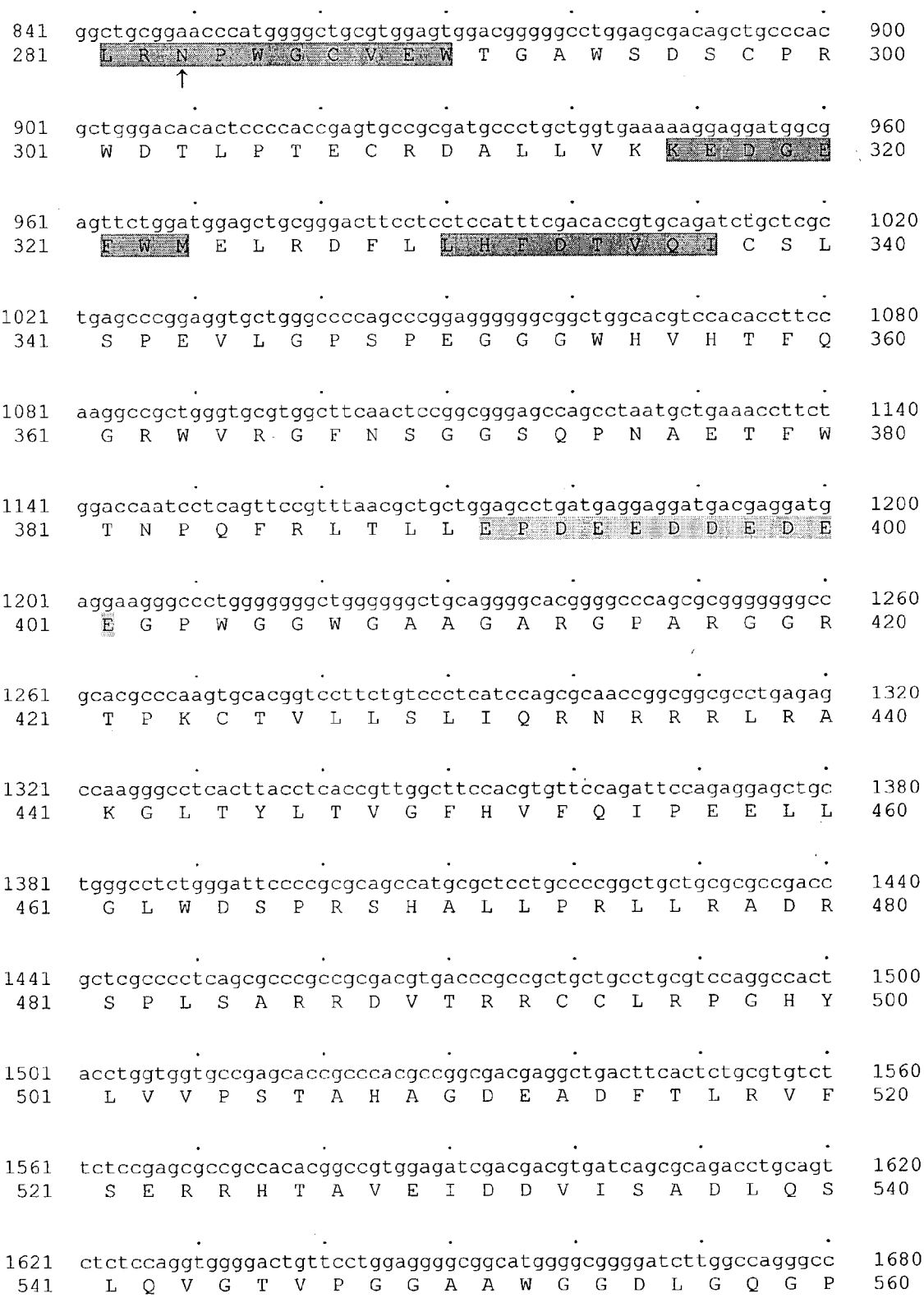

841 ggctgcggaacccatggggctgcgtggagtggacgggggcctggagcgacagctgcccac 900
281 L R N P W G C V E W T G A W S D S C P R 300
901 gctgggacacactccccaccgagtgccgcgatgccctgctggtgaaaaaggaggatggcg 960
301 W D T L P T E C R D A L L V K K E D G E 320
961 agttctggatggagctgcgggacttcctcctccatttcgacaccgtgcagatctgctcgc 1020
321 F W M E L R D F L L H F D T V Q I C S L 340
1021 tgagccccggaggtgctgggccccagcccggaggggggcggctggcacgtccacaccttcc 1080
341 S P E V L G P S P E G G G W H V H T F Q 360
1081 aaggccgctgggtgcgtggcttcaactccggcgggagccagcctaatgctgaaaccttct 1140
361 G R W V R G F N S G G S Q P N A E T F W 380
1141 ggaccaatcctcagttccgtttaacgctgctggagcctgatgaggaggatgacgaggatg 1200
381 T N P Q F R L T L L E P D E E D D E D E 400
1201 aggaagggccctgggggggctgggggctgcaggggcacggggcccagcgcgggggggcc 1260
401 E G P W G G W G A A G A R G P A R G G R 420
1261 gcacgcccaagtgcacggtccttctgtccctcatccagcgcaaccggcggcgcctgagag 1320
421 T P K C T V L L S L I Q R N R R R L R A 440
1321 ccaagggcctcacttacctcaccgttggcttccacgtgttccagattccagaggagctgc 1380
441 K G L T Y L T V G F H V F Q I P E E L L 460
1381 tgggcctctgggattccccgcgcagccatgcgctcctgccccggctgctgcgcgccgacc 1440
461 G L W D S P R S H A L L P R L L R A D R 480
1441 gctcgcccctcagcgcccgccgcgacgtgacccgccgctgctgcctgcgtccaggccact 1500
481 S P L S A R R D V T R R C C L R P G H Y 500
1501 acctggtggtgccgagcaccgcccacgccggcgacgaggctgacttcactctgcgtgtct 1560
501 L V V P S T A H A G D E A D F T L R V F 520
1561 tctccgagcgccgccacacggccgtggagatcgacgacgtgatcagcgcagacctgcagt 1620
521 S E R R H T A V E I D D V I S A D L Q S 540
1621 ctctccaggtggggactgttcctggaggggcggcatggggcggggatcttggccagggcc 1680
541 L Q V G T V P G G A A W G G D L G Q G P 560

FIG. 1C

```
1681  cctacctgcccctggagctggggttggagcagctgtttcaggagctggctggagaggagg  1740
 561    Y  L  P  L  E  L  G  L  E  Q  L  F  Q  E  L  A  G  E  E  E   580

1741  aagaactcaatgcctctcagctccaggccttactaagcattgccctggagcctgccaggg  1800
 581    E  L  N  A  S  Q  L  Q  A  L  L  S  I  A  L  E  P  A  R  A   600

1801  cccatacctccacccccagagagatcgggctcaggacctgtgagcagctgctgcagtgtt  1860
 601    H  T  S  T  P  R  E  I  G  L  R  T  C  E  Q  L  L  Q  C  F   620

1861  tcgggcatgggcaaagcctggccttacaccacttccagcagctctggggctacctcctgg  1920
 621    G  H  G  Q  S  L  A  L  H  H  F  Q  Q  L  W  G  Y  L  L  E   640

1921  agtggcaggccatattcaacaagttcgatgaggacacctctggaaccatgaactcctacg  1980
 641    W  Q  A  I  F  N  K  F  D  E  D  T  S  G  T  M  N  S  Y  E   660

1981  agctgaggctggcactgaatgcagcaggcttccacctgaacaaccagctgacccagaccc  2040
 661    L  R  L  A  L  N  A  A  G  F  H  L  N  N  Q  L  T  Q  T  L   680

2041  tcaccagccgctaccgggatagccgtctgcgtgtggacttcgagcggttcgtgtcctgtg  2100
 681    T  S  R  Y  R  D  S  R  L  R  V  D  F  E  R  F  V  S  C  V   700

2101  tggcccacctcacctgcatcttctgccactgcagccagcacctggatgggggtgaggggg  2160
 701    A  H  L  T  C  I  F  C  H  C  S  Q  H  L  D  G  G  E  G  V   720

2161  tcatctgcctgacccacagacagtggatggaggtggccaccttctcctaggatctccgga  2220
 721    I  C  L  T  H  R  Q  W  M  E  V  A  T  F  S                  735
```

FIG. 2A

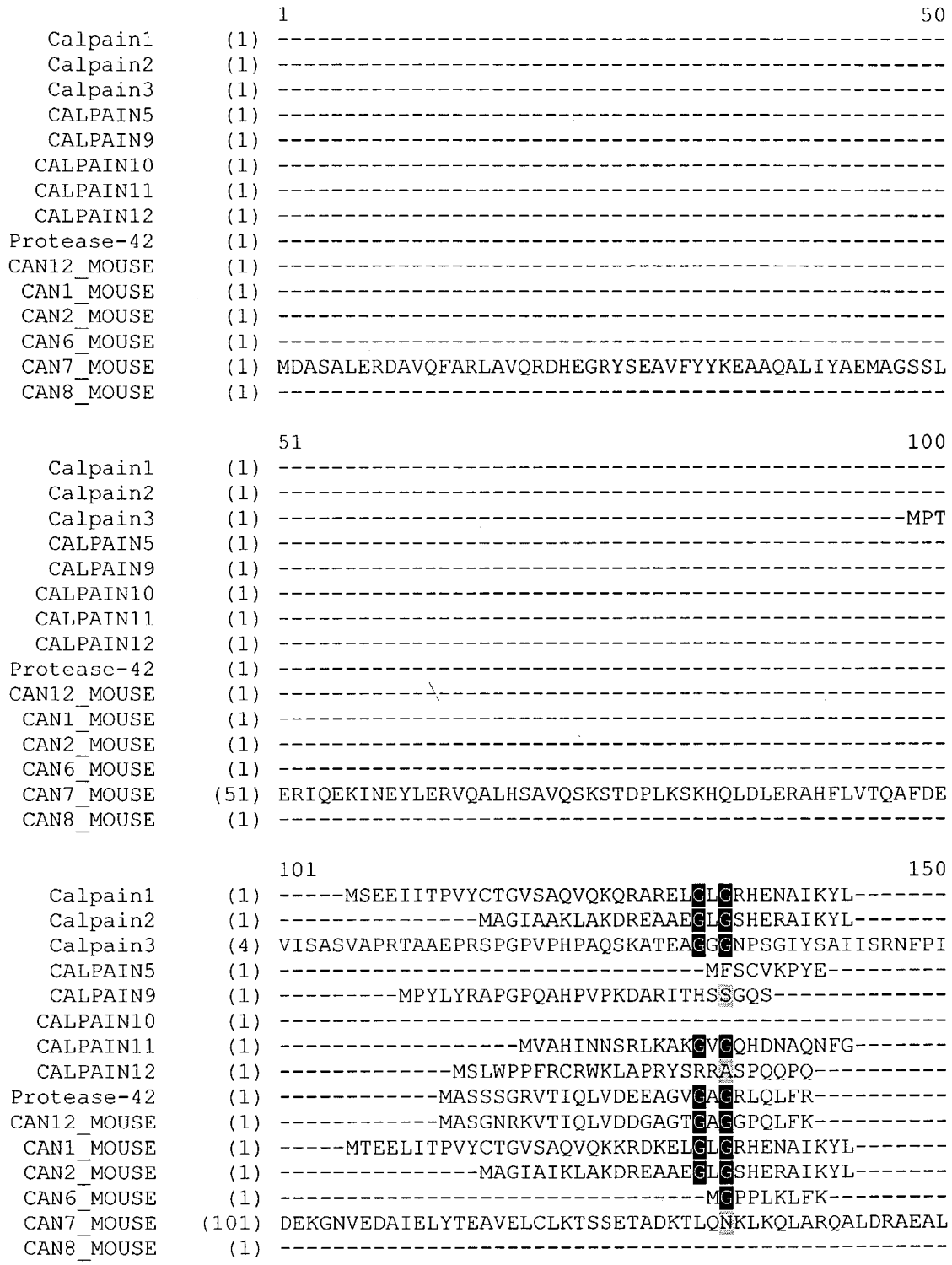

1 50
Calpain1 (1) ----------
Calpain2 (1) ----------
Calpain3 (1) ----------
CALPAIN5 (1) ----------
CALPAIN9 (1) ----------
CALPAIN10 (1) ----------
CALPAIN11 (1) ----------
CALPAIN12 (1) ----------
Protease-42 (1) ----------
CAN12_MOUSE (1) ----------
CAN1_MOUSE (1) ----------
CAN2_MOUSE (1) ----------
CAN6_MOUSE (1) ----------
CAN7_MOUSE (1) MDASALERDAVQFARLAVQRDHEGRYSEAVFYYKEAAQALIYAEMAGSSL
CAN8_MOUSE (1) ----------
51 100
Calpain1 (1) ----------
Calpain2 (1) ----------
Calpain3 (1) ----------MPT
CALPAIN5 (1) ----------
CALPAIN9 (1) ----------
CALPAIN10 (1) ----------
CALPAIN11 (1) ----------
CALPAIN12 (1) ----------
Protease-42 (1) ----------
CAN12_MOUSE (1) ----------
CAN1_MOUSE (1) ----------
CAN2_MOUSE (1) ----------
CAN6_MOUSE (1) ----------
CAN7_MOUSE (51) ERIQEKINEYLERVQALHSAVQSKSTDPLKSKHQLDLERAHFLVTQAFDE
CAN8_MOUSE (1) ----------
101 150
Calpain1 (1) -----MSEEIITPVYCTGVSAQVQKQRARELGLGRHENAIKYL-------
Calpain2 (1) ---------------MAGIAAKLAKDREAAEGLGSHERAIKYL-------
Calpain3 (4) VISASVAPRTAAEPRSPGPVPHPAQSKATEAGGGNPSGIYSAIISRNFPI
CALPAIN5 (1) ----------MFSCVKPYE---------
CALPAIN9 (1) ---------MPYLYRAPGPQAHPVPKDARITHSSGQS----------
CALPAIN10 (1) ----------
CALPAIN11 (1) ----------MVAHINNSRLKAKGVGQHDNAQNFG-------
CALPAIN12 (1) ----------MSLWPPFRCRWKLAPRYSRRASPQQPQ----------
Protease-42 (1) ----------MASSSGRVTIQLVDEEAGVGAGRLQLFR----------
CAN12_MOUSE (1) ----------MASGNRKVTIQLVDDGAGTGAGGPQLFK----------
CAN1_MOUSE (1) -----MTEELITPVYCTGVSAQVQKKRDKELGLGRHENAIKYL-------
CAN2_MOUSE (1) ---------------MAGIAIKLAKDREAAEGLGSHERAIKYL-------
CAN6_MOUSE (1) ----------MGPPLKLFK---------
CAN7_MOUSE (101) DEKGNVEDAIELYTEAVELCLKTSSETADKTLQNKLKQLARQALDRAEAL
CAN8_MOUSE (1) ----------

FIG. 2B

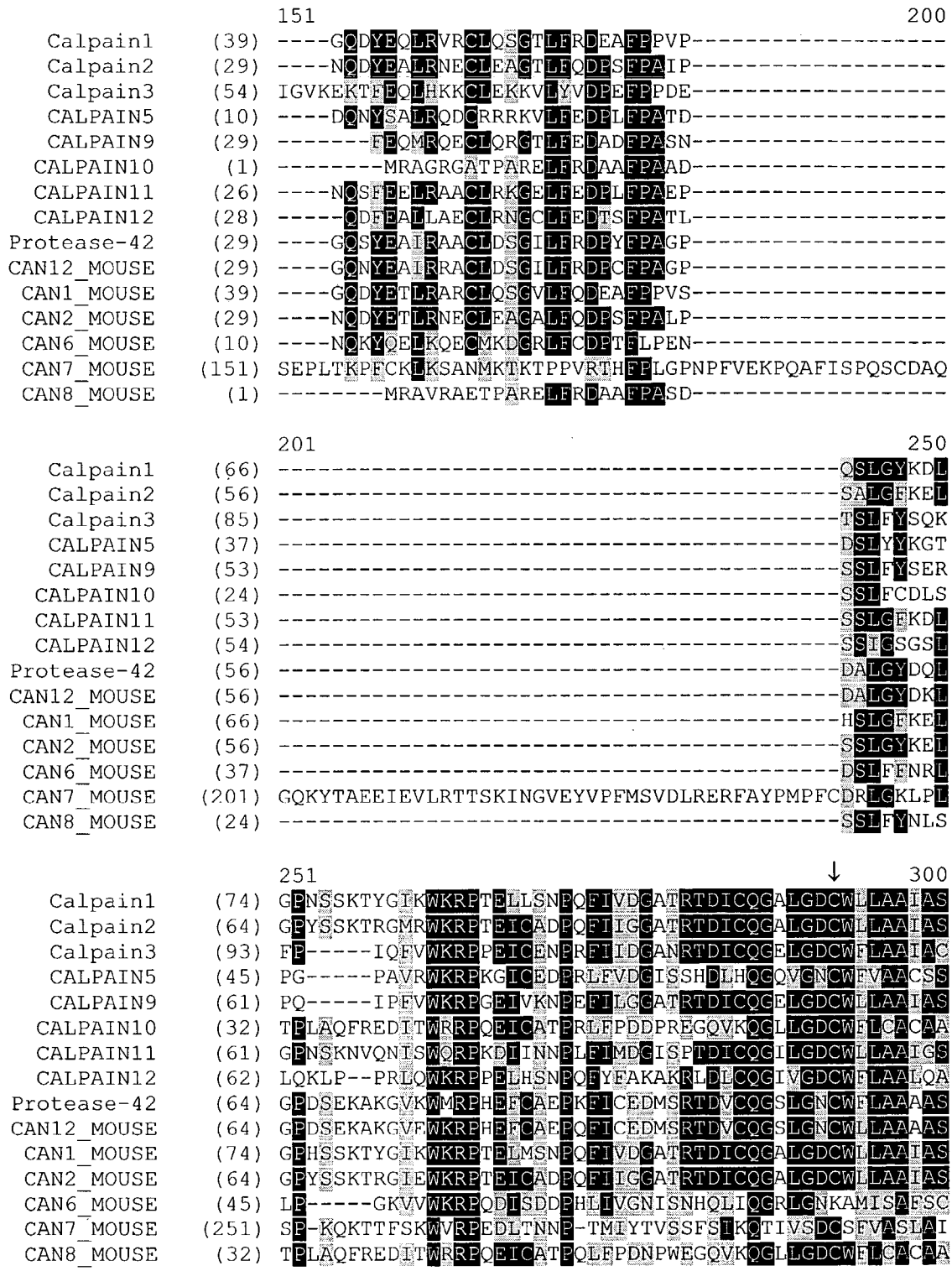

151 200
Calpain1 (39) ----GQDYEQLRVRCLQSGTLFRDEAFPPVP----------
Calpain2 (29) ----NQDYEALRNECLEAGTLFQDPSFPAIP----------
Calpain3 (54) IGVKEKTFEQLHKKCLEKKVLYVDPEFPPDE----------
CALPAIN5 (10) ----DQNYSALRQDCRRRKVLFEDPLFPATD----------
CALPAIN9 (29) -------FEQMRQECLQRGTLFEDADFPASN----------
CALPAIN10 (1) --------MRAGRGATPARELFRDAAFPAAD----------
CALPAIN11 (26) ----NQSFEELRAACLRKGELFEDPLFPAEP----------
CALPAIN12 (28) -----QDFEALLAECLRNGCLFEDTSFPATL----------
Protease-42 (29) ----GQSYEAIRAACLDSGILFRDPYFPAGP----------
CAN12_MOUSE (29) ----GQNYEAIRRACLDSGILFRDPCFPAGP----------
CAN1_MOUSE (39) ----GQDYETLRARCLQSGVLFQDEAFPPVS----------
CAN2_MOUSE (29) ----NQDYETLRNECLEAGALFQDPSFPALP----------
CAN6_MOUSE (10) ----NQKYQELKQECMKDGRLFCDPTFLPEN----------
CAN7_MOUSE (151) SEPLTKPFCKLKSANMKTKTPPVRTHFPLGPNPFVEKPQAFISPQSCDAQ
CAN8_MOUSE (1) --------MRAVRAETPARELFRDAAFPASD----------
201 250
Calpain1 (66) ----------QSLGYKDL
Calpain2 (56) ----------SALGFKEL
Calpain3 (85) ----------TSLFYSQK
CALPAIN5 (37) ----------DSLYYKGT
CALPAIN9 (53) ----------SSLFYSER
CALPAIN10 (24) ----------SSLFCDLS
CALPAIN11 (53) ----------SSLGFKDL
CALPAIN12 (54) ----------SSIGSGSL
Protease-42 (56) ----------DALGYDQL
CAN12_MOUSE (56) ----------DALGYDKL
CAN1_MOUSE (66) ----------HSLGFKEL
CAN2_MOUSE (56) ----------SSLGYKEL
CAN6_MOUSE (37) ----------DSLFFNRL
CAN7_MOUSE (201) GQKYTAEEIEVLRTTSKINGVEYVPFMSVDLRERFAYPMPFCDRLGKLPL
CAN8_MOUSE (24) ----------SSLFYNLS
251 300
Calpain1 (74) GPNSSKTYGIKWKRPTELLSNPQFIVDGATRTDICQGALGDCWLLAAIAS
Calpain2 (64) GPYSSKTRGMRWKRPTEICADPQFIIGGATRTDICQGALGDCWLLAAIAS
Calpain3 (93) FP-----IQFVWKRPPEICENPRFIIDGANRTDICQGELGDCWFLAAIAC
CALPAIN5 (45) PG-----PAVRWKRPKGICEDPRLFVDGISSHDLHQGQVGNCWFVAACSS
CALPAIN9 (61) PQ-----IPFVWKRPGEIVKNPEFILGGATRTDICQGELGDCWLLAAIAS
CALPAIN10 (32) TPLAQFREDITWRRPQEICATPRLFPDDPREGQVKQGLLGDCWFLCACAA
CALPAIN11 (61) GPNSKNVQNISWQRPKDIINNPLFIMDGISPTDICQGILGDCWLLAAIGS
CALPAIN12 (62) LQKLP--PRLQWKRPPELHSNPQFYFAKAKRLDLCQGIVGDCWFLAALQA
Protease-42 (64) GPDSEKAKGVKWMRPHEFCAEPKFICEDMSRTDVCQGSLGNCWFLAAAAS
CAN12_MOUSE (64) GPDSEKAKGVFWKRPHEFCAEPQFICEDMSRTDVCQGSLGNCWLLAAAAS
CAN1_MOUSE (74) GPHSSKTYGIKWKRPTELMSNPQFIVDGATRTDICQGALGDCWLLAAIAS
CAN2_MOUSE (64) GPYSSKTRGIEWKRPTEICADPQFIIGGATRTDICQGALGDCWLLAAIAS
CAN6_MOUSE (45) LP-----GKVVWKRPQDISDDPHLIVGNISNHQLIQGRLGNKAMISAFSC
CAN7_MOUSE (251) SP-KQKTTFSKWVRPEDLTNNP-TMIYTVSSFSIKQTIVSDCSFVASLAI
CAN8_MOUSE (32) TPLAQFREDITWRRPQEICATPQLFPDNPWEGQVKQGLLGDCWFLCACAA

FIG. 2C

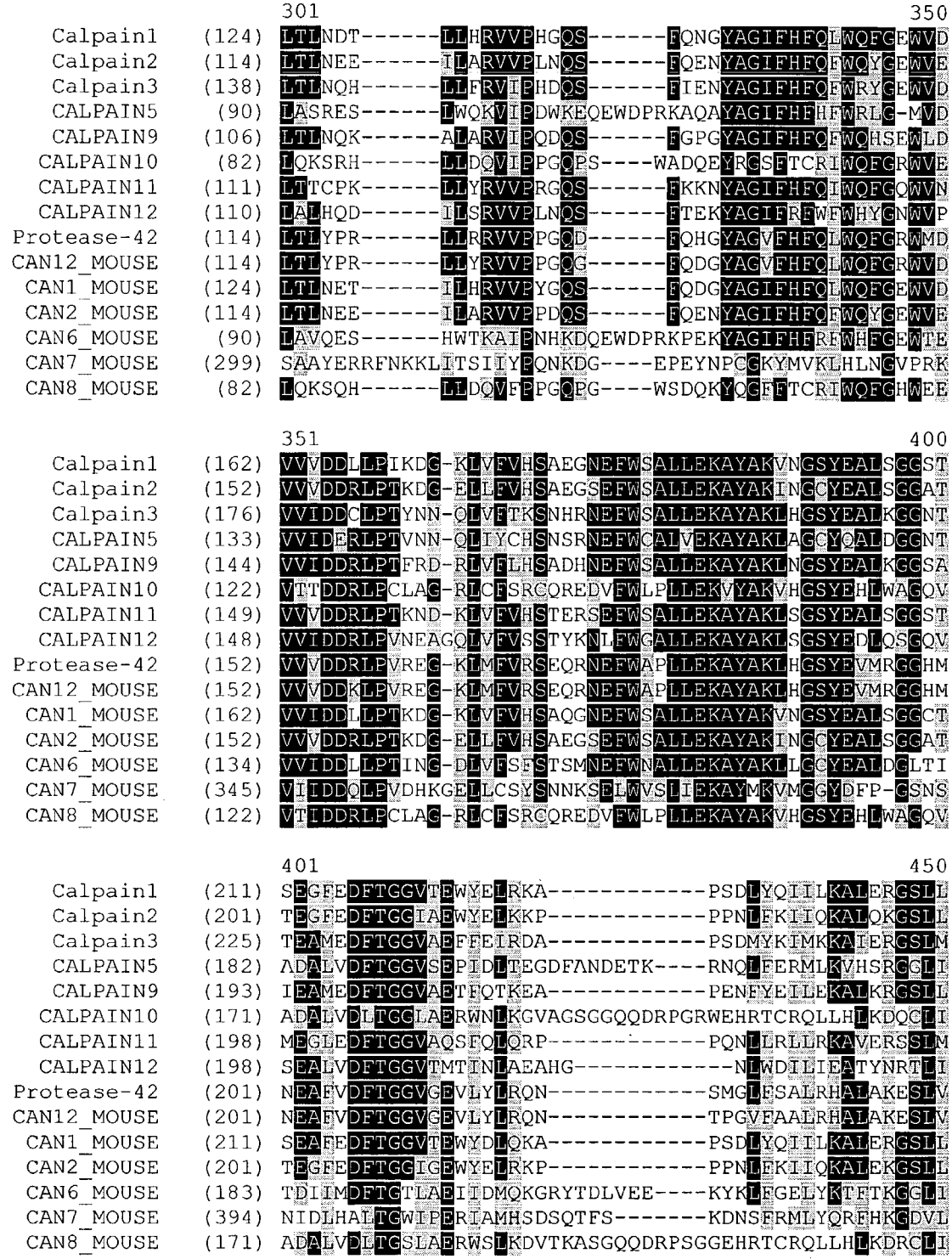

301
350
Calpain1 (124) LTLNDT------LLHRVVPHGQS------FQNGYAGIFHFQLWQFGEWVD
Calpain2 (114) LTLNEE------ILARVVPLNQS------FQENYAGIFHFQFWQYGEWVE
Calpain3 (138) LTLNQH------LLFRVIPHDQS------FIENYAGIFHFQFWRYGEWVD
CALPAIN5 (90) LASRES------LWQKVIPDWKEQEWDPRKAQAYAGIFHFHFWRLG-MVD
CALPAIN9 (106) LTLNQK------ALARVIPQDQS------FGPGYAGIFHFQFWQHSEWLD
CALPAIN10 (82) LQKSRH------LLDQVIPPGQPS----WADQEYRGSFTCRIWQFGRWVE
CALPAIN11 (111) LTTCPK------LLYRVVPRGQS------FKKNYAGIFHFQIWQFGQWVN
CALPAIN12 (110) LALHQD------ILSRVVPLNQS------FTEKYAGIFRFWFWHYGNWVP
Protease-42 (114) LTLYPR------LLRRVVPPGQD------FQHGYAGVFHFQLWQFGRWMD
CAN12_MOUSE (114) LTLYPR------LLYRVVPPGQG------FQDGYAGVFHFQLWQFGRWVD
CAN1_MOUSE (124) LTLNET------ILHRVVPYGQS------FQDGYAGIFHFQLWQFGEWVD
CAN2_MOUSE (114) LTLNEE------ILARVVPPDQS------FQENYAGIFHFQFWQYGEWVE
CAN6_MOUSE (90) LAVQES------HWTKAIPNHKDQEWDPRKPEKYAGIFHFRFWHFGEWTE
CAN7_MOUSE (299) SAAYERRFNKKLITSIIYPQNKDG----EPEYNPCGKYMVKLHLNGVPRK
CAN8_MOUSE (82) LQKSQH------LLDQVFPPGQPG----WSDQKYQGFFTCRIWQFGHWEE
351
400
Calpain1 (162) VVVDDLLPIKDG-KLVFVHSAEGNEFWSALLEKAYAKVNGSYEALSGGST
Calpain2 (152) VVVDDRLPTKDG-ELLFVHSAEGSEFWSALLEKAYAKINGCYEALSGGAT
Calpain3 (176) VVIDDCLPTYNN-QLVFTKSNHRNEFWSALLEKAYAKLHGSYEALKGGNT
CALPAIN5 (133) VVIDERLPTVNN-QLIYCHSNSRNEFWCALVEKAYAKLAGCYQALDGGNT
CALPAIN9 (144) VVIDDRLPTFRD-RLVFLHSADHNEFWSALLEKAYAKLNGSYEALKGGSA
CALPAIN10 (122) VTTDDRLPCLAG-RLCFSRCQREDVFWLPLLEKVYAKVHGSYEHLWAGQV
CALPAIN11 (149) VVVDDRLPTKND-KLVFVHSTERSEFWSALLEKAYAKLSGSYEALSGGST
CALPAIN12 (148) VVIDDRLPVNEAGQLVFVSSTYKNLFWGALLEKAYAKLSGSYEDLQSGQV
Protease-42 (152) VVVDDRLPVREG-KLMFVRSEQRNEFWAPLLEKAYAKLHGSYEVMRGGHM
CAN12_MOUSE (152) VVVDDKLPVREG-KLMFVRSEQRNEFWAPLLEKAYAKLHGSYEVMRGGHM
CAN1_MOUSE (162) VVIDDLLPTKDG-KLVFVHSAQGNEFWSALLEKAYAKVNGSYEALSGGCT
CAN2_MOUSE (152) VVVDDRLPTKDG-ELLFVHSAEGSEFWSALLEKAYAKINGCYEALSGGAT
CAN6_MOUSE (134) VVIDDLLPTING-DLVFSFSTSMNEFWNALLEKAYAKLLGCYEALDGLTI
CAN7_MOUSE (345) VIIDDQLPVDHKGELLCSYSNNKSELWVSLIEKAYMKVMGGYDFP-GSNS
CAN8_MOUSE (122) VTIDDRLPCLAG-RLCFSRCQREDVFWLPLLEKAYAKVHGSYEHLWAGQV
401
450
Calpain1 (211) SEGFEDFTGGVTEWYELRKA----------PSDLYQIILKALERGSLL
Calpain2 (201) TEGFEDFTGGIAEWYELKKP----------PPNLFKIIQKALQKGSLL
Calpain3 (225) TEAMEDFTGGVAEFFEIRDA----------PSDMYKIMKKAIERGSLM
CALPAIN5 (182) ADALVDFTGGVSEPIDLTEGDFANDETK----RNQLFERMLKVHSRGGLI
CALPAIN9 (193) IEAMEDFTGGVAETFQTKEA----------PENFYEILEKALKRGSLL
CALPAIN10 (171) ADALVDLTGGLAERWNLKGVAGSGGQQDRPGRWEHRTCRQLLHLKDQCLI
CALPAIN11 (198) MEGLEDFTGGVAQSFQLQRP----------PQNLLRLLRKAVERSSLM
CALPAIN12 (198) SEALVDFTGGVTMTINLAEAHG----------NLWDILIEATYNRTLI
Protease-42 (201) NEAFVDFTGGVGEVLYLRQN----------SMGLESALRHALAKESLV
CAN12_MOUSE (201) NEAFVDFTGGVGEVLYLRQN----------TPGVFAALRHALAKESLV
CAN1_MOUSE (211) SEAFEDFTGGVTEWYDLQKA----------PSDLYQIILKALERGSLL
CAN2_MOUSE (201) TEGFEDFTGGIGEWYELRKP----------PPNLFKIIQKALEKGSLL
CAN6_MOUSE (183) TDIIMDFTGTLAEIIDMQKGRYTDLVEE----KYKLFGELYKTFTKGGLI
CAN7_MOUSE (394) NIDLHALTGWIPERIAMHSDSQTFS-------KDNSFRMLYQRFHKGDVL
CAN8_MOUSE (171) ADALVDLTGSLAERWSLKDVTKASGQQDRPSGGEHRTCRQLLHLKDRCLI

FIG. 2D

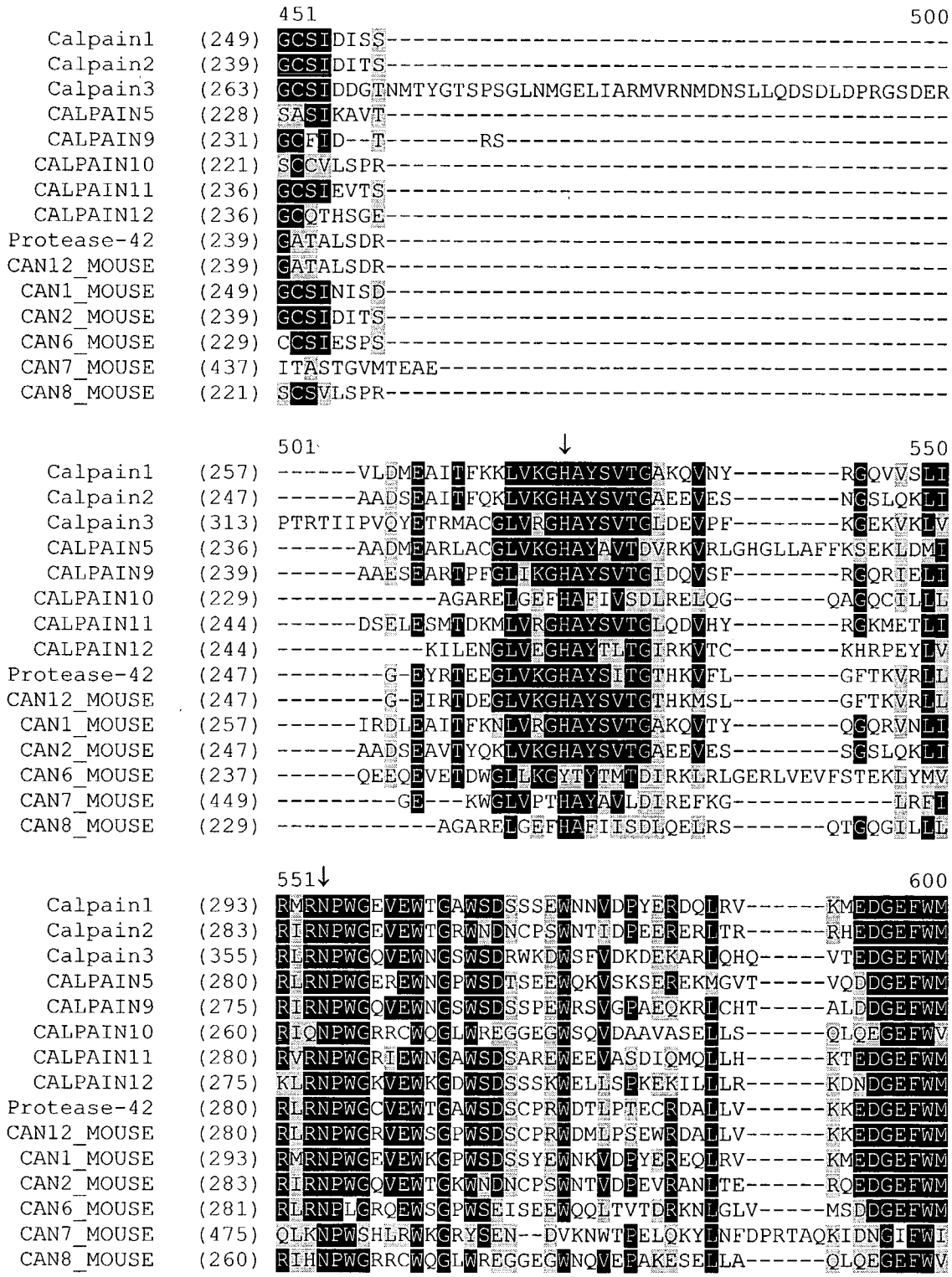

451
500
Calpain1 (249) GCSIDISS----------
Calpain2 (239) GCSIDITS----------
Calpain3 (263) GCSIDDGTNMTYGTSPSGLNMGELIARMVRNMDNSLLQDSDLDPRGSDER
CALPAIN5 (228) SASIKAVT----------
CALPAIN9 (231) GCFID--T-------RS----------
CALPAIN10 (221) SCCVLSPR----------
CALPAIN11 (236) GCSIEVTS----------
CALPAIN12 (236) GCQTHSGE----------
Protease-42 (239) GATALSDR----------
CAN12_MOUSE (239) GATALSDR----------
CAN1_MOUSE (249) GCSINISD----------
CAN2_MOUSE (239) GCSIDITS----------
CAN6_MOUSE (229) CCSIESPS----------
CAN7_MOUSE (437) ITASTGVMTEAE----------
CAN8_MOUSE (221) SCSVLSPR----------
501
550
Calpain1 (257) ------VLDMEAITFKKLVKGHAYSVTGAKQVNY--------RGQVVSLI
Calpain2 (247) ------AADSEAITFQKLVKGHAYSVTGAEEVES--------NGSLQKLI
Calpain3 (313) PTRTIIPVQYETRMACGLVRGHAYSVTGLDEVPF--------KGEKVKLV
CALPAIN5 (236) ------AADMEARLACGLVKGHAYAVTDVRKVRLGHGLLAFFKSEKLDMI
CALPAIN9 (239) ------AAESEARTPFGLIKGHAYSVTGIDQVSF--------RGQRIELI
CALPAIN10 (229) ------------AGARELGEFHAFIVSDLRELQG-------QAGQCILLL
CALPAIN11 (244) ------DSELESMTDKMLVRGHAYSVTGLQDVHY--------RGKMETLI
CALPAIN12 (244) -----------KILENGLVEGHAYTLTGIRKVTC--------KHRPEYLV
Protease-42 (247) --------G-EYRTEEGLVKGHAYSITGTHKVFL--------GFTKVRLL
CAN12_MOUSE (247) --------G-EIRTDEGLVKGHAYSVTGTHKMSL--------GFTKVRLL
CAN1_MOUSE (257) ------IRDLEAITFKNLVRGHAYSVTGAKQVTY--------QGQRVNLI
CAN2_MOUSE (247) ------AADSEAVTYQKLVKGHAYSVTGAEEVES--------SGSLQKLI
CAN6_MOUSE (237) ------QEEQEVETDWGLLKGYTYTMTDIRKLRLGERLVEVFSTEKLYMV
CAN7_MOUSE (449) ---------GE---KWGLVPTHAYAVLDIREFKG-----------LREI
CAN8_MOUSE (229) ------------AGARELGEFHAFIISDLQELRS-------QTGQGILLL
551
600
Calpain1 (293) RMRNPWGEVEWTGAWSDSSSEWNNVDPYERDQLRV------KMEDGEFWM
Calpain2 (283) RIRNPWGEVEWTGRWNDNCPSWNTIDPEERERLTR------RHEDGEFWM
Calpain3 (355) RLRNPWGQVEWNGSWSDRWKDWSFVDKDEKARLQHQ-----VTEDGEFWM
CALPAIN5 (280) RLRNPWGEREWNGPWSDTSEEWQKVSKSEREKMGVT-----VQDDGEFWM
CALPAIN9 (275) RIRNPWGQVEWNGSWSDSSPEWRSVGPAEQKRLCHT-----ALDDGEFWM
CALPAIN10 (260) RIQNPWGRRCWQGLWREGGEGWSQVDAAVASELLS------QLQEGEFWV
CALPAIN11 (280) RVRNPWGRIEWNGAWSDSAREWEEVASDIQMQLLH------KTEDGEFWM
CALPAIN12 (275) KLRNPWGKVEWKGDWSDSSSKWELLSPKEKILLLR------KDNDGEFWM
Protease-42 (280) RLRNPWGCVEWTGAWSDSCPRWDTLPTECRDALLV------KKEDGEFWM
CAN12_MOUSE (280) RLRNPWGRVEWSGPWSDSCPRWDMLPSEWRDALLV------KKEDGEFWM
CAN1_MOUSE (293) RMRNPWGEVEWKGPWSDSSYEWNKVDPYEREQLRV------KMEDGEFWM
CAN2_MOUSE (283) RIRNPWGQVEWTGKWNDNCPSWNTVDPEVRANLTE------RQEDGEFWM
CAN6_MOUSE (281) RLRNPLGRQEWSGPWSEISEEWQQLTVTDRKNLGLV-----MSDDGEFWM
CAN7_MOUSE (475) QLKNPWSHLRWKGRYSEN--DVKNWTPELQKYLNFDPRTAQKIDNGIFWI
CAN8_MOUSE (260) RIHNPWGRRCWQGLWREGGEGWNQVEPAKESELLA------QLQEGEFWV

FIG. 2E

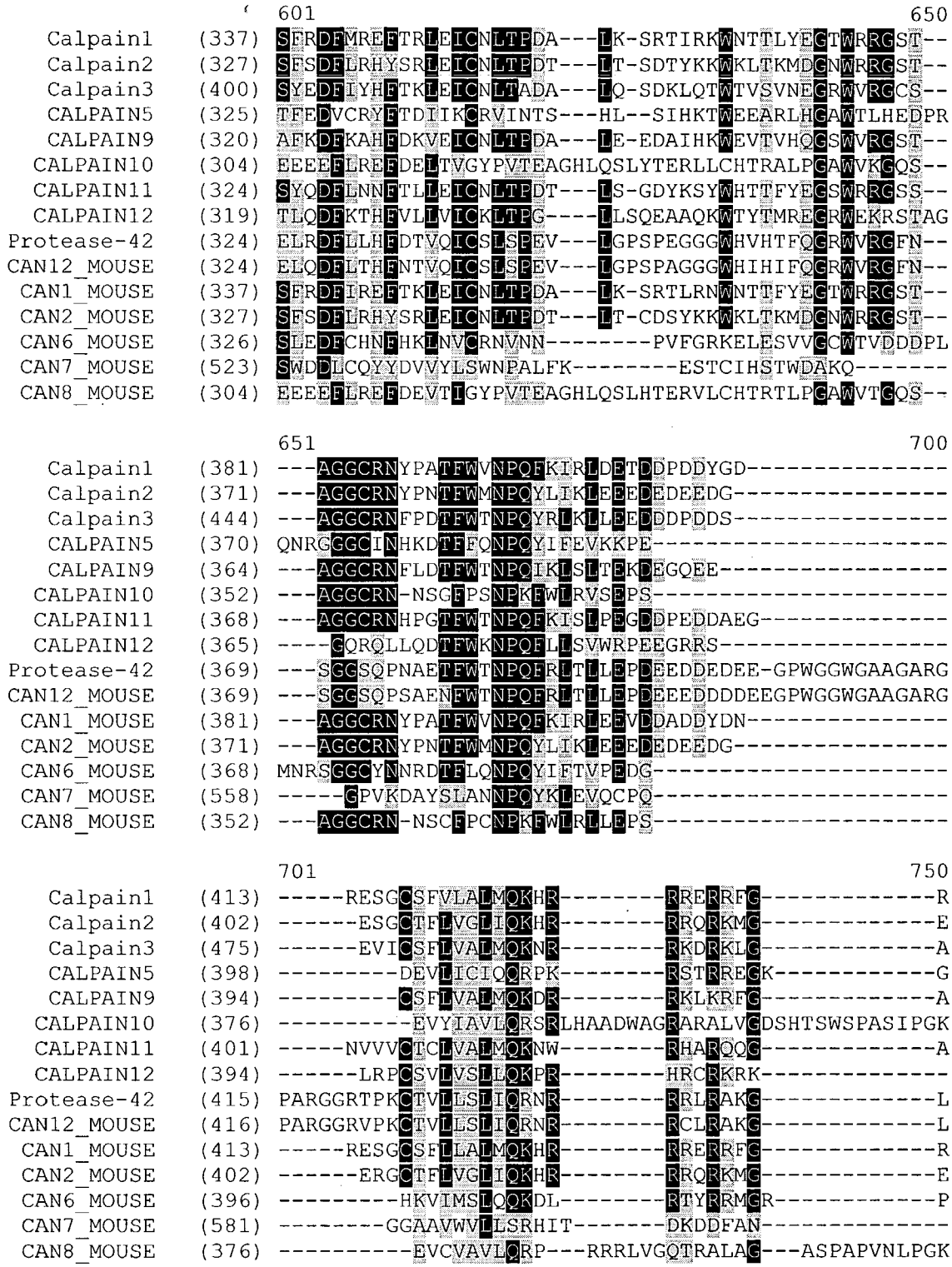

601 650
Calpain1 (337) SFRDFMREFTRLEICNLTPDA---LK-SRTIRKWNTTLYEGTWRRGST--
Calpain2 (327) SFSDFLRHYSRLEICNLTPDT---LT-SDTYKKWKLTKMDGNWRRGST--
Calpain3 (400) SYEDFIYHFTKLEICNLTADA---LQ-SDKLQTWTVSVNEGRWVRGCS--
CALPAIN5 (325) TFEDVCRYFTDIIKCRVINTS---HL--SIHKTWEEARLHGAWTLHEDPR
CALPAIN9 (320) AFKDFKAHFDKVEICNLTPDA---LE-EDAIHKWEVTVHQGSWVRGST--
CALPAIN10 (304) EEEEFLREFDELTVGYPVTEAGHLQSLYTERLLCHTRALPGAWVKGQS--
CALPAIN11 (324) SYQDFLNNFTLLEICNLTPDT---LS-GDYKSYWHTTFYEGSWRRGSS--
CALPAIN12 (319) TLQDFKTHFVLLVICKLTPG----LLSQEAAQKWTYTMREGRWEKRSTAG
Protease-42 (324) ELRDFLLHFDTVQICSLSPEV---LGPSPEGGGWHVHTFQGRWVRGFN--
CAN12_MOUSE (324) ELQDFLTHFNTVQICSLSPEV---LGPSPAGGGWHIHIFQGRWVRGFN--
CAN1_MOUSE (337) SFRDFIREFTKLEICNLTPDA---LK-SRTLRNWNTTFYEGTWRRGST--
CAN2_MOUSE (327) SFSDFLRHYSRLEICNLTPDT---LT-CDSYKKWKLTKMDGNWRRGST--
CAN6_MOUSE (326) SLEDFCHNFHKLNVCRNVNN--------PVFGRKELESVVGCWTVDDDPL
CAN7_MOUSE (523) SWDDLCQYYDVVYLSWNPALFK--------ESTCIHSTWDAKQ-------
CAN8_MOUSE (304) EEEEFLREFDEVTLGYPVTEAGHLQSLHTERVLCHTRTLPGAWVTGQS--
651 700
Calpain1 (381) ---AGGCRNYPATFWVNPQFKIRLDETDDPDDYGD----------
Calpain2 (371) ---AGGCRNYPNTFWMNPQYLIKLEEEDEDEEDG----------
Calpain3 (444) ---AGGCRNFPDTFWTNPQYRLKLLEEDDDPDDS----------
CALPAIN5 (370) QNRGGGCINHKDTFFQNPQYIFEVKKPE----------
CALPAIN9 (364) ---AGGCRNFLDTFWTNPQIKLSLTEKDEGQEE----------
CALPAIN10 (352) ---AGGCRN-NSGFPSNPKFWLRVSEPS----------
CALPAIN11 (368) ---AGGCRNHPGTFWTNPQFKISLPEGDDPEDDAEG----------
CALPAIN12 (365) ----GQRQLLQDTFWKNPQFLLSVWRPEEGRRS----------
Protease-42 (369) ---SGGSQPNAETFWTNPQFRLTLLEPDEEDDEDEE-GPWGGWGAAGARG
CAN12_MOUSE (369) ---SGGSQPSAENFWTNPQFRLTLLEPDEEEDDDDEEGPWGGWGAAGARG
CAN1_MOUSE (381) ---AGGCRNYPATFWVNPQFKIRLEEVDDADDYDN----------
CAN2_MOUSE (371) ---AGGCRNYPNTFWMNPQYLIKLEEEDEDEEDG----------
CAN6_MOUSE (368) MNRSGGCYNNRDTFLQNPQYIFTVPEDG----------
CAN7_MOUSE (558) -----GPVKDAYSLANNPQYKLEVQCPQ----------
CAN8_MOUSE (352) ---AGGCRN-NSCFPCNPKFWLRLLEPS----------
701 750
Calpain1 (413) -----RESGCSFVLALMQKHR--------RRERRFG----------R
Calpain2 (402) ------ESGCTFLVGLIQKHR--------RRQRKMG----------E
Calpain3 (475) ------EVICSFLVALMQKNR--------RKDRKLG----------A
CALPAIN5 (398) ---------DEVLICIQQRPK--------RSTRREGK----------G
CALPAIN9 (394) ---------CSFLVALMQKDR--------RKLKRFG----------A
CALPAIN10 (376) ----------EVYIAVLQRSRLHAADWAGRARALVGDSHTSWSPASIPGK
CALPAIN11 (401) -----NVVVCTCLVALMQKNW--------RHARQQG----------A
CALPAIN12 (394) ------LRPCSVLVSLLQKPR--------HRCRKRK----------
Protease-42 (415) PARGGRTPKCTVLLSLIQRNR--------RRLRAKG----------L
CAN12_MOUSE (416) PARGGRVPKCTVLLSLIQRNR--------RCLRAKG----------L
CAN1_MOUSE (413) -----RESGCSFLLALMQKHR--------RRERRFG----------R
CAN2_MOUSE (402) ------ERGCTFLVGLIQKHR--------RRQRKMG----------E
CAN6_MOUSE (396) ---------HKVLMSLQQKDL--------RTYRRMGR----------P
CAN7_MOUSE (581) --------GGAAVWVLLSRHIT-------DKDDFAN----------
CAN8_MOUSE (376) ----------EVCVAVLQRP---RRRLVGQTRALAG---ASPAPVNLPGK

FIG. 2F

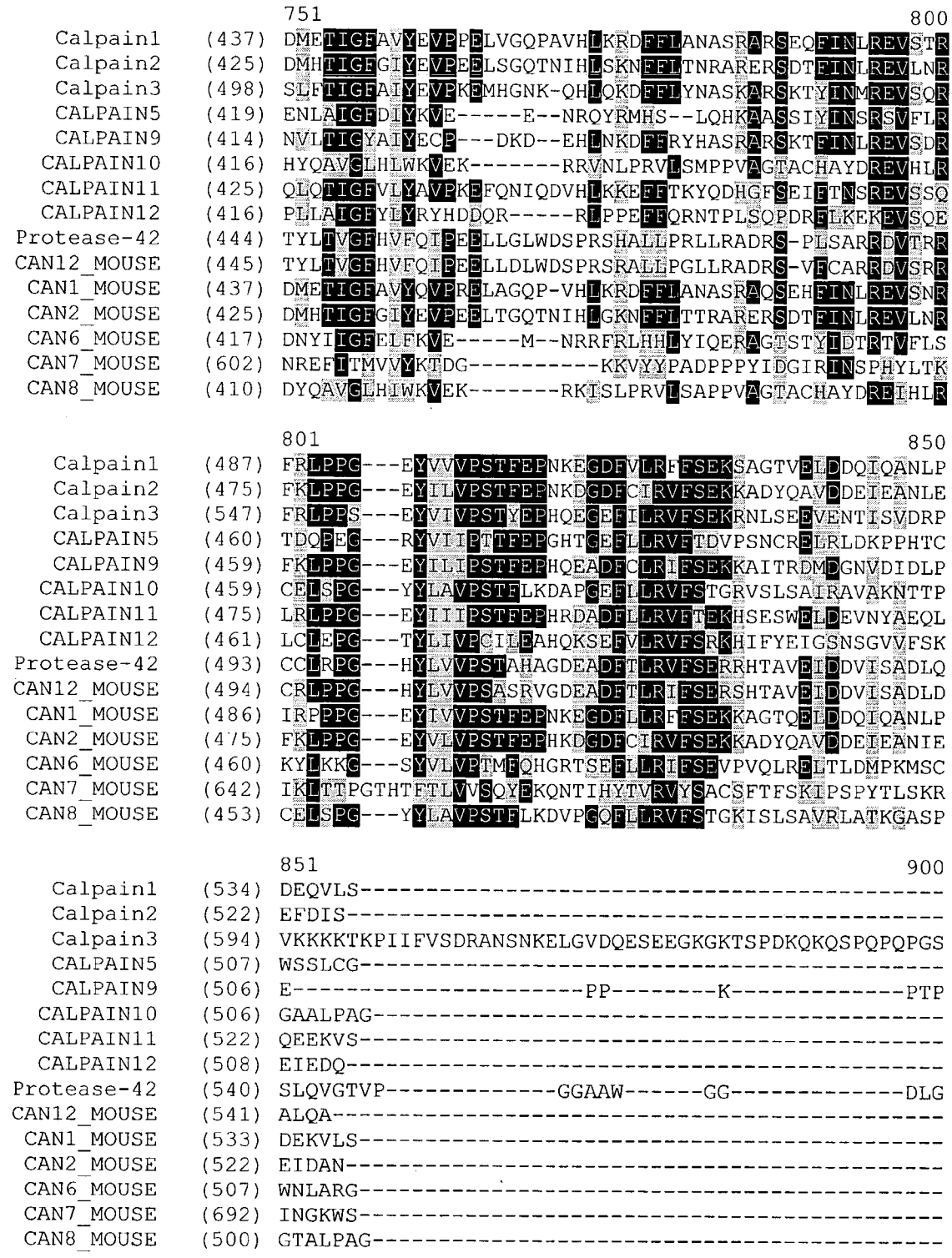

751
800
Calpain1 (437) DMETIGFAVYEVPPELVGQPAVHLKRDFFLANASRARSEQFINLREVSTR
Calpain2 (425) DMHTIGFGIYEVPEELSGQTNIHLSKNFFLTNRARERSDTFINLREVLNR
Calpain3 (498) SLFTIGFAIYEVPKEMHGNK-QHLQKDFFLYNASKARSKTYINMREVSQR
CALPAIN5 (419) ENLAIGFDIYKVE-----E--NRQYRMHS--LQHKAASSIYINSRSVFLR
CALPAIN9 (414) NVLTIGYAIYECP---DKD--EHLNKDFFRYHASRARSKTFINLREVSDR
CALPAIN10 (416) HYQAVGLHLWKVEK-------RRVNLPRVLSMPPVAGTACHAYDREVHLR
CALPAIN11 (425) QLQTIGFVLYAVPKEFQNIQDVHLKKEFFTKYQDHGFSEIFTNSREVSSQ
CALPAIN12 (416) PLLAIGFYLYRYHDDQR-----RLPPEFFQRNTPLSQPDRFLKEKEVSQE
Protease-42 (444) TYLTVGFHVFQIPEELLGLWDSPRSHALLPRLLRADRS-PLSARRDVTRR
CAN12_MOUSE (445) TYLTVGFHVFQIPEELLDLWDSPRSRALLPGLLRADRS-VFCARRDVSRR
CAN1_MOUSE (437) DMETIGFAVYQVPRELAGQP-VHLKRDFFLANASRAQSEHFINLREVSNR
CAN2_MOUSE (425) DMHTIGFGIYEVPEELTGQTNIHLGKNFFLTTRARERSDTFINLREVLNR
CAN6_MOUSE (417) DNYIIGFELFKVE-----M--NRRFRLHHLYIQERAGTSTYIDTRTVFLS
CAN7_MOUSE (602) NREFITMVVYKTDG----------KKVYYPADPPPYIDGIRINSPHYLTK
CAN8_MOUSE (410) DYQAVGLHIWKVEK-------RKISLPRVLSAPPVAGTACHAYDREIHLR
801
850
Calpain1 (487) FRLPPG---EYVVVPSTFEPNKEGDFVLRFFSEKSAGTVELDDQIQANLP
Calpain2 (475) FKLPPG---EYILVPSTFEPNKDGDFCIRVFSEKKADYQAVDDEIEANLE
Calpain3 (547) FRLPPS---EYVIVPSTYEPHQEGEFILRVFSEKRNLSEEVENTISVDRP
CALPAIN5 (460) TDQPEG---RYVIIPTTFEPGHTGEFLLRVFTDVPSNCRELRLDKPPHTC
CALPAIN9 (459) FKLPPG---EYILIPSTFEPHQEADFCLRIFSEKKAITRDMDGNVDIDLP
CALPAIN10 (459) CELSPG---YYLAVPSTFLKDAPGEFLLRVFSTGRVSLSAIRAVAKNTTP
CALPAIN11 (475) LRLPPG---EYIIIPSTFEPHRDADFLLRVFTEKHSESWELDEVNYAEQL
CALPAIN12 (461) LCLEPG---TYLIVPCILEAHQKSEFVLRVFSRKHIFYEIGSNSGVVFSK
Protease-42 (493) CCLRPG---HYLVVPSTAHAGDEADFTLRVFSERRHTAVEIDDVISADLQ
CAN12_MOUSE (494) CRLPPG---HYLVVPSASRVGDEADFTLRIFSERSHTAVEIDDVISADLD
CAN1_MOUSE (486) IRPPPG---EYIVVPSTFEPNKEGDFLLRFFSEKKAGTQELDDQIQANLP
CAN2_MOUSE (475) FKLPPG---EYVLVPSTFEPHKDGDFCIRVFSEKKADYQAVDDEIEANIE
CAN6_MOUSE (460) KYLKKG---SYVLVPTMFQHGRTSEFLLRIFSEVPVQLRELTLDMPKMSC
CAN7_MOUSE (642) IKLTTPGTHTFTLVVSQYEKQNTIHYTVRVYSACSFTFSKIPSPYTLSKR
CAN8_MOUSE (453) CELSPG---YYLAVPSTFLKDVPGQFLLRVFSTGKISLSAVRLATKGASP
851
900
Calpain1 (534) DEQVLS----------
Calpain2 (522) EFDIS----------
Calpain3 (594) VKKKKTKPIIFVSDRANSNKELGVDQESEEGKGKTSPDKQKQSPQPQPGS
CALPAIN5 (507) WSSLCG----------
CALPAIN9 (506) E----------PP--------K----------PTP
CALPAIN10 (506) GAALPAG----------
CALPAIN11 (522) QEEKVS----------
CALPAIN12 (508) EIEDQ----------
Protease-42 (540) SLQVGTVP----------GGAAW------GG----------DLG
CAN12_MOUSE (541) ALQA----------
CAN1_MOUSE (533) DEKVLS----------
CAN2_MOUSE (522) EIDAN----------
CAN6_MOUSE (507) WNLARG----------
CAN7_MOUSE (692) INGKWS----------
CAN8_MOUSE (500) GTALPAG----------

FIG. 2G

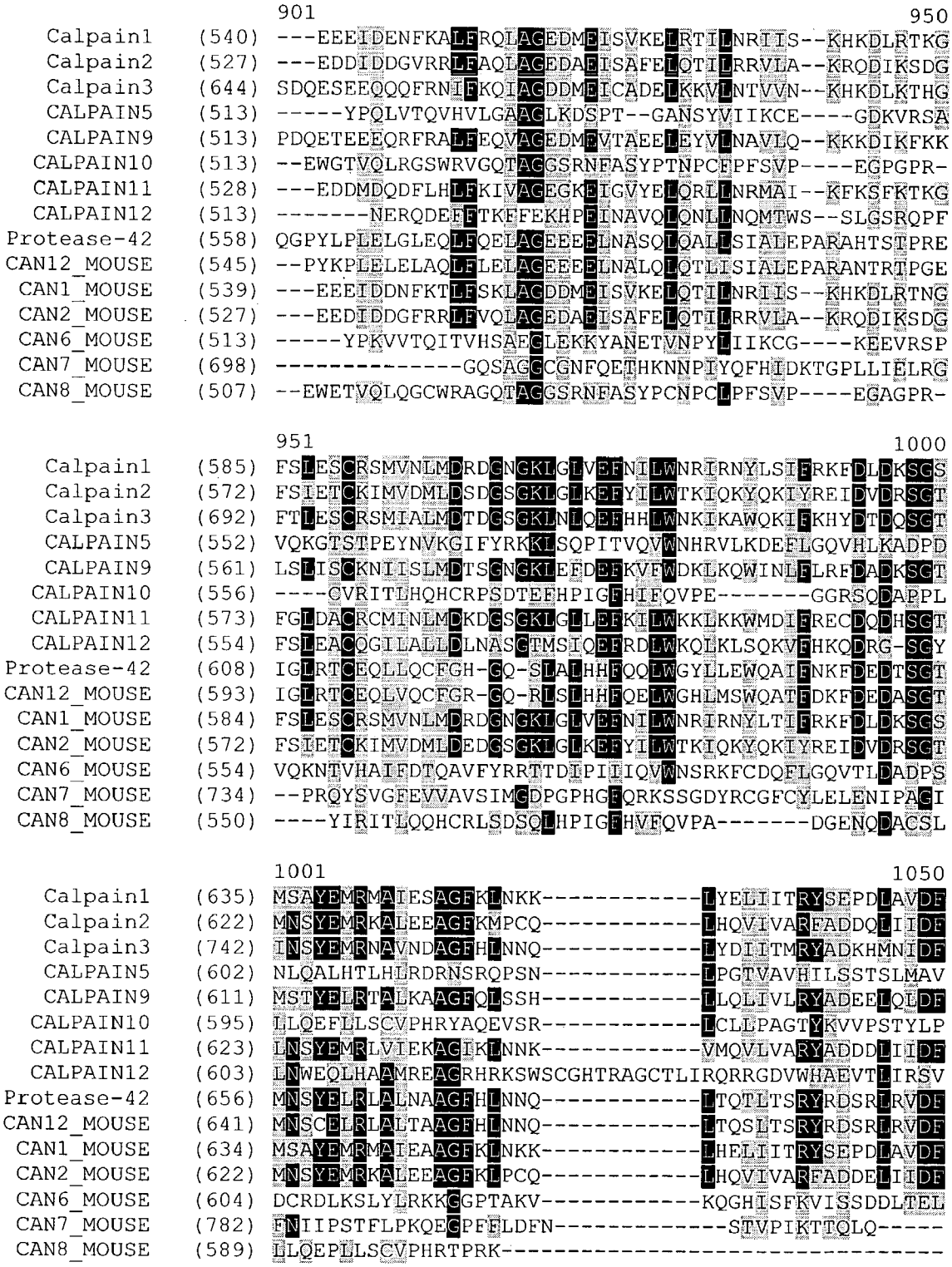
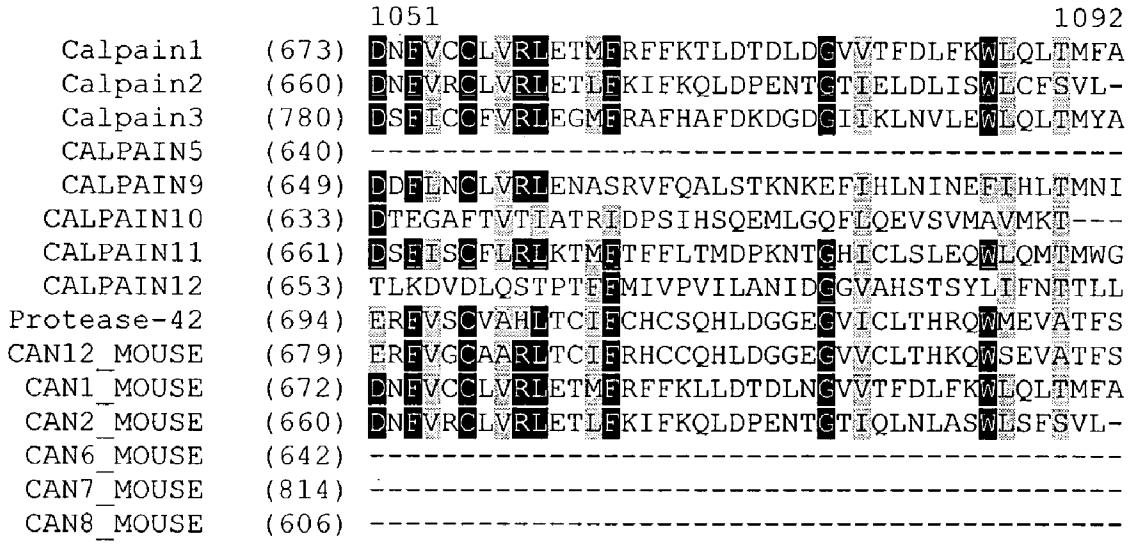

901 950
Calpain1 (540) ---EEEIDENFKALFRQLAGEDMEISVKELRTILNRIIS--KHKDLRTKG
Calpain2 (527) ---EDDIDDGVRRLFAQLAGEDAEISAFELQTILRRVLA--KRQDIKSDG
Calpain3 (644) SDQESEEQQQFRNIFKQIAGDDMEICADELKKVLNTVVN--KHKDLKTHG
CALPAIN5 (513) -----YPQLVTQVHVLGAAGLKDSPT--GANSYVIIKCE----GDKVRSA
CALPAIN9 (513) PDQETEEEQRFRALFEQVAGEDMEVTAEELEYVLNAVLQ--KKKDIKFKK
CALPAIN10 (513) --EWGTVQLRGSWRVGQTAGGSRNFASYPTNPCFPFSVP----EGPGPR-
CALPAIN11 (528) ---EDDMDQDFLHLFKIVAGEGKEIGVYELQRLLNRMAI--KFKSFKTKG
CALPAIN12 (513) -------NERQDEIFTKFFEKHPEINAVQLQNLLNQMTWS--SLGSRQPF
Protease-42 (558) QGPYLPLELGLEQLFQELAGEEEELNASQLQALLSIALEPARAHTSIPRE
CAN12_MOUSE (545) --PYKPLELELAQLFLELAGEEEELNALQLQTLISIALEPARANTRTPGE
CAN1_MOUSE (539) ---EEEIDDNFKTLFSKLAGDDMEISVKELQTILNRIIS--KHKDLRTNG
CAN2_MOUSE (527) ---EEDIDDGFRRLFVQLAGEDAEISAFELQTILRRVLA--KRQDIKSDG
CAN6_MOUSE (513) -----YPKVVTQITVHSAEGLEKKYANETVNPYLIIKCG----KEEVRSP
CAN7_MOUSE (698) --------------GQSAGGCGNFQETHKNNPIYQFHIDKTGPLLIELRG
CAN8_MOUSE (507) --EWETVQLQGCWRAGQTAGGSRNFASYPCNPCLPFSVP----EGAGPR-
951 1000
Calpain1 (585) FSLESCRSMVNLMDRDCNGKLGLVEFNILWNRIRNYLSIFRKFDLDKSGS
Calpain2 (572) FSIETCKIMVDMLDSDGSGKLGLKEFYILWTKIQKYQKIYREIDVDRSGT
Calpain3 (692) FTLESCRSMIALMDTDGSGKLNLQEFHHLWNKIKAWQKIFKHYDTDQSGT
CALPAIN5 (552) VQKGTSTPEYNVKGIFYRKKLSQPITVQVWNHRVLKDEFLGQVHLKADPD
CALPAIN9 (561) LSLISCKNIISLMDTSGNGKLEFDEFKVFWDKLKQWINLFLRFDADKSGT
CALPAIN10 (556) ----CVRITLHQHCRPSDTEFHPIGFHIFQVPE--------GGRSQDAPPL
CALPAIN11 (573) FGLDACRCMINLMDKDGSGKLGLLEFKILWKKLKKWMDIFRECDQDHSGT
CALPAIN12 (554) FSLEACQGILALLDLNASGTMSIQEFRDLWKQLKLSQKVFHKQDRG-SGY
Protease-42 (608) IGLRTCEQLLQCFGH-GQ-SLALHHFQQLWGYLLEWQAIFNKFDEDTSGT
CAN12_MOUSE (593) IGLRTCEQLVQCFGR-GQ-RLSLHHFQELWGHLMSWQATFDKFDEDASGT
CAN1_MOUSE (584) FSLESCRSMVNLMDRDGNGKLGLVEFNILWNRIRNYLTIFRKFDLDKSGS
CAN2_MOUSE (572) FSIETCKIMVDMLDEDGSGKLGLKEFYILWTKIQKYQKIYREIDVDRSGT
CAN6_MOUSE (554) VQKNTVHAIFDTQAVFYRRTTDIPIIIQVWNSRKFCDQFLGQVTLDADPS
CAN7_MOUSE (734) --PRQYSVGFEVVAVSIMGDPGPHGFQRKSSGDYRCGFCYLELENIPAGI
CAN8_MOUSE (550) ----YIRITLQQHCRLSDSQLHPIGFHVFQVPA-------DGENQDACSL
1001 1050
Calpain1 (635) MSAYEMRMAIESAGFKLNKK----------LYELIITRYSEPDLAVDF
Calpain2 (622) MNSYEMRKALEEAGFKMPCQ----------LHQVIVARFADDQLIIDF
Calpain3 (742) INSYEMRNAVNDAGFHLNNQ----------LYDIITMRYADKHMNIDF
CALPAIN5 (602) NLQALHTLHLRDRNSRQPSN----------LPGTVAVHILSSTSLMAV
CALPAIN9 (611) MSTYELRTALKAAGFQLSSH----------LLQLIVLRYADEELQLDF
CALPAIN10 (595) LLQEFLLSCVPHRYAQEVSR----------LCLLPAGTYKVVPSTYLP
CALPAIN11 (623) LNSYEMRLVIEKAGIKLNNK----------VMQVLVARYADDDLIIDF
CALPAIN12 (603) LNWEQLHAAMREAGRHRKSWSCGHTRAGCTLIRQRRGDVWHAEVTLIRSV
Protease-42 (656) MNSYELRLALNAAGFHLNNQ----------LTQTLTSRYRDSRLRVDF
CAN12_MOUSE (641) MNSCELRLALTAAGFHLNNQ----------LTQSLTSRYRDSRLRVDF
CAN1_MOUSE (634) MSAYEMRMAIEAAGFKLNKK----------LHELIITRYSEPDLAVDF
CAN2_MOUSE (622) MNSYEMRKALEEAGFKLPCQ----------LHQVIVARFADDELIIDF
CAN6_MOUSE (604) DCRDLKSLYLRKKGGPTAKV----------KQGHISFKVISSDDLTEL
CAN7_MOUSE (782) FNIIPSTFLPKQEGPFFLDFN----------STVPIKTTQLQ-----
CAN8_MOUSE (589) LLQEPLLSCVPHRTPRK----------

FIG. 2H


1051
1092
Calpain1 (673) DNFVCCLVRLETMFRFFKTLDTDLDGVVTFDLFKWLQLTMFA
Calpain2 (660) DNFVRCLVRLETLFKIFKQLDPENTGTIELDLISWLCFSVL-
Calpain3 (780) DSFICCFVRLEGMFRAFHAFDKDGDGIIKLNVLEWLQLTMYA
CALPAIN5 (640) ----------
CALPAIN9 (649) DDFLNCLVRLENASRVFQALSTKNKEFIHLNINEFIHLTMNI
CALPAIN10 (633) DTEGAFTVTIATRIDPSIHSQEMLGQFLQEVSVMAVMKT---
CALPAIN11 (661) DSFISCFLRLKTMFTFFLTMDPKNTGHICLSLEQWLQMTMWG
CALPAIN12 (653) TLKDVDLQSTPTFFMIVPVILANIDGGVAHSTSYLIFNTTLL
Protease-42 (694) ERFVSCVAHLTCIFCHCSQHLDGGEGVICLTHRQWMEVATFS
CAN12_MOUSE (679) ERFVGCAARLTCIFRHCCQHLDGGEGVVCLTHKQWSEVATFS
CAN1_MOUSE (672) DNFVCCLVRLETMFRFFKLLDTDLNGVVTFDLFKWLQLTMFA
CAN2_MOUSE (660) DNFVRCLVRLETLFKIFKQLDPENTGTIQLNLASWLSFSVL-
CAN6_MOUSE (642) ----------
CAN7_MOUSE (814) ----------
CAN8_MOUSE (606) ----------

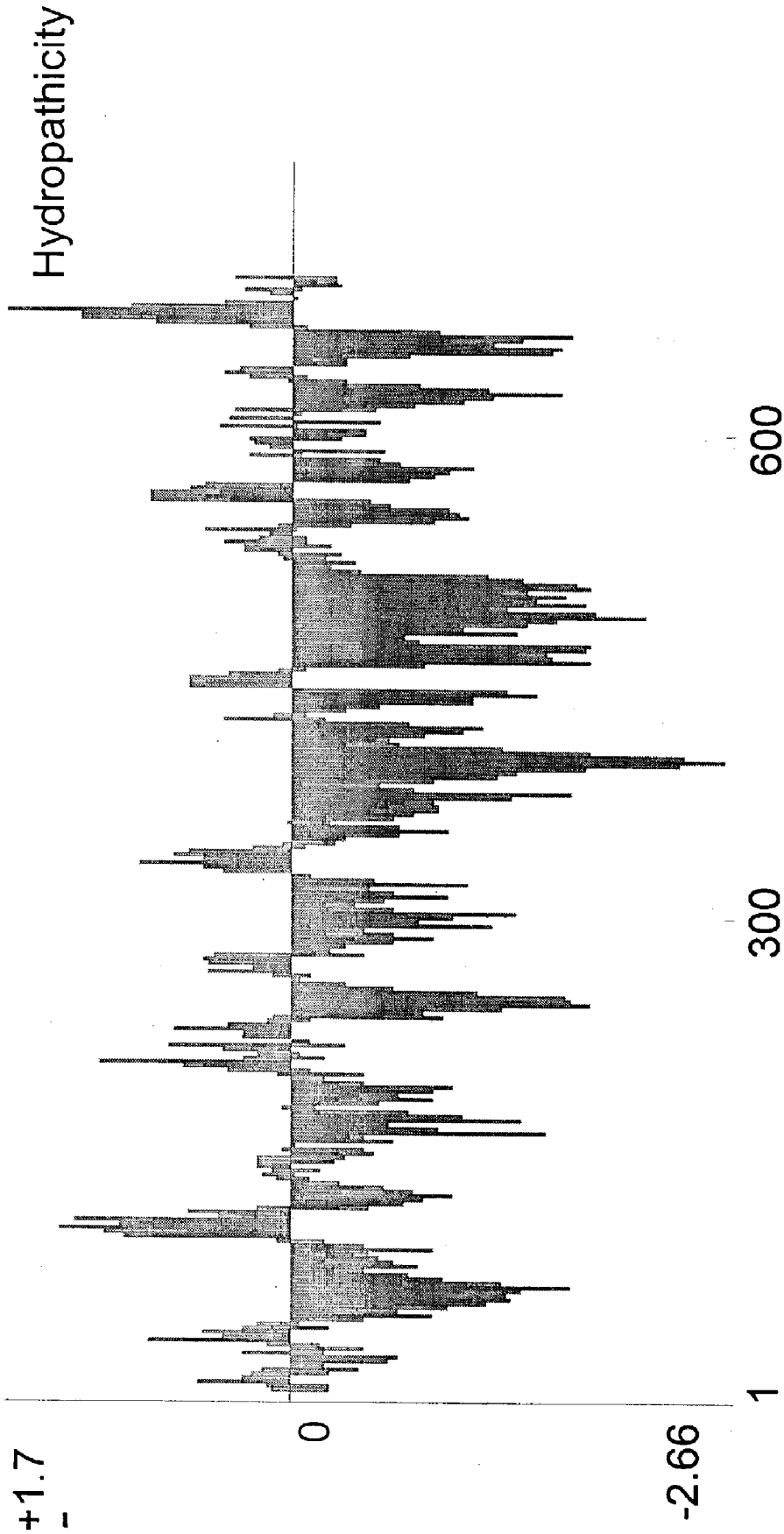


FIG. 3

FIG. 4


Relative Difference
0
10
20
30
40
50
60
70
lymph node
liver
prostate
kidney
heart
lung
spleen
uterus
thymus
bone marrow
brain
small intestine
pancreas
spinal cord
stomach
Tissue RNA Source

FIG. 5

Protein	Genbank ID / Application Serial No.	Identities	Similarities
Human CALPAIN 1 protein	gi|12408656	45.3%	53.8%
human CAN2 protein	gi|4502563	44.4%	55.0%
human calpain 3 protein	gi|4557405	40.5%	51.9%
human CAN5 protein	gi|NP_004046	32.8%	43.2%
human CAN9 protein	gi|5729758	42.6%	52.1%
human CAN10 protein	gi|NP_075574	32.4%	38.6%
human CAN11 protein	gi|NP_008989	42.6%	51.9%
human CAN12 protein	US 60/300,620	37.3%	44.6%
mouse CALPAIN 1 protein	gi|3462902	44.2%	53.5%
mouse CALPAIN 2 protein	gi|2570158	44.8%	55.0%
mouse CALPAIN 6 protein	gi|13959310	32.1%	41.4%
mouse CALPAIN 7 protein	gi|6753258	25.5%	34.3%
mouse CALPAIN 8 protein	gi|5305702	34.4%	40.6%
Mouse CAN12	gi|10303329	86.2%	88.3%

FIG. 6

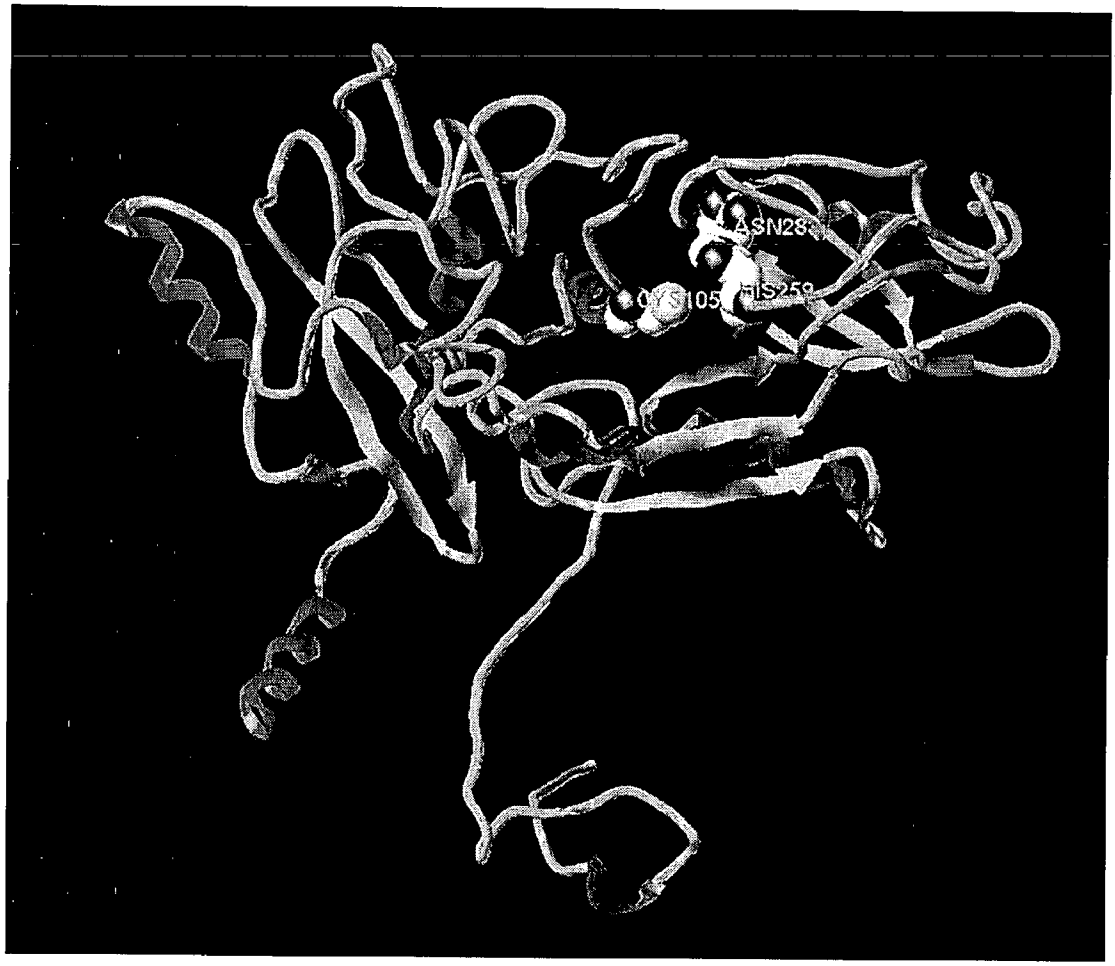


FIG. 7

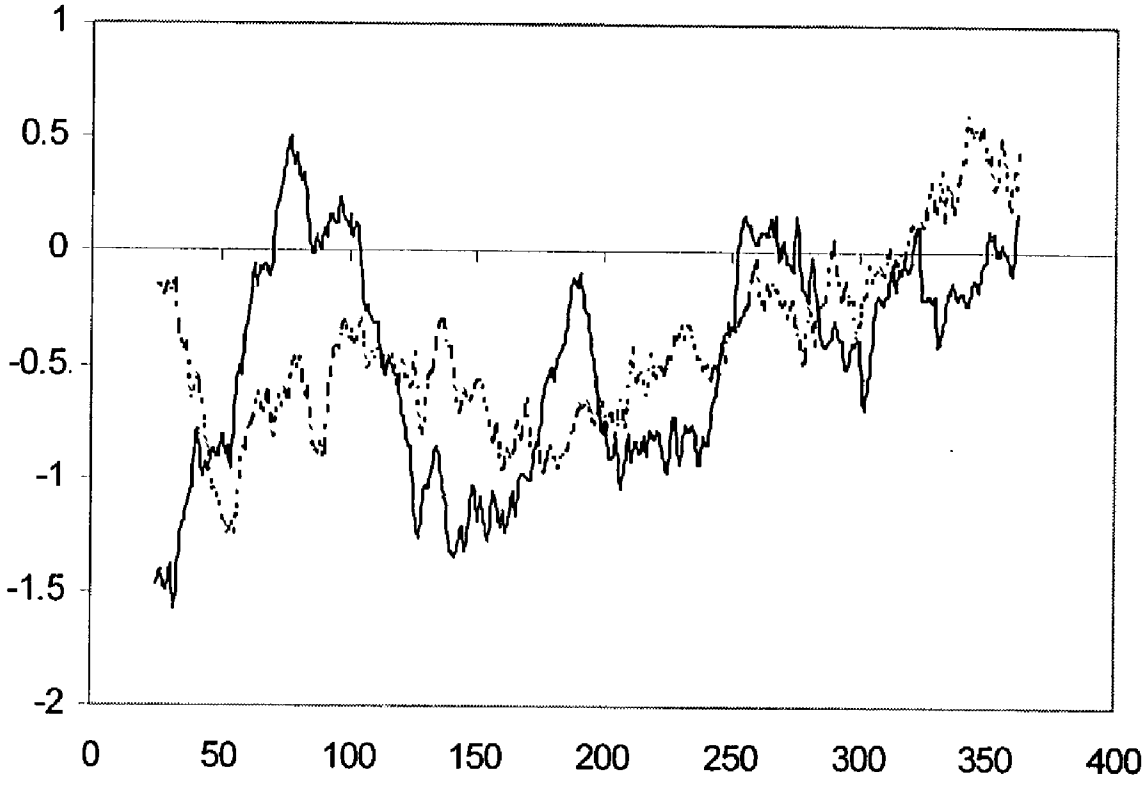


FIG. 8

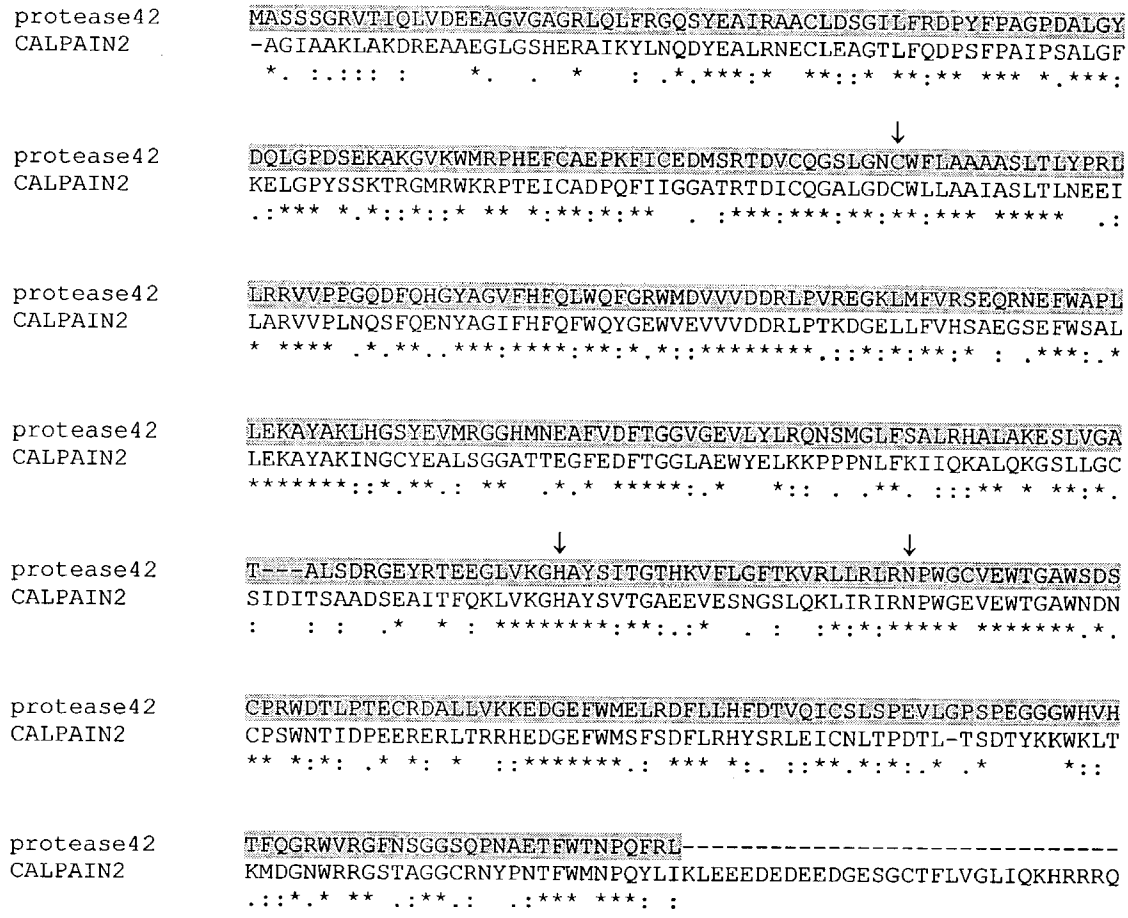

protease42
CALPAIN2
MASSSGRVTIQLVDEEAGVGAGRLQLFRGQSYEAIRAACLDSGILFRDPYFPAGPDALGY
-AGIAAKLAKDREAAEGLGSHERAIKYLNQDYEALRNECLEAGTLFQDPSFPAIPSALGF
protease42
CALPAIN2
DQLGPDSEKAKGVKWMRPHEFCAEPKFICEDMSRTDVCQGSLGNCWFLAAAASLTLYPRL
KELGPYSSKTRGMRWKRPTEICADPQFIIGGATRTDICQGALGDCWLLAAIASLTLNEEI
protease42
CALPAIN2
LRRVVPPGQDFQHGYAGVFHFQLWQFGRWMDVVVDDRLPVREGKLMFVRSEQRNEFWAPL
LARVVPLNQSFQENYAGIFHFQFWQYGEWVEVVVDDRLPTKDGELLFVHSAEGSEFWSAL
protease42
CALPAIN2
LEKAYAKLHGSYEVMRGGHMNEAFVDFTGGVGEVLYLRQNSMGLFSALRHALAKESLVGA
LEKAYAKINGCYEALSGGATTEGFEDFTGGLAEWYELKKPPPNLFKIIQKALQKGSLLGC
protease42
CALPAIN2
T---ALSDRGEYRTEEGLVKGHAYSITGTHKVFLGFTKVRLLRLRNPWGCVEWTGAWSDS
SIDITSAADSEAITFQKLVKGHAYSVTGAEEVESNGSLQKLIRIRNPWGEVEWTGAWNDN
protease42
CALPAIN2
CPRWDTLPTECRDALLVKKEDGEFWMELRDFLLHFDTVQICSLSPEVLGPSPEGGGWHVH
CPSWNTIDPEERERLTRRHEDGEFWMSFSDFLRHYSRLEICNLTPDTL-TSDTYKKWKLT
protease42
CALPAIN2
TFQGRWVRGFNSGGSQPNAETFWTNPQFRL----------
KMDGNWRRGSTAGGCRNYPNTFWMNPQYLIKLEEEDEDEEDGESGCTFLVGLIQKHRRRQ

FIG. 9

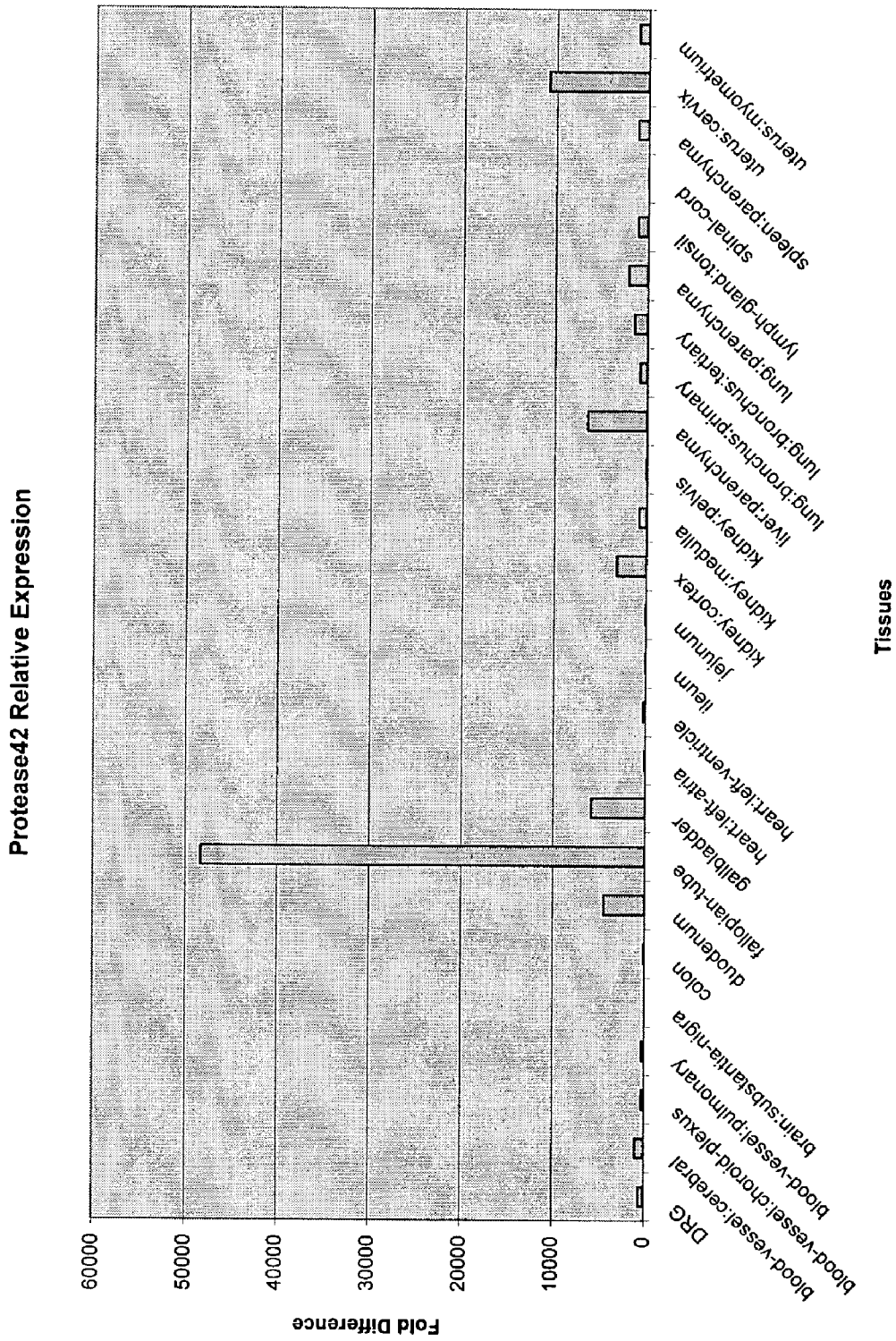
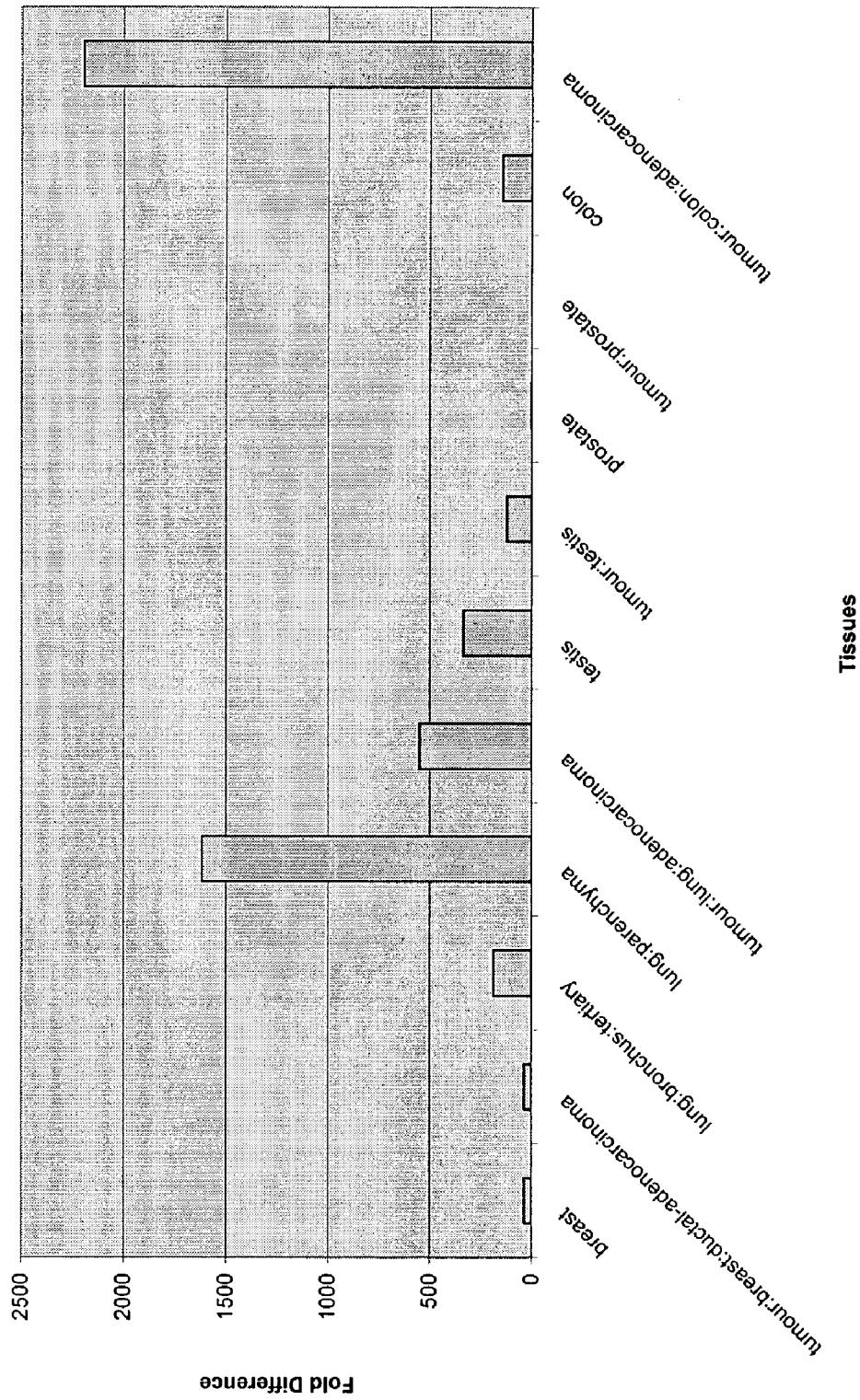

Protease42 Relative Expression
Fold Difference
60000
50000
40000
30000
20000
10000
0
Tissues
DRG
blood-vessel:cerebral
blood-vessel:choroid-plexus
blood-vessel:pulmonary
brain:substantia-nigra
colon
duodenum
fallopian-tube
gallbladder
heart:left-atria
heart:left-ventricle
ileum
jejunum
kidney:cortex
kidney:medulla
kidney:pelvis
liver:parenchyma
lung:bronchus:primary
lung:bronchus:tertiary
lung:parenchyma
lymph-gland:tonsil
spinal-cord
spleen:parenchyma
uterus:cervix
uterus:myometrium

FIG. 10


Protease42 Relative Expression In Various Tumor Types
Fold Difference
2500
2000
1500
1000
500
0
Tissues
breast
tumour:breast:ductal-adenocarcinoma
lung:bronchus:tertiary
lung:parenchyma
tumour:lung:adenocarcinoma
testis
tumour:testis
prostate
tumour:prostate
colon
tumour:colon:adenocarcinoma

POLYNUCLEOTIDE ENCODING A NOVEL CYSTEINE PROTEASE OF THE CALPAIN SUPERFAMILY, PROTEASE-42

[0001] This application is a continuation-in-part application of provisional application U.S. Serial No. 60/364,941, filed Mar. 14, 2002, and claims benefit of the same under 35 U.S.C. 119(e). The entire teachings of the referenced application are incorporated herein by reference.

FIELD OF THE INVENTION

[0002] The present invention provides novel polynucleotides encoding Protease-42 polypeptides, fragments and homologues thereof. Also provided are vectors, host cells, antibodies, and recombinant and synthetic methods for producing said polypeptides. The invention further relates to diagnostic and therapeutic methods for applying these novel Protease-42 polypeptides to the diagnosis, treatment, and/or prevention of various diseases and/or disorders related to these polypeptides. The invention further relates to screening methods for identifying agonists and antagonists of the polynucleotides and polypeptides of the present invention.

BACKGROUND OF THE INVENTION

[0003] Cysteine or thiol proteases contain a reactive sulphydral moiety activated by an adjacent histidine. Hydrolysis of the substrates peptide bond is initiated when the cysteine sulfur attacks the carbon in the peptide bond forming a thiol-enzyme intermediate, liberating the amino portion of the peptide. The thiol-enzyme intermediate is hydrolyzed by water releasing the substrates C-terminus and restoring the enzyme. There are over 20 some families of cysteine proteases. [Rawlings N. D., & Barrett A. J. Families of cysteine peptidases. Methods in Enzymol. 244 461-486 (1994)]. The present invention relates to a thiol protease of the C2 family that includes the calpain superfamily.

[0004] Calpains are calcium-activated intracellular neutral cysteine proteases (EC 3.4.22.17)(for reviews see Sorimachi et al., Structure and physiological function of calpains. Biochem J. 328:721-32, 1997; Carafoli E and Molinari M. Calpain: a protease in search of a function? Biochem Biophys Res Commun 247:193-203, 1998). Some calpains are expressed ubiquitously while others are tissue-specific. μ-Calpain and m-calpains appear in all tissues, p94 is skeletal muscle specific while nCL-2 is stomach specific. (Sorimachi et al., Structure and physiological function of calpains. Biochem J. 328:721-32, 1997). The best characterized are μ-calpain and m-calpains which consist of two subunits. An 80 kDa large subunit contains both Ca^{2+} binding sites and the catalytic activity and small 30 kDa subunit with a separate set of Ca^{2+} binding sites. All the proteolytic activity is contained in the larger subunit of both μ-and m-calpain. In the presence of PEG or chaperones the large subunit is catalytically activated in the absence of the smaller subunit. Other calpains, for example nCL-2 and p94, are proteolytically active monomers with homology to the [μ-calpain and m-calpains large subunit.

[0005] The large (catalytic) subunit has four domains (I-IV)(Hosfield et al., Crystal structure of calpain reveals the structural basis for Ca(2+)-dependent protease activity and a novel mode of enzyme activation. EMBO J. 18:6880-9, 1999; Strobl et al., The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. Proc Natl Acad Sci USA. 97:588-92, 2000). The N-terminus (domain I) contains an alpha helical region and a site of autocatalytic cleavage. Domain II contains the catalytically active domain with the active site amino acids (m-calpain residues Cys105, His262, & Asn286). Domain III contains the linker between the Ca^{2+} binding domain (in domain IV) and links Ca^{2+} binding to proteolytic activity. Domain IV contains a calmodulin-like Ca^{2+} binding regions with EF hands. p94 (also called calpain 3) is similarly organized with domains I-IV, but, also contains a proline-rich N-terminus and two unique insertion loops (IS1 and IS2). nCL-2 is also active as a large monomer with domains I-IV; however, a splice variant (nCL-2') lacks domains III & IV, but maintains proteolytic activity.

[0006] Calpains are responsible for limited intracellular proteolytic cleavage, as opposed to complete proteolytic digestion. The proteolysis modifies protein function both specifically and irreversibly. Numerous proteins have been identified as calpain substrates (Carafoli E and Molinari M Calpain: a protease in search of a function? Biochem Biophys Res Commun 247:193-203, 1998; Hayes et al., Drug News Perspect 11:215-222, 1998). The best-characterized substrates are large cytostructural and/or membrane associated proteins, calmodulin-binding proteins and transcriptional factors. Physiologically significant substrates for calpain include kinases, phosphatases, channel proteins and cytoskeletal proteins that link transmembrane receptors to the membrane skeleton. Proteolytic modification of these proteins may have fundamental roles in development, differentiation, and cellular transformation in response to cell signaling, cell-cell and/or cell-extracellular matrix interactions. In platelets, calpain activation appears to be linked to clustering of the integrin receptor aIIb3 (Fox J E On the role of calpain and Rho proteins in regulating integrin-induced signaling. Thromb Haemost 82:385-91, 1999).

[0007] Calpains have been implicated in cell signaling through activation of protein kinases and phosphatases (cleaving between regulatory and catalytic domains resulting in changes in activity after hydrolysis) and modulation of their intracellular localization. Calpains have been shown to modify specific enzymes and cytoskeletal proteins as part of calcium-mediated signal pathways. They are also involved in remodeling and disassembling the cytoskeleton, especially where the cytoskeleton attaches to membranes or other subcellular structures.

[0008] Several nuclear transcription factors have been suggested as calpain substrates. Calpains are also involved in the progression of cells through the cell cycle (Carafoli E and Molinari M Calpain: a protease in search of a function? Biochem Biophys Res Commun 247:193-203, 1998) in that calpain activity accelerates some cells through the cell cycle by cleavage of p53. Calpain is also thought to play a role in long term potentiation (memory) and rat strains deficient in the endogenous calpain inhibitor, calpastatin, have increased long term potentiation.

Calpains in Disease

[0009] Several diseases have been associated with calpain deficiencies. For example, limb-girdle muscular dystrophy (LGMD) is a group of disorders that primarily cause weakness of the shoulder and pelvic regions. A subtype of LGMD called LGMD2A is caused by defects in the gene for p94

(also called calpain 3)(Richard et al., Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 81:27-40, 1995).

[0010] Positional cloning has recently identified single-nucleotide polymorphisms (SNPs) in an intron of the gene coding for calpain-10 that appears to confer insulin resistance in diabetics. Presence of this mutation correlates with reduced levels of calpain 10 in patients susceptible to type II diabetes (Horikawa et al., Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet. 26:163-75, 2000). The same calpain-10 SNP also correlates with type II diabetes in a high-risk population of Pima Indians (Baier et al., A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest. 106:R69-73, 2000).

Over Activation of Calpains—Ischemic and Traumatic Damage

[0011] Intracellular calcium levels and calpain activity are normally tightly regulated. Under stress, such as follows neuronal excitotoxicity, ischemic stroke, hemoragic stroke, hypoxic stress and/or trauma, intracellular calcium levels rise causing inappropriate calpain proteolytic activity. Calpain activity has been implicated in further cell destruction and non-specific calpain inhibitors have been shown to be protective in animal models (Lee et al., Proc. Natl. Acad. Sci. USA, 88:7233-7237, 1991; Wang K K and Yuen P W. Calpain inhibition: an overview of its therapeutic potential. Trends Pharmacol. Sci. 15:412-9, 1994; Lee, K S, et al., Calcium-activated proteolysis as a therapeutic target in cerebrovascular disease. Annal NY Acad. Sci. 825, 95-103, 1997).

[0012] Calpains are activated in neurons following ischemia-induced damage in animal models of stroke. (Lee et al., Proc. Natl. Acad. Sci. USA, 88:7233-7237, 1991). Inhibition of calcium-activated proteolysis by means of high doses of (usually non-specific) calpain inhibitors protect against the degeneration of vulnerable hippocampal neurons after ischemia (Rami et al., Brain Research, 609:67-70, 1993; Wang et al., An alpha-mercaptoacrylic acid derivative is a selective nonpeptide cell-permeable calpain inhibitor and is neuroprotective. Proc Natl Acad Sci USA. 93:6687-92, 1996). After an ischemic insult, neuronal death is delayed for hours to days. This time interval represents a potential therapeutic window in which to apply effective therapies to minimize brain damage after stroke.

[0013] In addition to neuronal damage, calpains are thought to contribute to cardiac ischemic damage (Iwamoto H et al., Calpain inhibitor-1 reduces infarct size and DNA fragmentation of myocardium in ischemic/reperfused rat heart. J Cardiovasc Pharmacol 33:580-6, 1999) and hepatocyte necrosis during and following anoxia (Arora A S et al., Hepatocellular carcinoma cells resist necrosis during anoxia by preventing phospholipase-mediated calpain activation. J Cell Physiol 167:434-42, 1996).

[0014] Recently, McDonald, et al., provided evidence that hemorrhage and resuscitation with shed blood resulted in an increase in calpain activity (heart), activation of NF-kB (kidney), expression of iNOS and COX-2 (kidney), and the development of multiple organ injury and dysfunction, all of which were attenuated by calpain inhibitor I (10 mg/kg i.p.), administered 30 min prior to hemorrhage in rat (McDonald, M C et al., 2001).

[0015] Those studies indicate the potential utility of calpain inhibitors (especially those calpains expressed in lung, kidney, liver, pancreas, digestive track) in treating ischemia/reperfusion injury.

Calpains in Neurodegenerative Diseases

[0016] Calpains have been implicated in neurodegenerative diseases ncluding, Alzheimer's disease, Multiple sclerosis, Huntington's disease, Parkinson's disease and amyotrophy. Calpain activation is increased during normal aging and a strong case can be made for the involvement of calpain in the abnormal proteolysis underlying the accumulation of plaque and neurofibriles in brain tissue from people who suffered Alzheimer-type dementia (Iwamoto et al., Brain Research, 561:177-180 1991; Nixon et al., Calcium-activated neutral proteinase (calpain) system in aging and Alzheimer's disease. Ann NY Acad Sci;747:77-91, 1994; Grynspan et al., Active site-directed antibodies identify calpain II as an early-appearing and pervasive component of neurofibrillary pathology in Alzheimer's disease. Brain Res 763:145-58, 1997). Calpains are significantly activated in human postmortem brain from patients with Alzheimer's disease, and the degree of activation correlated with those regions of the brain showing the greatest amount of degeneration (Saito et al., Proc. Natl. Acad. Sci. USA, 90:2628-2632, 1993). More recently, it has been recognized that in Alzheimer's disease cyclin-dependent kinase 5 (cdk5) and its neuron-specific activator p35 are involved in neurite outgrowth and cortical lamination. Calpain cleavage of p35 produces p25, which accumulates in the brains of patients with Alzheimers disease. Conversion of p35 to p25 causes prolonged activation and mislocalization of cdk5 which hyperphosphorylates tau, disrupts the cytoskeleton and promotes the death (apoptosis) of primary neurons (Lee et al., Neurotoxicity induces cleavage of p35 to p25 by calpain. Nature. 18;405:360-4, 2000). Compounds that inhibit calpain activity could prove useful in reducing or delaying neurodegeneration caused to Alzheimer's disease.

Calpains in Damage Following Trauma

[0017] Traumatic injury also causes calpain activation associated with further cell death, atrophy and shrinkage of the brain. A forceful blow trigger cell damage and increased calpain activity that can cleave structural proteins in the brain for up to weeks afterward (Hayes et al., Potential Contribution of Proteases to Neuronal Damage Drug News & Perspectives 11, 1998).

[0018] Calpain activation has also been implicated in spinal cord injury following trauma (for reviews see: Banik et al., Role of calpain and its inhibitors in tissue degeneration and neuroprotection in spinal cord injury. Ann. N.Y. Acad. Sci. 825:120-7 1997; Banik et al., Role of calpain in spinal cord injury: effects of calpain and free radical inhibitors. Ann. N.Y. Acad. Sci. 844:131-7, 1998). Analogous to brain trauma, secondary pathophysiological alterations occur in the traumatized spinal cord well after the initiating insult. These secondary events ultimately cause cell death and tissue damage. Non-specific calpain inhibitors have shown utility in preventing further damage due to spinal chord

injury in animal models (Ray et al., Increased calpain expression is associated with apoptosis in rat spinal cord injury: calpain inhibitor provides neuroprotection. Neurochem Res. 25:1191-8, 2000).

[0019] These studies indicate the potential utility of calpain inhibitors (especially those calpains located in the spinal cord) in treating traumatic injury resulting from automobile crashes, gunshot wounds, and sports accidents.

Calpains in Degeneration of Cochlear Tissues Following Noise Exposure

[0020] Calpains are activated during acoustic trauma and calpain inhibitors protect against hearing loss caused by noise (Stracher A Calpain inhibitors as therapeutic agents in nerve and muscle degeneration. Ann NY Acad Sci 884:52-9, 1999).

Calpains in Inflammation

[0021] Calpains also regulate integrin-mediated interaction of T-cells with the extracellular matrix (ECM) and calpain inhibitors prevent acute and chronic inflammation in animal models (Cuzzocrea S et al., Calpain inhibitor I reduces the development of acute and chronic inflammation Am J Pathol 157:2065-79, 2000). In human models of allergic inflammation, the nuclear localization of the transcription factor nuclear factor (NK)-kappa B, which binds to and affects the function of several genes encoding proteins mediating inflammation can be suppressed by calpain inhibitor or calpastatin (Wilson S J et al., 1999). In human bronchial epithelial cells, in which calpain was constitutively inhibited by the overexpression of calpastatin, there was a reduced basal and induced IkappaBalpha degradation and NF-kappaB activation (Chen, F et al., 2000). In calpain 3-deficient mouse, affected muscles manifest a similar perturbation of the IkappaBalpha/nuclear factor kappaB pathway as seen in LGMD2A patients (Richard I, et al., 2000). Recently, calpain 3 dependent IkappaBalpha degradation was reconstituted in vitro supporting a possible in vivo sequence of events in skeletal muscle of LGMD2A patients leading to IkappaBalpha accumulation, prevention of nuclear translocation of NF-kappaB, and ultimately apoptosis (Baghdiguian J et al., 1999). In addition, calpain inhibitors have been shown to significantly reduced degree of colon injury, rise in myeloperoxidase activity (mucosa) as well as upregulation of ICAM-1 and P-selectin in an animal model of inflammatory bowel disease (Cuzzocrea S et al., 2001).

Calpains in Multiple Sclerosis

[0022] Multiple sclerosis is characterized by the progressive loss of the myelin of the brain and spinal cord. In autoimmune demyelinating diseases such as multiple sclerosis and experimental allergic encephalomyelitis, the degradation of myelin proteins results in the destabilization of the myelin sheath. Calpains have been implicated in that calpain degrades all major myelin proteins and increased calpain activity is observed in multiple sclerosis (Shields D C et al., A putative mechanism of demyelination in multiple sclerosis by a proteolytic enzyme, calpain. Proc. Natl. Acad. Sci. USA 96:11486-91, 1999).

Calpains in Cataract Formation

[0023] In the lens, crystallins prevent thermal denaturation and aggregation of other proteins. Crystallins are also substrates for calpains and cataract formation in a rat model of selenite-induced cataract formation is thought to result from calpain activation and cleavage of crystallins (Shearer T R, David L L, Anderson R S, Azuma M. Review of selenite cataract. Curr Eye Res 1992; 11:357-369). In this model the crystallin cleavage could be blocked by calpain inhibitors (Azuma M et al., Cysteine protease inhibitor E64 reduces the rate of formation of selenite cataract in the whole animal. Curr Eye Res 10:657-666, 1991). In a genetic model cataract-prone rats also showed enhanced proteolysis of crystallins and lens cytoskeletin proteins thought to be mediated by calpain (Inomata M et al., Evidence for the involvement of calpain in cataractogenesis in Shumiya cataract rat (SCR). Biochim Biophys Acta 1362:11-23 1997). Calpain activation is also thought to play a role in cataracts induced by buthionine sulfoximine, calcium ionophore A23187, hydrogen peroxide, diamide, xylose, galactose and streptozotocin (Kadoya et al., Role of calpain in hydrogen peroxide cataract. Curr Eye Res 1993; 12:341-346; David et al., Buthionine sulfoximine induced cataracts in mice contain insolubilized crystallins with calpain II cleavage sites, Exp Eye Res 1994; 59:501-504.). These models of cataract formation in rats suggest that calpain-induced proteolysis is a common underlying mechanism. Fragments of alpha-crystallin, consistent with calpain cleavage, have been also observed in cataractous human lens.

Calpains in Reovirus Induced Myocarditis

[0024] Infection of neonatal mice with reovirus produces histological myocarditis. This is due to a direct viral injury and apoptosis of myocytes. Calpain inhibitors block reovirus-induced apoptosis in vitro and prevented viral-induced induced myocarditis (DeBiasi et al., Calpain inhibition protects against virus-induced apoptotic myocardial injury. Virol 75:351-61, 2001).

Calpains in Cancers

[0025] There is growing body of literature that implicates the role of calpain in various aspect of carcinogenesis, including cell-cycle progression, cellular differentiation and apoptosis (Wang, K. K, 2000). Many products of oncogenes and tumor suppressor genes (i.e. c-fos, c-jun, p53, pp60src, estrogen receptor, integrin) are substrates of calpains (Liu, K et al., 2000). Association between abnormal calpain activity and tumorigenesis has been observed in several studies. For example, calpain-I expression is correlated with increased malignancy in renal cell carcinoma (Brau, C et al., 1999). Recently, Yoshikawa, et al., reported that expression of calpain 9 was downregulated in gastric cancer tissues and cell lines of both differentiated and poorly differentiated type (Yoshikawa, Y et al., 2000). Independently, Liu et al., showed that depletion of calpain 9 mRNA in NIH3T3 fibroblast cells resulted in cellular transformation and tumorigenesis (Liu, K et al., 2000). Together these studies suggest that calpain 9 might be acting as a tumor suppressor through proteolytic degradation of digestive track specific oncogenes.

Calpains in Hair Growth

[0026] Dear et al., identified a novel calpain, calpain 12 expressed only at high levels in mouse skin. Expression of calpain 12 was localized to the cortex of the pelage hair follicle. Calpain 12 mRNA reached its highest levels at the

midpoint of the anagen phase (proliferating phase) but were absent from the latter stage of the hair cycle (telogen phase). Most common forms of hair loss (alopecia) are caused by aberrant hair follicle cycling and changes in hair follicle morphology. However, current treatments do not specifically target these processes. The end product of hair follicle proliferation and differentiation is the hair shaft, which together with its surrounding root sheaths, is derived from epithelial cells. The size and the length of the hair shafts correspond to the size of the hair follicle and to the duration of anagen, respectively.

[0027] The major goals in the treatment of alopecia include prolonging anagen, converting telogen follicles to anagen, and possibly generating new follicles. There have been recent dramatic advances in our understanding of the molecules and pathways regulating hair follicle formation and hair growth. In particular, p53 has been shown to be involved in the development of chemotherapy-induced alopecia, as p53 knockout mice treated with chemotherapeutic agents remarkably do not lose their hair (Botchkarev et al., 2000). Interestingly, P53 is a known substrate of calpain (Benetti R, et al). The death substrate Gas2 binds m-calpain and increases susceptibility to p53-dependent apoptosis (EMBO J. Jun. 1, 2001;20(11):2702-14). A specific calpain agonist could possibly prevent hair loss by cleavage of p53, promoting hair growth.

[0028] The inventors of the present invention describe herein, the polynucleotides corresponding to the full-length novel Protease-42 calpain, and its encoded polypeptide. Also provided are polypeptide alignments illustrating the strong conservation of the Protease-42 polypeptides to known proteases and a model of the active conformation of Protease-42. Based on this strong conservation, the inventors have ascribed the Protease-42 polypeptides as having calpain proteolytic activity. Data is also provided illustrating the unique tissue expression profile of the Protease-42 polypeptide in brain, liver, spleen, lung, kidney, and in digestive track tissues, which has not been appreciated heretofore.

[0029] In fact, calpains have been the subject of significant research and development programs designed to identify inhibitors of this disease associated protein class. For example, the following, non-limiting examples of drugs, therapies, or regimens directed to inhibiting calpains are currently known: BDA 410 (Mitsubishi Tokyo); AK 295 (Alkermes; CAS® Registry Number: 160399-35-9, 144231-82-3, and 145731-49-3; (1-(((1-ethyl-3-((3-(4-morpholinyl-)propyl)amino)-2,3-dioxopropyl)amino)carbonyl)-3-methylbutyl)carbamic acid phenylmethyl ester stereois); AK 275 (Alkermes; CAS® Registry Number: 158798-83-5, and 150519-08-7; N-((phenylmethoxy)carbonyl)-L-leucyl-N-ethyl-L-2-aminobutanamide); inhibitor I (University of Indiana; acetyl-leu-leu-norleucinal); calpeptin (University of Indiana; benzyloxycarbonyl-leu-norleucinal); VASOLEX (Cortex); RESTENEX (Cortex); MDL 28170 (Aventis; CBZ-Val-Phe-H); P1 (Sankyo; CAS® Registry Number: 128102-74-9, and 128102-75-0; L-phenylalanyl-L-glutaminyl-L-valyl-L-valyl-3-((3-nitro-2-pyridiny 1)dithio)-L-alanylglycinamide); MDL 28170 (Hoechst Marion Roussel); BDA-410 (Mitsubishi-Tokyo); SJA-6017 (Senju; CAS® Registry Number: 190274-53-4; Butanamide,2-(((4-fluorophenyl)sulfonyl)amino)-N-((1S)-1-formyl-3-methylbutyl.).-3-methyl-, (2S)—); Pharmaprojects No. 5123 (Pfizer; 2-Chloro-acetic acid(3-oxo-4-phenyl-3,4-dihydro-1H-quinoxalin-2-ylidene)hydrazide; WO96-25403); CEP-4143 (Cephalon; WO96-14067); MDL-104903 (Aventis; CAS® Registry Number: 180799-56-8; Carbamic acid,((1S)-1-(((4S,5R)-5-hydroxy-4-(phenylmethyl)-3-oxazolidinyl)carbonyl)-2-methylpropyl)-,phenylmethyl ester)); MDL-28170 (Aventis; CAS® Registry Number: 19542-51-9; Alanine, N-(N-carboxy-L-valyl)-3-phenyl-N-benzyl ester, L-); CX-275 (Cortex Pharmaceuticals; Phenylmethy1N-((1R)-1-((((1S)-1-ethyl-3-(ethylamino)-2,3-dioxopropyl)amino)carbonyl)-3-methylbutyl)carbamate); NS 7 (Nippon Shinyaku; 4-(4-Fluorophenyl)-2-methyl-6-(5-piperidinopentyloxy)pyrimidine hydrochloride); Calpain inhibitor 1 (Suntory; N-Acetyl-L-leucinyl-L-leucinyl-L-norleucinal); E 64 (Taisho Pharmaceutical (; CAS® Registry Number: 66701-25-5); and CEP 4143 (Cephalon); SJA 6017 (Senju; N-(4-Fluorophenylsulfonyl)-L-valyl-L-leucinal). The present invention is directed to antagonists specific to the Protease-42 polypeptides. Modulating the activity of the calpain polypeptides of the present invention may result in fewer toxicities and better efficacy than the drugs, therapies, or regimens presently known to regulate other known calpains.

[0030] The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells, in addition to their use in the production of Protease-42 polypeptides or peptides using recombinant techniques. Synthetic methods for producing the polypeptides and polynucleotides of the present invention are provided. Also provided are diagnostic methods for detecting diseases, disorders, and/or conditions related to the Protease-42 polypeptides and polynucleotides, and therapeutic methods for treating such diseases, disorders, and/or conditions. The invention further relates to screening methods for identifying binding partners of the polypeptides, particularly activators and inhibitors of the novel Protease-42 polypeptides of the present invention.

BRIEF SUMMARY OF THE INVENTION

[0031] The present invention provides isolated nucleic acid molecules, that comprise, or alternatively consist of, a polynucleotide encoding the Protease-42 protein having the amino acid sequence shown in FIGS. 1A-C (SEQ ID NO:2) or the amino acid sequence encoded by the cDNA clone, Protease-42 deposited as ATCC Deposit Number PTA-3745 on Oct. 1, 2001.

[0032] The present invention also relates to recombinant vectors, which include the isolated nucleic acid molecules of the present invention, and to host cells containing the recombinant vectors, as well as to methods of making such vectors and host cells, in addition to their use in the production of Protease-42 polypeptides or peptides using recombinant techniques. Synthetic methods for producing the polypeptides and polynucleotides of the present invention are provided. Also provided are diagnostic methods for detecting diseases, disorders, and/or conditions related to the Protease-42 polypeptides and polynucleotides, and therapeutic methods for treating such diseases, disorders, and/or conditions. The invention further relates to screening methods for identifying binding partners of the polypeptides.

[0033] The invention further provides an isolated Protease-42 polypeptide having an amino acid sequence encoded by a polynucleotide described herein.

[0034] The invention further relates to a polynucleotide encoding a polypeptide fragment of SEQ ID NO:2, or a polypeptide fragment encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

[0035] The invention further relates to a polynucleotide encoding a polypeptide domain of SEQ ID NO:2 or a polypeptide domain encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

[0036] The invention further relates to a polynucleotide encoding a polypeptide epitope of SEQ ID NO:2 or a polypeptide epitope encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

[0037] The invention further relates to a polynucleotide encoding a polypeptide of SEQ ID NO:2 or the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1, having biological activity.

[0038] The invention further relates to a polynucleotide which is a variant of SEQ ID NO:1.

[0039] The invention further relates to a polynucleotide which is an allelic variant of SEQ ID NO:1.

[0040] The invention further relates to a polynucleotide which encodes a species homologue of the SEQ ID NO:2.

[0041] The invention further relates to a polynucleotide which represents the complimentary sequence (antisense) of SEQ ID NO:1.

[0042] The invention further relates to a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified herein, wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

[0043] The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:2, wherein the polynucleotide fragment comprises a nucleotide sequence encoding a calpain protein.

[0044] The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:1 wherein the polynucleotide fragment comprises a nucleotide sequence encoding the sequence identified as SEQ ID NO:2 or the polypeptide encoded by the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

[0045] The invention further relates to an isolated nucleic acid molecule of of SEQ ID NO:1 wherein the polynucleotide fragment comprises the entire nucleotide sequence of SEQ ID NO:1 or the cDNA sequence included in the deposited clone, which is hybridizable to SEQ ID NO:1.

[0046] The invention further relates to an isolated nucleic acid molecule of SEQ ID NO:1, wherein the nucleotide sequence comprises sequential nucleotide deletions from either the C-terminus or the N-terminus.

[0047] The invention further relates to an isolated polypeptide comprising an amino acid sequence that comprises a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

[0048] The invention further relates to a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in the deposited clone, having biological activity.

[0049] The invention further relates to a polypeptide domain of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

[0050] The invention further relates to a polypeptide epitope of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

[0051] The invention further relates to a full length protein of SEQ ID NO:2 or the encoded sequence included in the deposited clone.

[0052] The invention further relates to a variant of SEQ ID NO:2.

[0053] The invention further relates to an allelic variant of SEQ ID NO:2. The invention further relates to a species homologue of SEQ ID NO:2.

[0054] The invention further relates to the isolated polypeptide of of SEQ ID NO:2, wherein the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

[0055] The invention further relates to an isolated antibody that binds specifically to the isolated polypeptide of SEQ ID NO:2.

[0056] The invention further relates to a method for preventing, treating, or ameliorating a medical condition, comprising administering to a mammalian subject a therapeutically effective amount of the polypeptide of SEQ ID NO:2 or the polynucleotide of SEQ ID NO:1.

[0057] The invention further relates to a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising the steps of (a) determining the presence or absence of a mutation in the polynucleotide of SEQ ID NO:1; and (b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

[0058] The invention further relates to a method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising the steps of (a) determining the presence or amount of expression of the polypeptide of of SEQ ID NO:2 in a biological sample; and diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

[0059] The invention further relates to a method for identifying a binding partner to the polypeptide of SEQ ID NO:2 comprising the steps of (a) contacting the polypeptide of SEQ ID NO:2 with a binding partner; and (b) determining whether the binding partner effects an activity of the polypeptide.

[0060] The invention further relates to a gene corresponding to the cDNA sequence of SEQ ID NO:1.

[0061] The invention further relates to a method of identifying an activity in a biological assay, wherein the method comprises the steps of (a) expressing SEQ ID NO:1 in a cell, (b) isolating the supernatant; (c) detecting an activity in a biological assay; and (d) identifying the protein in the supernatant having the activity.

[0062] The invention further relates to a process for making polynucleotide sequences encoding gene products having altered activity selected from the group consisting of SEQ ID NO:2 activity comprising the steps of (a) shuffling a nucleotide sequence of SEQ ID NO:1, (b) expressing the resulting shuffled nucleotide sequences and, (c) selecting for altered activity selected from the group consisting of SEQ ID NO:2 activity as compared to the activity selected from the group consisting of SEQ ID NO:2 activity of the gene product of said unmodified nucleotide sequence.

[0063] The invention further relates to a shuffled polynucleotide sequence produced by a shuffling process, wherein said shuffled DNA molecule encodes a gene product having enhanced tolerance to an inhibitor of any one of the activities selected from the group consisting of SEQ ID NO:2 activity.

[0064] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a gastrointenstinal disorder

[0065] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a pulmonary disorder.

[0066] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a disorder related to aberrant calcium regulation.

[0067] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a disorder related to aberrant protease regulation.

[0068] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a neural disorder.

[0069] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a renal disorder.

[0070] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is an inflammatory condition.

[0071] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a tumorigenesis process in a gastroointestinal organ or tissue, particularly in colon cancer.

[0072] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a hepatic-disorder.

[0073] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a female reproductive disorder.

[0074] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a uterine disorder.

[0075] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a fallopian tube disorder.

[0076] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is ischemia-reperfusion injury.

[0077] The invention further relates to a method for preventing, treating, or ameliorating a medical condition with the polypeptide provided as SEQ ID NO:2, in addition to, its encoding nucleic acid, wherein the medical condition is a condition associated with tissue damage caused by calpain activation, either directly or indirectly.

[0078] The invention further relates to a method of identifying a compound that modulates the biological activity of Protease-42, comprising the steps of, (a) combining a candidate modulator compound with Protease-42 having the sequence set forth in one or more of SEQ ID NO:2; and measuring an effect of the candidate modulator compound on the activity of Protease-42.

[0079] The invention further relates to a method of identifying a compound that modulates the biological activity of a calpain, comprising the steps of, (a) combining a candidate modulator compound with a host cell expressing Protease-42 having the sequence as set forth in SEQ ID NO:2; and, (b) measuring an effect of the candidate modulator compound on the activity of the expressed Protease-42.

[0080] The invention further relates to a compound that modulates the biological activity of human Protease-42 as identified by the methods described herein for the treatment of diseases and tissue damage caused by calpain activation or inactivation, which would include inflammation, cancer, and hair loss.

[0081] The invention further relates to a method of identifying a compound that modulates the biological activity of Protease-42, comprising the steps of, (a) combining a candidate modulator compound with a host cell containing a vector described herein, wherein Protease-42 is expressed by the cell; and, (b) measuring an effect of the candidate modulator compound on the activity of the expressed Protease-42.

[0082] The invention further relates to a method of screening for a compound that is capable of modulating the

biological activity of Protease-42, comprising the steps of: (a) providing a host cell described herein; (b) determining the biological activity of Protease-42 in the absence of a modulator compound; (c) contacting the cell with the modulator compound; and (d) determining the biological activity of Protease-42 in the presence of the modulator compound; wherein a difference between the activity of Protease-42 in the presence of the modulator compound and in the absence of the modulator compound indicates a modulating effect of the compound.

[0083] The invention further relates to a compound that modulates the biological activity of human Protease-42 as identified by the methods described herein.

[0084] The invention also provides a machine readable storage medium which comprises the structure coordinates of Protease-42, including all or any parts conserved calpain regions. Such storage medium encoded with these data are capable of displaying on a computer screen or similar viewing device, a three-dimensional graphical representation of a molecule or molecular complex which comprises said regions or similarly shaped homologous regions.

[0085] The invention also provides a machine-readable data storage medium, comprising a data storage material encoded with machine readable data, wherein the data is defined by the structure coordinates of the model Protease-42 according to Table IV or a homologue of said model, wherein said homologue comprises any kind of surrogate atoms that have a root mean square deviation from the backbone atoms of the complex of not more than about 4.0, 3.0. 2.0, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, or 0.1 Angstroms.

[0086] The invention also provides a machine-readable data storage medium, comprising a data storage material encoded with machine readable data, wherein the data is defined by the structure coordinates of the model Protease-42 according to Table IV or a homologue of said model, wherein said homologue comprises any kind of surrogate atoms that have a root mean square deviation from the backbone atoms of the complex of not more than about 4.0, 3.0. 2.0, 1.0, 0.9, 0.8, 0.7, 0.6, 0.5, 0.4, 0.3, 0.2, or 0.1 Angstroms

[0087] The invention also provides a model comprising all or any part of the model defined by structure coordinates of Protease-42 according to Table IV, or a mutant or homologue of said molecule or molecular complex.

[0088] The invention also provides a method for identifying a mutant of Protease-42 with altered biological properties, function, or reactivity, the method comprising one or more of the following steps: (a) use of the model or a homologue of said model according to Table IV, for the design of protein mutants with altered biological function or properties which exhibit any combination of therapeutic effects described herein; and/or (b) use of the model or a homologue of said model, for the design of a protein with mutations in the active site region comprised of the amino acids S93, R94, T95, D96, V97, C98, Q99, G100, S101, L102, G103, N104, C105, W106, F107, L108, A109, A110, A111, A112, S113, L121, F167, V168, W177, E182, H199, M200, N201, A203, F204, F207, T208, G209, G210, V211, G212, E213, V214, L215, Y216, L217, R218, L237, V238, G239, A240, T241, A242, L243, S244, D245, R246, L255, V256, K257, G258, H259, A260, Y261, S262, 1263, T264, G265, L279, R280, L281, R282, N283, P284, W285, G286, C287, V288, E289, W290, K316, E317, D318, G319, E320, F321, W322, M323, L330, H331, F332, D333, T334, V335, Q336, and/or 1337 of SEQ ID NO:2 according to Table IV with altered biological function or properties which exhibit any combination of therapeutic effects described herein.

[0089] The method also relates to a method for identifying modulators of Protease-42 biological properties, function, or reactivity, the method comprising the step of modeling test compounds that fit spatially into the active site region defined by all or any portion of residues G103, N_1O_4, C105, W106, F107, L108, A109, A110, F167, T241, A242, V256, K257, G258, H259, A260, Y261, S262, L281, R282, N283, P284, W285, G286, C287, V288, D318, G319, and/or F321 of the three-dimensional structural model according to Table IV, or using a homologue or portion thereof, or analogue in which the original C, N, and O atoms have been replaced with other elements

[0090] The invention also provides methods for designing, evaluating and identifying compounds which bind to all or parts of the aforementioned regions. The methods include three dimensional model building (homology modeling) and methods of computer assisted-drug design which can be used to identify compounds which bind or modulate the forementioned regions of the Protease-42 polypeptide. Such compounds are potential inhibitors of Protease-42 or its homologues.

[0091] The invention also relates to method for identifying modulators of Protease-42 biological properties, function, or reactivity, the method comprising the step of modeling test compounds that fit spatially into the EF-hand calcium binding region defined by D633 to E644 of SEQ ID NO:2 using a homologue or portion thereof or analogue in which the original C, N, and O atoms have been replaced with other elements.

[0092] The invention also relates to method for identifying modulators of Protease-42 biological properties, function, or reactivity, the method comprising the step of modeling test compounds that fit spatially into the acidic loop region region defined by E391 to E401 of SEQ ID NO:2 using a homologue or portion thereof or analogue in which the original C, N, and O atoms have been replaced with other elements.

[0093] The invention also relates to a method of using said structure coordinates as set forth in Table IV to identify structural and chemical features of Protease-42; employing identified structural or chemical features to design or select compounds as potential Protease-42 modulators; employing the three-dimensional structural model to design or select compounds as potential Protease-42 modulators; synthesizing the potential Protease-42 modulators; screening the potential Protease-42 modulators in an assay characterized by binding of a protein to the Protease-42. The invention also relates to said method wherein the potential Protease-42 modulator is selected from a database. The invention further relates to said method wherein the potential Protease-42 modulator is designed de novo. The invention further relates to a method wherein the potential Protease-42 modulator is designed from a known modulator of activity.

BRIEF DESCRIPTION OF THE FIGURES/DRAWINGS

[0094] FIGS. 1A-C shows the polynucleotide sequence (SEQ ID NO:1) and deduced amino acid sequence (SEQ ID NO:2) of the novel human calpain, Protease-42, of the present invention. The standard one-letter abbreviation for amino acids is used to illustrate the deduced amino acid sequence. The polynucleotide sequence contains a sequence of 2220 nucleotides (SEQ ID NO:1), encoding a polypeptide of 735 amino acids (SEQ ID NO:2). An analysis of the Protease-42 polypeptide determined that it comprised the following features: a predicted EF-hand calcium binding domain located from about amino acid 644 to about amino acid 666 (SEQ ID NO:53) of SEQ ID NO:2 represented by dotted underlining; a predicted thiol (cysteine) protease domain located from about amino acid 94 to about amino acid 115 (SEQ ID NO:54) of SEQ ID NO:2 represented by double underlining; a predicted highly acidic region domain that is thought to interact with Ca^{2+} and function as an "electrostatic switch" for protease activation located from about amino acid 391 to about amino acid 401 (SEQ ID NO:54) of SEQ ID NO:2 represented by light shading; a predicted active site domain amino acids located from about amino acid S93 to about amino acid S113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2 represented by dark shading; and predicted catalytic amino acid residues within the Protease-42 active site located at amino acid C105, H259, and N283 of SEQ ID NO:2 (FIGS. 1A-C) denoted by an arrow ("↑"). The predicted active site domain amino acids are believed to form the active site binding pocket of the Protease-42 polypeptide and facilitate catalysis of appropriate calpain substrates.

[0095] FIGS. 2A-H show the regions of identity and similarity between Protease-42 and other calpains, specifically, the large catalytic subunit of the human CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase, CANP, μ-TYPE) (Calpain1; Genbank Accession No: gil12408656; SEQ ID NO:3); the human CAN2 protein (Calpain2; Genbank Accession No: gil4502563; SEQ ID NO:4); the large subunit of the human calpain 3 protein (EC 3.4.22.17) (also referred to as CALPAIN L3, CALPAIN P94, Calcium-Activated Neutral Proteinase 3, CANP 3; muscle-specific calcium-activated neutral protease 3 large subunit) (Calpain3; Genbank Accession No: gil14557405; SEQ ID NO:5); the human CAN5 protein (Calpain5; Genbank Accession No: gilNP_004046; SEQ ID NO:6); the human CAN9 protein (Calpain9; Genbank Accession No: gil5729758; SEQ ID NO:7); the human CAN10 protein (type II diabetes linked) (Calpain10; Genbank Accession No: gilNP_075574; SEQ ID NO:8); the human CAN11 protein (Calpain11; Genbank Accession No: gilNP_008989; SEQ ID NO:9); the human CAN12 protein (Calpain12; Co-pending U.S. Provisional Application: 60/300,620; SEQ ID NO:10); the large catalytic subunit of the mouse CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase) (CANP) μ-TYPE) (CAN1_MOUSE; Genbank Accession No: gil3462902; SEQ ID NO:11); the mouse CALPAIN 2 protein (CAN2_MOUSE; Genbank Accession No: gil2570158; SEQ ID NO:12); the mouse CALPAIN 6 protein (CAN6_MOUSE; Genbank Accession No: gil13959310; SEQ ID NO:13); the mouse CALPAIN 7 protein (CAN7_MOUSE; Genbank Accession No: gil6753258; SEQ ID NO:14); the mouse CALPAIN 8 protein (CAN8_MOUSE; Genbank Accession No: gil5305702; SEQ ID NO:15); and the mouse CAN12 protein (CAN12_MOUSE; Genbank Accession No. gil10303329; SEQ ID NO:16). The alignment was performed using the CLUSTALW algorithm described elsewhere herein, as available within the Vector NTI AlignX program (CLUSTALW parameters: gap opening penalty: 10; gap extension penalty: 0.5; gap separation penalty range: 8; percent identity for alignment delay: 40%; and transition weighting: 0). The darkly shaded amino acids represent regions of matching identity. The lightly shaded amino acids represent regions of matching similarity. Lines between residues indicate gapped regions for the aligned polypeptides. The arrow ("↓") denotes the characteristic active site cysteine (Cys105), histidine (His259), and asparagine (Asn283) residues of calpain proteases.

[0096] **FIG. 3** shows a hydropathy plot of the novel human calpain, Protease-42. The hydropathy plot was created using the TmPred algorithm (J. Biol. Chem. 347:166, 1993).

[0097] **FIG. 4** shows an expression profile of the novel human calpain, Protease-42. The figure illustrates the relative expression level of Protease-42 amongst various mRNA tissue sources. As shown, transcripts corresponding to Protease-42 expressed predominately in the brain, liver, and spleen. The Protease-42 polypeptide was also expressed significantly in kidney, lung, and to a lesser extent, in other tissues as shown. Expression data was obtained by measuring the steady state Protease-42 mRNA levels by quantitative PCR using the PCR primer pair provided as SEQ ID NO:17 and 18 as described herein.

[0098] **FIG. 5** shows a table illustrating the percent identity and percent similarity between the Protease-42 polypeptide of the present invention with the large catalytic subunit of the human CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase, CANP, μ-TYPE) (Calpain 1; Genbank Accession No: gil12408656; SEQ ID NO:3); the human CAN2 protein (Calpain2; Genbank Accession No: gil4502563; SEQ ID NO:4); the large subunit of the human calpain 3 protein (EC 3.4.22.17) (also referred to as CALPAIN L3, CALPAIN P94, Calcium-Activated Neutral Proteinase 3, CANP 3; muscle-specific calcium-activated neutral protease 3 large subunit) (Calpain3; Genbank Accession No: gil4557405; SEQ ID NO:5); the human CAN5 protein (Calpain5; Genbank Accession No: gilNP_004046; SEQ ID NO:6); the human CAN9 protein (Calpain9; Genbank Accession No: gil5729758; SEQ ID NO:7); the human CAN10 protein (type II diabetes linked) (Calpain10; Genbank Accession No: gilNP_075574; SEQ ID NO:8); the human CAN11 protein (Calpain11; Genbank Accession No: gilNP_008989; SEQ ID NO:9); the human CAN12 protein (Calpain12; Co-pending U.S. Provisional Application: 60/300,620; SEQ ID NO:10); the large catalytic subunit of the mouse CALPAIN 1 protein (also referred

to as Calcium-Activated Neutral Proteinase) (CANP) μ-TYPE) (CAN1_MOUSE; Genbank Accession No: gil3462902; SEQ ID NO:11); the mouse CALPAIN 2 protein (CAN2_MOUSE; Genbank Accession No: gil2570158; SEQ ID NO:12); the mouse CALPAIN 6 protein (CAN6_MOUSE; Genbank Accession No: gil13959310; SEQ ID NO:13); the mouse CALPAIN 7 protein (CAN7_MOUSE; Genbank Accession No: gil6753258; SEQ ID NO:14); the mouse CALPAIN 8 protein (CAN8_MOUSE; Genbank Accession No: gil5305702; SEQ ID NO:15); and the mouse CAN12 protein (CAN12_MOUSE; Genbank Accession No. gil10303329; SEQ ID NO:16). The percent identity and percent similarity values were determined based upon the GAP algorithm (GCG suite of programs; and Henikoff, S. and Henikoff, J. G., Proc. Natl. Acad. Sci. USA 89: 10915-10919(1992)) using the following parameters: gap weight=8, and length weight=2.

[0099] **FIG. 6** shows a three-dimensional homology model of the Protease-42 polypeptide based upon the homologous structure of a portion of the human m-calpain, also referred to as, CAN2 (hCAN2; PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19). The predicted catalytic active site amino acids of the human Protease-42 polypeptide are labeled. The predicted regions of alpha helix structure, beta sheet structure, and flexible loop structure are shown. The catalytic amino acid residues are shown in a CPK/space filled rendering of the side chain atoms. The structural coordinates of the Protease-42 polypeptide are provided in Table IV herein. The homology model of Protease-42 was derived from generating a sequence alignment with the human m-calpain, CAN2 protein (hCAN2; PDB code 1dkv; Genbank Accession No. gi16980465; SEQ ID NO:19) using the Proceryon suite of software (Proceryon Biosciences, Inc. N.Y., N.Y.), and the overall atomic model including plausible sidechain orientations using the program LOOK (V3.5.2, Molecular Applications Group).

[0100] **FIG. 7** shows an energy graph for the Protease-42 model of the present invention (dotted line) and the human m-calpain template (PDB code 1dkv) (solid line) from which the model was generated. The energy distribution for each protein fold is displayed on the y-axis, while the amino acid residue position of the protein fold is displayed on the x-axis. As shown, the Protease-42 model and 1dkv template have similar energies over the aligned region, suggesting that the structural model of Protease-42 represents a "native-like" conformation of the Protease-42 polypeptide. This graph supports the motif and sequence alignments in confirming that the three dimensional structure coordinates of Protease-42 are an accurate and useful representation of the structure of the Protease-42 polypeptide.

[0101] **FIG. 8** shows the regions of identity and similarity between Protease-42 and the human m-calpain template polypeptide sequence (PDB code 1dkv) from which the Protease-42 homology model was generated. In the alignment, the asterisk ("*") refers to identical amino acid residues; the period (".") refers to similar residues that may not be chemically related but are sterically of similar size; the colon (":") refer to similar amino acid residues that are chemically similar to each other (e.g., acidic/acidic, hydrophobic/hydrophobic etc.). The portion of the Protease-42 polypeptide that was used to create the Protease-42 homology model shown in **FIG. 6** is represented by light shading. The catalytic residues Cys105, His259 and Asn283-are indicated by the arrows ("↓").

[0102] **FIG. 9** shows an expanded expression profile of the novel full-length human calpain Protease-42 protein in normal tissues. The figure illustrates the relative expression level of Protease-42 amongst various mRNA tissue sources. As shown, the Protease-42 polypeptide was expressed predominately in the fallopian tube, the uterus (cervix), the duodenum, the gallbladder, and to a lesser extent in other tissues as shown. Expression data was obtained by measuring the steady state Protease-42 mRNA levels by quantitative PCR using the PCR primer pair provided as SEQ ID NO:55 and 56, and Taqman probe (SEQ ID NO:57) as described in Example 5 herein.

[0103] **FIG. 10** shows an expanded expression profile of the novel full-length human calpain Protease-42 protein in tumor tissues. The figure illustrates the relative expression level of Protease-42 amongst various mRNA tissue sources. As shown, the Protease-42 polypeptide was overexpressed predominately in colon tumors, relative to normal colon tissues. Expression data was obtained by measuring the steady state Protease-42 mRNA levels by quantitative PCR using the PCR primer pair provided as SEQ ID NO:55 and 56, and Taqman probe (SEQ ID NO:57) as described in Example 5 herein.

[0104] Table I provides a summary of the novel polypeptides and their encoding polynucleotides of the present invention.

[0105] Table II illustrates the preferred hybridization conditions for the polynucleotides of the present invention. Other hybridization conditions may be known in the art or are described elsewhere herein.

[0106] Table III provides a summary of various conservative substitutions encompassed by the present invention.

[0107] Table IV provides the structural coordinates of the homology model of the Protease-42 polypeptide provided in **FIG. 6**. A description of the headings are as follows: "Atom No" refers to the atom number within the Protease-42 homology model; "Atom Name" refers to the element whose coordinates are measured, the first letter in the column defines the element; "Residue" refers to the amino acid of the Protease-42 polypeptide within which the atom resides, in addition to the amino acid position in which the atom resides; "X Coord", "Y Coord", and "Z Coord" structurally define the atomic position of the element measured in three dimensions.

DETAILED DESCRIPTION OF THE INVENTION

[0108] The present invention may be understood more readily by reference to the following detailed description of the preferred embodiments of the invention and the Examples included herein.

[0109] The invention provides a novel human sequence that encodes a calpain with substantial homology to the large subunits of a variety of known calpains. Calpains affect a variety of cellular processes based upon their involvement in modulating signal transduction. Aberrations in the large subunit polypeptides of calpains have been implicated in a

number of diseases and disorders which include, for example, incidence of type II diabetes (Horikawa et al., Nat Genet. 26:163-75 (2000)), limb-girdle muscular dystrophy (Richard et al., Cell 81:27-40 (1995)), ischemia-induced damage in neurons and heart tissue, neurodegnerative disorders such as Alzheimer's disease, Multiple sclerosis, Huntington's disease, Parkinson's disease and amyotrophy, inflammatory disorders, susceptibility to infectious diseases, etc. Protease-42 polynucleotides and polypeptides, including agonists and antagonists thereof are expected to be useful in ameliorating at least some of these disorders. In addition, expression analysis indicates the Protease-42 has strong preferential expression in brain, liver, spleen, kidney, lung, and to a lesser extent, in other tissues as shown. Based on this information, we have provisionally named the gene and protein Protease-42.

[0110] In the present invention, "isolated" refers to material removed from its original environment (e.g., the natural environment if it is naturally occurring), and thus is altered "by the hand of man" from its natural state. For example, an isolated polynucleotide could be part of a vector or a composition of matter, or could be contained within a cell, and still be "isolated" because that vector, composition of matter, or particular cell is not the original environment of the polynucleotide. The term "isolated" does not refer to genomic or cDNA libraries, whole cell total or mRNA preparations, genomic DNA preparations (including those separated by electrophoresis and transferred onto blots), sheared whole cell genomic DNA preparations or other compositions where the art demonstrates no distinguishing features of the polynucleotide/sequences of the present invention.

[0111] In specific embodiments, the polynucleotides of the invention are at least 15, at least 30, at least 50, at least 100, at least 125, at least 500, or at least 1000 continuous nucleotides but are less than or equal to 300 kb, 200 kb, 100 kb, 50 kb, 15 kb, 10 kb, 7.5 kb, 5 kb, 2.5 kb, 2.0 kb, or 1 kb, in length. In a further embodiment, polynucleotides of the invention comprise a portion of the coding sequences, as disclosed herein, but do not comprise all or a portion of any intron. In another embodiment, the polynucleotides comprising coding sequences do not contain coding sequences of a genomic flanking gene (i.e., 5' or 3' to the gene of interest in the genome). In other embodiments, the polynucleotides of the invention do not contain the coding sequence of more than 1000, 500, 250, 100, 50, 25, 20, 15, 10, 5, 4, 3, 2, or 1 genomic flanking gene(s).

[0112] As used herein, a "polynucleotide" refers to a molecule having a nucleic acid sequence contained in SEQ ID NO:1 or the cDNA contained within the clone deposited with the ATCC. For example, the polynucleotide can contain the nucleotide sequence of the full length cDNA sequence, including the 5' and 3' untranslated sequences, the coding region, with or without a signal sequence, the secreted protein coding region, as well as fragments, epitopes, domains, and variants of the nucleic acid sequence. Moreover, as used herein, a "polypeptide" refers to a molecule having the translated amino acid sequence generated from the polynucleotide as broadly defined.

[0113] In the present invention, the full length sequence identified as SEQ ID NO:1 was often generated by overlapping sequences contained in multiple clones (contig analysis). A representative clone containing all or most of the sequence for SEQ ID NO:1 was deposited with the American Type Culture Collection ("ATCC"). As shown in Table 1, each clone is identified by a cDNA Clone ID (Identifier) and the ATCC Deposit Number. The ATCC is located at 10801 University Boulevard, Manassas, Va. 20110-2209, USA. The ATCC deposit was made pursuant to the terms of the Budapest Treaty on the international recognition of the deposit of microorganisms for purposes of patent procedure. The deposited clone is inserted in the pGEM-T-Easy plasmid (Promega) using the 'TA' cloning methodology in the reverse orientation as described herein.

[0114] Unless otherwise indicated, all nucleotide sequences determined by sequencing a DNA molecule herein were determined using an automated DNA sequnencer (such as the Model 373, preferably a Model 3700, from Applied Biosystems, Inc.), and all amino acid sequences of polypeptides encoded by DNA molecules determined herein were pridcted by translation of a DNA sequence determined above. Therefore, as is known in the art for any DNA seuqnece detemrined by this automated approach, any nucleotide seqence determined herein may contain some errors. Nucleotide sequences determined by automation are typically at least about 90% identical, more typically at least about 95% to at least about 99.9% identical to the actual nucleotide seqnece of the sequenced DNA molecule. The actual sequence can be more precisely determined by other approaches including manual DNA sequencing methods well known in the art. As is also known in the art, a single insertion or deletion in a detemrined nucleotide sequence compared to the actual sequence will cause a frame shift in translation of the nucleotide sequence such that the predicted amino acid sequence encoded by a determined nucleotide sequence will be completely different from the amino acid sequence actually encoded bt the sequenced DNA molecule, beginning at the point of such an insertion or deletion.

[0115] Using the information provided herein, such as the nucletide sequence in FIGS. 1A-C (SEQ ID NO:1), a nucleic acid molecule of the present invention encoding the Protease-42 polypeptide may be obtained using standard cloning and screening procedures, such as those for cloning cDNAs using mRNA as starting material. Illustrative of the invention, the nucleic acid molecule described in FIGS. 1A-C (SEQ ID NO:1) was discovered in a mixture of cDNA libraries derived from human brain and testis.

[0116] The determined nucleotide sequence of the Protease-42 cDNA in FIGS. 1A-C (SEQ ID NO:1) contains an open reading frame encoding a protein of about 735 amino acid residues, with a deduced molecular weight of about 82.5 kDa. The amino acid sequence of the predicted Protease-42 polypeptide is shown in FIGS. 1A-C (SEQ ID NO:2).

[0117] A "polynucleotide" of the present invention also includes those polynucleotides capable of hybridizing, under stringent hybridization conditions, to sequences contained in SEQ ID NO:1, the complement thereof, or the cDNA within the clone deposited with the ATCC. "Stringent hybridization conditions" refers to an overnight incubation at 42 degree C. in a solution comprising 50% formamide, 5×SSC (750 mM NaCl, 75 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5× Denhardt's solution, 10% dextran sulfate, and

20 μg/ml denatured, sheared salmon sperm DNA, followed by washing the filters in 0.1×SSC at about 65 degree C.

[0118] Also contemplated are nucleic acid molecules that hybridize to the polynucleotides of the present invention at lower stringency hybridization conditions. Changes in the stringency of hybridization and signal detection are primarily accomplished through the manipulation of formamide concentration (lower percentages of formamide result in lowered stringency); salt conditions, or temperature. For example, lower stringency conditions include an overnight incubation at 37 degree C. in a solution comprising 6×SSPE (20×SSPE=3M NaCl; 0.2M NaH2PO4; 0.02M EDTA, pH 7.4), 0.5% SDS, 30% formamide, 100 ug/ml salmon sperm blocking DNA; followed by washes at 50 degree C. with 1×SSPE, 0.1% SDS. In addition, to achieve even lower stringency, washes performed following stringent hybridization can be done at higher salt concentrations (e.g. 5×SSC).

[0119] Note that variations in the above conditions may be accomplished through the inclusion and/or substitution of alternate blocking reagents used to suppress background in hybridization experiments. Typical blocking reagents include Denhardt's reagent, BLOTTO, heparin, denatured salmon sperm DNA, and commercially available proprietary formulations. The inclusion of specific blocking reagents may require modification of the hybridization conditions described above, due to problems with compatibility.

[0120] Of course, a polynucleotide which hybridizes only to polyA+ sequences (such as any 3' terminal polyA+ tract of a cDNA shown in the sequence listing), or to a complementary stretch of T (or U) residues, would not be included in the definition of "polynucleotide," since such a polynucleotide would hybridize to any nucleic acid molecule containing a poly (A) stretch or the complement thereof (e.g., practically any double-stranded cDNA clone generated using oligo dT as a primer).

[0121] The polynucleotide of the present invention can be composed of any polyribonucleotide or polydeoxribonucleotide, which may be unmodified RNA or DNA or modified RNA or DNA. For example, polynucleotides can be composed of single- and double-stranded DNA, DNA that is a mixture of single- and double-stranded regions, single- and double-stranded RNA, and RNA that is mixture of single- and double-stranded regions, hybrid molecules comprising DNA and RNA that may be single-stranded or, more typically, double-stranded or a mixture of single- and double-stranded regions. In addition, the polynucleotide can be composed of triple-stranded regions comprising RNA or DNA or both RNA and DNA. A polynucleotide may also contain one or more modified bases or DNA or RNA backbones modified for stability or for other reasons. "Modified" bases include, for example, tritylated bases and unusual bases such as inosine. A variety of modifications can be made to DNA and RNA; thus, "polynucleotide" embraces chemically, enzymatically, or metabolically modified forms.

[0122] The polypeptide of the present invention can be composed of amino acids joined to each other by peptide bonds or modified peptide bonds, i.e., peptide isosteres, and may contain amino acids other than the 20 gene-encoded amino acids. The polypeptides may be modified by either natural processes, such as posttranslational processing, or by chemical modification techniques which are well known in the art. Such modifications are well described in basic texts and in more detailed monographs, as well as in a voluminous research literature. Modifications can occur anywhere in a polypeptide, including the peptide backbone, the amino acid side-chains and the amino or carboxyl termini. It will be appreciated that the same type of modification may be present in the same or varying degrees at several sites in a given polypeptide. Also, a given polypeptide may contain many types of modifications. Polypeptides may be branched, for example, as a result of ubiquitination, and they may be cyclic, with or without branching. Cyclic, branched, and branched cyclic polypeptides may result from posttranslation natural processes or may be made by synthetic methods. Modifications include acetylation, acylation, ADP-ribosylation, amidation, covalent attachment of flavin, covalent attachment of a heme moiety, covalent attachment of a nucleotide or nucleotide derivative, covalent attachment of a lipid or lipid derivative, covalent attachment of phosphotidylinositol, cross-linking, cyclization, disulfide bond formation, demethylation, formation of covalent cross-links, formation of cysteine, formation of pyroglutamate, formylation, gamma-carboxylation, glycosylation, GPI anchor formation, hydroxylation, iodination, methylation, myristoylation, oxidation, pegylation, proteolytic processing, phosphorylation, prenylation, racemization, selenoylation, sulfation, transfer-RNA mediated addition of amino acids to proteins such as arginylation, and ubiquitination. (See, for instance, Proteins—Structure and Molecular Properties, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York (1993); Posttranslational Covalent Modification of Proteins, B. C. Johnson, Ed., Academic Press, New York, pgs. 1-12 (1983); Seifter et al., Meth Enzymol 182:626-646 (1990); Rattan et al., Ann NY Acad Sci 663:48-62 (1992)).

[0123] It is another aspect of the present invention to provide modulators of the Protease-19 protein and Protease-19 peptide targets which can affect the function or activity of Protease-19 in a cell in which Protease-19 function or activity is to be modulated or affected. In addition, modulators of Protease-19 can affect downstream systems and molecules that are regulated by, or which interact with, Protease-19 in the cell. Modulators of Protease-19 include compounds, materials, agents, drugs, and the like, that antagonize, inhibit, reduce, block, suppress, diminish, decrease, or eliminate Protease-19 function and/or activity. Such compounds, materials, agents, drugs and the like can be collectively termed "antagonists". Alternatively, modulators of Protease-19 include compounds, materials, agents, drugs, and the like, that agonize, enhance, increase, augment, or amplify Protease-19 function in a cell. Such compounds, materials, agents, drugs and the like can be collectively termed "agonists".

[0124] As used herein the terms "modulate" or "modulates" refer to an increase or decrease in the amount, quality or effect of a particular activity, DNA, RNA, or protein. The definition of "modulate" or "modulates" as used herein is meant to encompass agonists and/or antagonists of a particular activity, DNA, RNA, or protein.

[0125] As will be appreciated by the skilled practitioner, should the amino acid fragment comprise an antigenic epitope, for example, biological function per se need not be maintained. The terms Protease-19 polypeptide and Protease-19 protein are used interchangeably herein to refer to the encoded product of the Protease-19 nucleic acid sequence according to the present invention.

[0126] "SEQ ID NO:1" refers to a polynucleotide sequence while "SEQ ID NO:2" refers to a polypeptide sequence, both sequences identified by an integer specified in Table 1.

[0127] "A polypeptide having biological activity" refers to polypeptides exhibiting activity similar, but not necessarily identical to, an activity of a polypeptide of the present invention, including mature forms, as measured in a particular biological assay, with or without dose dependency. In the case where dose dependency does exist, it need not be identical to that of the polypeptide, but rather substantially similar to the dose-dependence in a given activity as compared to the polypeptide of the present invention (i.e., the candidate polypeptide will exhibit greater activity or not more than about 25-fold less and, preferably, not more than about tenfold less activity, and most preferably, not more than about three-fold less activity relative to the polypeptide of the present invention).

[0128] The term "organism" as referred to herein is meant to encompass any organism referenced herein, though preferably to eukaryotic organisms, more preferably to mammals, and most preferably to humans.

[0129] The present invention encompasses the identification of proteins, nucleic acids, or other molecules, that bind to polypeptides and polynucleotides of the present invention (for example, in a receptor-ligand interaction). The polynucleotides of the present invention can also be used in interaction trap assays (such as, for example, that discribed by Ozenberger and Young (Mol Endocrinol., 9(10):1321-9, (1995); and Ann. N.Y. Acad. Sci., 7;766:279-81, (1995)).

[0130] The polynucleotide and polypeptides of the present invention are useful as probes for the identification and isolation of full-length cDNAs and/or genomic DNA which correspond to the polynucleotides of the present invention, as probes to hybridize and discover novel, related DNA sequences, as probes for positional cloning of this or a related sequence, as probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides, as probes to quantify gene expression, and as probes for microarays.

[0131] In addition, polynucleotides and polypeptides of the present invention may comprise one, two, three, four, five, six, seven, eight, or more membrane domains.

[0132] Also, in preferred embodiments the present invention provides methods for further refining the biological fuction of the polynucleotides and/or polypeptides of the present invention.

[0133] Specifically, the invention provides methods for using the polynucleotides and polypeptides of the invention to identify orthologs, homologs, paralogs, variants, and/or allelic variants of the invention. Also provided are methods of using the polynucleotides and polypeptides of the invention to identify the entire coding region of the invention, non-coding regions of the invention, regulatory sequences of the invention, and secreted, mature, pro-, prepro-, forms of the invention (as applicable).

[0134] In preferred embodiments, the invention provides methods for identifying the glycosylation sites inherent in the polynucleotides and polypeptides of the invention, and the subsequent alteration, deletion, and/or addition of said sites for a number of desirable characteristics which include, but are not limited to, augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion.

[0135] In further preferred embodiments, methods are provided for evolving the polynucleotides and polypeptides of the present invention using molecular evolution techniques in an effort to create and identify novel variants with desired structural, functional, and/or physical characteristics.

[0136] The present invention further provides for other experimental methods and procedures currently available to derive functional assignments. These procedures include but are not limited to spotting of clones on arrays, micro-array technology, PCR based methods (e.g., quantitative PCR), anti-sense methodology, gene knockout experiments, and other procedures that could use sequence information from clones to build a primer or a hybrid partner.

Polynucleotides and Polypeptides of the Invention

Features of the Polypeptide Encoded by Gene No:1

[0137] The polypeptide of this gene provided as SEQ ID NO:2 (FIGS. **1**A-C), encoded by the polynucleotide sequence according to SEQ ID NO:1 (FIGS. **1**A-C), and/or encoded by the polynucleotide contained within the deposited clone, Protease-42, has significant homology at the nucleotide and amino acid level to the large catalytic subunit of the human CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase, CANP, μ-TYPE) (Calpain1; Genbank Accession No: gil2408656; SEQ ID NO:3); the human CAN2 protein (Calpain2; Genbank Accession No: gil4502563; SEQ ID NO:4); the large subunit of the human calpain 3 protein (EC 3.4.22.17) (also referred to as CALPAIN L3, CALPAIN P94, Calcium-Activated Neutral Proteinase 3, CANP 3; muscle-specific calcium-activated neutral protease 3 large subunit) (Calpain3; Genbank Accession No: gil4557405; SEQ ID NO:5); the human CAN5 protein (Calpain5; Genbank Accession No: gilNP_004046; SEQ ID NO:6); the human CAN9 protein (Calpain9; Genbank Accession No: gil5729758; SEQ ID NO:7); the human CAN10 protein (type II diabetes linked) (Calpain10; Genbank Accession No: gilNP_075574; SEQ ID NO:8); the human CAN11 protein (Calpain11; Genbank Accession No: gilNP_008989; SEQ ID NO:9); the human CAN12 protein (Calpain12; Co-pending U.S. Provisional Application: 60/300,620; SEQ ID NO:10); the large catalytic subunit of the mouse CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase) (CANP) μ-TYPE) (CAN1_MOUSE; Genbank Accession No: gil3462902; SEQ ID NO:11); the mouse CALPAIN 2 protein (CAN2_MOUSE; Genbank Accession No: gil2570158; SEQ ID NO:12); the mouse CALPAIN 6 protein (CAN6_MOUSE; Genbank Accession No: gil13959310; SEQ ID NO:13); the mouse CALPAIN 7 protein (CAN7_MOUSE; Genbank Accession No: gil6753258; SEQ ID NO:14); the mouse CALPAIN 8 protein (CAN8_MOUSE; Genbank Accession No: gil5305702; SEQ ID NO:15); and the mouse CAN12 protein (CAN12_MOUSE; Genbank Accession No. gil10303329; SEQ ID NO:16). An alignment of the Protease-42 polypeptide with these proteins is provided in FIGS. **2**A-H.

[0138] The Protease-42 polypeptide was determined to share 45.3% identity and 53.8% similarity with the large catalytic subunit of the human CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase, CANP, μ-TYPE) (Calpain1; Genbank Accession No: gil12408656; SEQ ID NO:3); to share 44.4% identity and 55.0% similarity with the human CAN2 protein (Calpain2; Genbank Accession No: gil4502563; SEQ ID NO:4); to share 40.5% identity and 51.9% similarity with the large subunit of the human calpain 3 protein (EC 3.4.22.17) (also referred to as CALPAIN L3, CALPAIN P94, Calcium-Activated Neutral Proteinase 3, CANP 3; muscle-specific calcium-activated neutral protease 3 large subunit) (Calpain3; Genbank Accession No: gil4557405; SEQ ID NO:5); to share 32.8% identity and 43.2% similarity with the human CAN5 protein (Calpain5; Genbank Accession No: gilNP_004046; SEQ ID NO:6); to share 42.6% identity and 52.1% similarity with the human CAN9 protein (Calpain9; Genbank Accession No: gil5729758; SEQ ID NO:7); to share 32.4% identity and 38.6% similarity with the human CAN100protein (type II diabetes linked) (Calpain10; Genbank Accession No: gilNP_075574; SEQ ID NO:8); to share 42.6% identity and 51.9% similarity with the human CAN11 protein (Calpain11; Genbank Accession No: gilNP_008989; SEQ ID NO:9); to share 37.3% identity and 44.6% similarity with the human CAN12 protein (Calpain12; Co-pending U.S. Provisional Application: 60/300,620; SEQ ID NO:10); to share 44.2% identity and 53.5% similarity with the large catalytic subunit of the mouse CALPAIN 1 protein (also referred to as Calcium-Activated Neutral Proteinase) (CANP) μ-TYPE) (CAN1_MOUSE; Genbank Accession No: gil3462902; SEQ ID NO:11); to share 44.8% identity and 55.0% similarity with the mouse CALPAIN 2 protein (CAN2_MOUSE; Genbank Accession No: gil2570158; SEQ ID NO:12); to share 32.1% identity and 41.4% similarity with the mouse CALPAIN 6 protein (CAN6_MOUSE; Genbank Accession No: gil13959310; SEQ ID NO:13); to share 25.5% identity and 34.3% similarity with the mouse CALPAIN 7 protein (CAN7_MOUSE; Genbank Accession No: gil6753258; SEQ ID NO:14); to share 34.4% identity and 40.6% similarity with the mouse CALPAIN 8 protein (CAN8_MOUSE; Genbank Accession No: gil5305702; SEQ ID NO:15); and to share 86.2% identity and 88.3% similarity with the mouse CAN12 protein (CAN12_MOUSE; Genbank Accession No. gil10303329; SEQ ID NO:16) as shown in **FIG. 5**.

[0139] The Protease-42 polypeptide of the present invention is believed to represent the human ortholog of the mouse CAN12 polypeptide (CAN12_MOUSE; Genbank Accession No. gil10303329; SEQ ID NO:16). The significant identity between the mouse CAN12 polypeptide and Protease-42 (SEQ ID NO:2) is consistent with this result.

[0140] Protease-42 polypeptides and polynucleotides are useful for diagnosing diseases related to the over and/or under expression of Protease-42 by identifying mutations in the Protease-42 gene using Protease-42 sequences as probes or by determining Protease-42 protein or mRNA expression levels. Protease-42 polypeptides will be useful in screens for compounds that affect the activity of the protein. Protease-42 peptides can also be used for the generation of specific antibodies and as bait in yeast two hybrid screens to find proteins the specifically interact with Protease-42. Based on the expression pattern of this novel sequence, diseases that can be treated with agonists and/or antagonists for Protease-42 including, but not limited to, epilepsy, Bartter's syndrome, persistent hyperinsulinemic hypoglycemia of infancy, hyperkalemia and hypokalemia, cystic fibrosis and hypercalciuric nephrolithiasis.

[0141] Protease-42 polynucleotides and polypeptides, in addition to fragments and/or modulators thereof, are useful in treating ischemia-reperfusion injury or tumorigenesis processes in a variety of tissues, particularly the brain, liver, spleen, lung, kidney, and in the digestive track. In addition, identification of endogenous substrate(s) of Protease-42 might help define the underlying mechanisms in hair proliferation and differentiation and lead to the development of a novel drug target for the treatment of alopecia. In preferred embodiments, Protease-42 polynucleotides and polypeptides, in addition to fragments and/or modulators thereof, are useful in treating alopecia, male pattern baldness, chemotherapy induced hair loss, and other hair-related conditions.

[0142] Protein threading and molecular modeling of Protease-42 suggests that Protease-42 has a structural fold similar to representative m-calpains. Moreover, the structural and threading alignments of the present invention suggest that amino acids 105 ("C"), 259 ("H"), and 283 ("N") of SEQ ID NO:2 (FIGS. **1**A-C) may represent the catalytic amino acids within the active site domain. Thus, based upon the sequence and structural homology to known calpains, particularly the presence of the thiol cysteine protease active site domain, the novel Protease-42 is believed to represent a novel human calpain.

[0143] As discussed more particularly herein, calpains are a group of structurally diverse, high molecular weight (400 to 500 amino acids) proteins that have a catalytic cysteine amino acid and one or more calcium binding domains. Despite the structural heterogeneity, calpains share some well defined structural-functional characteristics, particularly in their active site domains.

[0144] In preferred embodiments, the Protease-42 polypeptide of the present invention is directed to a polypeptide having structural similarity to calpains.

[0145] Expression profiling designed to measure the steady state mRNA levels encoding the Protease-42 polypeptide showed predominately high expression levels in brain, liver, and spleen; significantly in kidney, lung, and to a lesser extent, in other tissues (as shown in **FIG. 4**).

[0146] SYBR green quantitative PCR analysis of Protease42 on a limited number of tissues indicated that this putative novel calpain protease is expressed at low levels and in a restricted manner (**FIG. 4**). Additional expression profiling on an expanded mRNA panel of tissues using TaqMan™ quantitative PCR revealed additional tissues in which the Protease-42 transcripts were expressed which include the fallopian tube, the uterus (cervix), the duodenum, and the gallbladder (as shown in **FIG. 9**). These data suggest that polynucleotides and polypeptides of Protease-42, including fragments and modulators thereof, may be useful in the treatment, diagnosis, and/or amelioration of female reproductive tract disorders, including infertility and carcinomas.

[0147] Moreover, additional expression profiling on an expanded mRNA panel of tumor and normal tissues using TaqMan™ quantitative PCR indicated Protease-42 transcripts were preferentially overexpressed in colon tumor

tissues relative to normal colon tissues as shown in **FIG. 10**. These data suggest that Protease-42 may play an important role in tumor progression modulators of Protease-42, particularly inhibitors, may have utility in the treatment and/or amelioration of colon cancer.

[**0148**] Moreover, polynucleotides encoding the Protease-42 polypeptide of the present invention were found to map to chromosome 19q13.1. Polynucleotides and polypeptides, including fragments and/or modulators thereof are useful for the treatment, amelioration, and/or diagnosis of diseases or disorders that map at or near the chromosome 19q13.1 locus.

[**0149**] Based upon the strong homology to members of the calpain family, the Protease-42 polypeptide is expected to share at least some biological activity with calpains, preferably with m-calpain family members, and more preferable to the large subunits of m-calpain family members, in addition to other calpains and calpain subunits referenced herein and/or otherwise known in the art.

[**0150**] Moreover, the tissue distribution of Protease-42, in conjunction with the strong homology to calpains and their associated functions, suggests that the Protease-42 polynucleotides and polypeptides could participate in remodeling/disassembly of cytoskeletal/plasma membrane interactions, or could be involved in various pathological states such as acute or chronic inflammation, ischemia-reperfusion injury, and cancers.

[**0151**] The Protease-42 polynucleotides and polypeptides of the present invention, including agonists and/or fragments thereof, have uses that include modulating cellular adhesion events, cellular proliferation, and inflammation, in various cells, tissues, and organisms, and particularly in mammalian brain, liver, spleen, kidney, and lung tissue, preferably human tissue. Protease-42 polynucleotides and polypeptides of the present invention, including agonists and/or fragments thereof, may be useful in diagnosing, treating, prognosing, and/or preventing neural, hepatic, immune, hematopoietic, renal, pulmonary, and/or proliferative diseases or disorders.

[**0152**] The strong homology to calpain proteins, combined with the localized expression in brain tissue suggests suggests the Protease-42 polynucleotides and polypeptides may be useful in treating, diagnosing, prognosing, and/or preventing neurodegenerative disease states, behavioral disorders, or inflammatory conditions. Representative uses are described in the "Regeneration" and "Hyperproliferative Disorders" sections below, in the Examples, and elsewhere herein. Briefly, the uses include, but are not limited to the detection, treatment, and/or prevention of Alzheimer's Disease, Parkinson's Disease, Huntington's Disease, Tourette Syndrome, meningitis, encephalitis, demyelinating diseases, peripheral neuropathies, neoplasia, trauma, congenital malformations, spinal cord injuries, ischemia and infarction, aneurysms, hemorrhages, schizophrenia, mania, dementia, paranoia, obsessive compulsive disorder, depression, panic disorder, learning disabilities, ALS, psychoses, autism, and altered behaviors, including disorders in feeding, sleep patterns, balance, and perception. In addition, elevated expression of this gene product in regions of the brain indicates it plays a role in normal neural function. Potentially, this gene product is involved in synapse formation, neurotransmission, learning, cognition, homeostasis, or neuronal differentiation or survival. Furthermore, the protein may also be used to determine biological activity, to raise antibodies, as tissue markers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

[**0153**] The strong homology to calpain proteins, combined with the localized expression in liver tissue suggests the potential utility for Protease-42 polynucleotides and polypeptides in treating, diagnosing, prognosing, and/or preventing hepatic disorders. Representative uses are described in the "Hyperproliferative Disorders", "Infectious Disease", and "Binding Activity" sections below, and elsewhere herein. Briefly, the protein can be used for the detection, treatment, amelioration, and/or prevention of hepatoblastoma, jaundice, hepatitis, liver metabolic diseases and conditions that are attributable to the differentiation of hepatocyte progenitor cells, cirrhosis, hepatic cysts, pyrogenic abscess, amebic abcess, hydatid cyst, cystadenocarcinoma, adenoma, focal nodular hyperplasia, hemangioma, hepatocellulae carcinoma, cholangiocarcinoma, and angiosarcoma, granulomatous liver disease, liver transplantation, hyperbilirubinemia, jaundice, parenchymal liver disease, portal hypertension, hepatobiliary disease, hepatic parenchyma, hepatic fibrosis, anemia, gallstones, cholestasis, carbon tetrachloride toxicity, beryllium toxicity, vinyl chloride toxicity, choledocholithiasis, hepatocellular necrosis, aberrant metabolism of amino acids, aberrant metabolism of carbohydrates, aberrant synthesis proteins, aberrant synthesis of glycoproteins, aberrant degradation of proteins, aberrant degradation of glycoproteins, aberrant metabolism of drugs, aberrant metabolism of hormones, aberrant degradation of drugs, aberrant degradation of drugs, aberrant regulation of lipid metabolism, aberrant regulation of cholesterol metabolism, aberrant glycogenesis, aberrant glycogenolysis, aberrant glycolysis, aberrant gluconeogenesis, hyperglycemia, glucose intolerance, hyperglycemia, decreased hepatic glucose uptake, decreased hepatic glycogen synthesis, hepatic resistance to insulin, portal-systemic glucose shunting, peripheral insulin resistance, hormonal abnormalities, increased levels of systemic glucagon, decreased levels of systemic cortisol, increased levels of systemic insulin, hypoglycemia, decreased gluconeogenesis, decreased hepatic glycogen content, hepatic resistance to glucagon, elevated levels of systemic aromatic amino acids, decreased levels of systemic branched-chain amino acids, hepatic encephalopathy, aberrant hepatic amino acid transamination, aberrant hepatic amino acid oxidative deamination, aberrant ammonia synthesis, aberant albumin secretion, hypoalbuminemia, aberrant cytochromes b5 function, aberrant P450 function, aberrant glutathione S-acyltransferase function, aberrant cholesterol synthesis, and aberrant bile acid synthesis.

[**0154**] Moreover, polynucleotides and polypeptides, including fragments and/or antagonists thereof, have uses which include, directly or indirectly, treating, preventing, diagnosing, and/or prognosing the following, non-limiting, hepatic infections: liver disease caused by sepsis infection, liver disease caused by bacteremia, liver disease caused by *Pneomococcal pneumonia* infection, liver disease caused by Toxic shock syndrome, liver disease caused by Listeriosis, liver disease caused by Legionnaries' disease, liver disease caused by Brucellosis infection, liver disease caused by *Neisseria gonorrhoeae* infection, liver disease caused by

Yersinia infection, liver disease caused by Salmonellosis, liver disease caused by Nocardiosis, liver disease caused by Spirochete infection, liver disease caused by *Treponema pallidum* infection, liver disease caused by *Brrelia burgdorferi* infection, liver disease caused by Leptospirosis, liver disease caused by *Coxiella burnetii* infection, liver disease caused by *Rickettsia richettsii* infection, liver disease caused by *Chlamydia trachomatis* infection, liver disease caused by *Chlamydia psittaci* infection, liver disease caused by hepatitis virus infection, liver disease caused by Epstein-Barr virus infection in addition to any other hepatic disease and/or disorder implicated by the causative agents listed above or elsewhere herein.

[**0155**] The strong homology to calpain proteins, combined with the localized expression in spleen tissue suggests the Protease-42 polynucleotides and polypeptides may be useful in treating, diagnosing, prognosing, and/or preventing immune diseases and/or disorders. Representative uses are described in the "Immune Activity", "Chemotaxis", and "Infectious Disease" sections below, and elsewhere herein. Briefly, the strong expression in immune tissue indicates a role in regulating the proliferation; survival; differentiation; and/or activation of hematopoietic cell lineages, including blood stem cells.

[**0156**] The Protease-42 polypeptide may also be useful as a preventative agent for immunological disorders including arthritis, asthma, immunodeficiency diseases such as AIDS, leukemia, rheumatoid arthritis, granulomatous disease, inflammatory bowel disease, sepsis, acne, neutropenia, neutrophilia, psoriasis, hypersensitivities, such as T-cell mediated cytotoxicity; immune reactions to transplanted organs and tissues, such as host-versus-graft and graft-versus-host diseases, or autoimmunity disorders, such as autoimmune infertility, lense tissue injury, demyelination, systemic lupus erythematosis, drug induced hemolytic anemia, rheumatoid arthritis, Sjogren's disease, and scleroderma. The Protease-42 polypeptide may be useful for modulating cytokine production, antigen presentation, or other processes, such as for boosting immune responses, etc.

[**0157**] Moreover, the protein may represent a factor that influences the differentiation or behavior of other blood cells, or that recruits hematopoietic cells to sites of injury. Thus, this gene product is thought to be useful in the expansion of stem cells and committed progenitors of various blood lineages, and in the differentiation and/or proliferation of various cell types. Furthermore, the protein may also be used to determine biological activity, raise antibodies, as tissuemarkers, to isolate cognate ligands or receptors, to identify agents that modulate their interactions, in addition to its use as a nutritional supplement. Protein, as well as, antibodies directed against the protein may show utility as a tumor marker and/or immunotherapy targets for the above listed tissues.

[**0158**] The strong homology to calpain proteins, combined with the localized expression in kidney tissue suggests suggests the Protease-42 polynucleotides and polypeptides may be useful in treating, diagnosing, prognosing, and/or preventing renal diseases and/or disorders, which include, but are not limited to: nephritis, renal failure, nephrotic syndrome, urinary tract infection, hematuria, proteinuria, oliguria, polyuria, nocturia, edema, hypertension, electrolyte disorders, sterile pyuria, renal osteodystrophy, large kidneys, renal transport defects, nephrolithiasis, azotemia, anuria, urinary retention, slowing of urinary stream, large prostate, flank tenderness, full bladder sensation after voiding, enuresis, dysuria, bacteriuria, kideny stones, glomerulonephritis, vasculitis, hemolytic uremic syndromes, thrombotic thrombocytopenic purpura, malignant hypertension, casts, tubulointerstitial kidney diseases, renal tubular acidosis, pyelonephritis, hydronephritis, nephrotic syndrome, crush syndrome, and/or renal colic, in addition to Wilm's Tumor Disease, and congenital kidney abnormalities such as horseshoe kidney, polycystic kidney, and Falconi's syndrome for example.

[**0159**] The strong homology to calpain proteins, combined with the localized expression in lung suggests the Protease-42 polynucleotides and polypeptides may be useful in treating, diagnosing, prognosing, and/or preventing pulmonary diseases and disorders which include the following, not limiting examples: ARDS, emphysema, cystic fibrosis, interstitial lung disease, chronic obstructive pulmonary disease, bronchitis, lymphangioleiomyomatosis, pneumonitis, eosinophilic pneumonias, granulomatosis, pulmonary infarction, pulmonary fibrosis, pneumoconiosis, alveolar hemorrhage, neoplasms, lung abscesses, empyema, and increased susceptibility to lung infections (e.g., immumocompromised, HIV, etc.), for example.

[**0160**] Moreover, polynucleotides and polypeptides, including fragments and/or antagonists thereof, have uses which include, directly or indirectly, treating, preventing, diagnosing, and/or prognosing the following, non-limiting, pulmonary infections: pnemonia, bacterial pnemonia, viral pnemonia (for example, as caused by Influenza virus, Respiratory syncytial virus, Parainfluenza virus, Adenovirus, Coxsackievirus, Cytomegalovirus, *Herpes simplex* virus, Hantavirus, etc.), mycobacteria pnemonia (for example, as caused by *Mycobacterium tuberculosis*, etc.) mycoplasma pnemonia, fungal pnemonia (for example, as caused by *Pneumocystis carinii, Histoplasma capsulatum, Coccidioides immitis, Blastomyces dermatitidis,* Candida sp., *Cryptococcus neoformans*, Aspergillus sp., Zygomycetes, etc.), Legionnaires' Disease, Chlamydia pnemonia, aspiration pnemonia, Nocordia sp. Infections, parasitic pnemonia (for example, as caused by Strongyloides, *Toxoplasma gondii*, etc.) necrotizing pnemonia, in addition to any other pulmonary disease and/or disorder (e.g., non-pneumonia) implicated by the causative agents listed above or elsewhere herein.

[**0161**] The strong homology to calpain proteins, combined with the localized expression in fallopian tubes and uterine tissue suggests a potential utility for Protease-42 polynucleotides and polypeptides in treating, diagnosing, prognosing, and/or preventing female reproductive disorders, particularly of the uterus In preferred embodiments, Protease-42 polynucleotides and polypeptides including agonists and fragments thereof, have uses which include treating, diagnosing, prognosing, and/or preventing the following, non-limiting, female reproductive diseases or disorders: dysfunctional uterine bleeding, amenorrhea, primary dysmenorrhea, sexual dysfunction, infertility, pelvic inflammatory disease, endometriosis, placental aromatase deficiency, premature menopause, placental dysfunction, pelvic inflammatory disease, tubal pregnancy, and Chlamydial infection.

[0162] Protease-42 polynucleotides and polypeptides, including fragments and/or antagonsists thereof, may have uses which include identification of modulators of Protease-42 function including antibodies (for detection or neutralization), naturally-occurring modulators and small molecule modulators. Antibodies to domains (including Protease-42 epitopes provided herein) of the Protease-42 protein could be used as diagnostic agents of inflammatory conditions in patients, are useful in monitoring the activation and presence of cognate proteases, and can be used as a biomarker for the protease involvement in disease states and in the evaluation of inhibitors of the cognate protease in vivo.

[0163] Protease-42 polypeptides and polynucleotides are useful for diagnosing diseases related to over or under expression of Protease-42 proteins by identifying mutations in the Protease-42 gene using Protease-42 probes, or determining Protease-42 protein or mRNA expression levels. Protease-42 polypeptides are also useful for screening for compounds, which affect activity of the protein. Diseases that can be treated with Protease-42 include, the following, non-limiting examples: neuro-regeneration, neuropathic pain, obesity, anorexia, HIV infections, cancers, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, osteoporosis, angina pectoris, myocardial infarction, psychotic, neural, metabolic, hepatic, immune, hematopoietic, pulmonary, renal, and neurological disorders.

[0164] The Protease-42 polynucleotides and polypeptides also have uses which include, but are not limited to treating, diagnosing, prognosing, and/or preventing proliferative disorders which include the following non-limiting examples: carcinoid tumor, islet cell carcinoma, Zollinger-Ellison gastrinoma, insulinoma, vipoma, glucagonoma, somatostatinoma, grfoma, crfoma, ppoma, neurotensinoma, and small cell carcinoma.

[0165] In addition, antagonists of the Protease-42 polynucleotides and polypeptides may have uses that include diagnosing, treating, prognosing, and/or preventing diseases or disorders related to hyper calpain activity, which neurological, metabolic, hepatic, immune, hematopoietic, pulmonary, renal, and/or proliferative diseases or disorders.

[0166] Alternatively, Protease-42 polypeptides of the invention, or agonists thereof, are administered to treat, prevent, prognose, and/or diagnose disorders involving excessive smooth muscle tone or excitability, which include, but are not limited to asthma, angina, hypertension, incontinence, pre-term labor, and irratible bowel syndrome.

[0167] Molecular genetic manipulation of the structure of the active site domain, particularly the predicted catalytic amino acids, and of other functional domains in the calpain family (e.g., active site domain binding pocket) enables the production of calpains with tailor-made activities. Thus, the Protease-42 polypeptides, and fragments thereof, as well as any homologous product resulting from genetic manipulation of the structure, are useful for NMR-based design of modulators of Protease-42 biological activity, and calpains, in general.

[0168] Protease-42 polypeptides and polynucleotides have additional uses which include diagnosing diseases related to the over and/or under expression of Protease-42 by identifying mutations in the Protease-42 gene by using Protease-42 sequences as probes or by determining Protease-42 protein or mRNA expression levels. Protease-42 polypeptides may be useful for screening compounds that affect the activity of the protein. Protease-42 peptides can also be used for the generation of specific antibodies and as bait in yeast two hybrid screens to find proteins the specifically interact with Protease-42 (described elsewhere herein).

[0169] The Protease-42 polynucleotides and polypeptides, including fragments and agonists thereof, may have uses which include detecting, diagnosing, treating, ameliorating, and/or preventing metabolic diseases and disorders, such as diabetes. Moreover, expressed human Protease-42 may be useful in the detection of patients susceptible to diabetes. Also paradigms that would simulate intracellular Protease-42 activity would be useful in treating diabetes.

[0170] The Protease-42 polynucleotides and polypeptides, including fragments thereof, may have uses which include identifying inhibitors of intracellular calpain inhibitors (calpastatins) leading to an effective increase in calpain activity.

[0171] Although it is believed the encoded polypeptide may share at least some biological activities with human calpains (particularly m-calpains), a number of methods of determining the exact biological function of this clone are either known in the art or are described elsewhere herein. Briefly, the function of this clone may be determined by applying microarray methodology. Nucleic acids corresponding to the Protease-42 polynucleotides, in addition to, other clones of the present invention, may be arrayed on microchips for expression profiling. Depending on which polynucleotide probe is used to hybridize to the slides, a change in expression of a specific gene may provide additional insight into the function of this gene based upon the conditions being studied. For example, an observed increase or decrease in expression levels when the polynucleotide probe used comes from diseased brain tissue, as compared to, normal tissue might indicate a function in modulating neurological function, for example. In the case of Protease-42, brain, liver, spleen, kidney, and/or lung tissue should be used to extract RNA to prepare the probe.

[0172] In addition, the function of the protein may be assessed by applying quantitative PCR methodology, for example. Real time quantitative PCR would provide the capability of following the expression of the Protease-42 gene throughout development, for example. Quantitative PCR methodology requires only a nominal amount of tissue from each developmentally important step is needed to perform such experiments. Therefore, the application of quantitative PCR methodology to refining the biological function of this polypeptide is encompassed by the present invention. In the case of Protease-42, a disease correlation related to Protease-42 may be made by comparing the mRNA expression level of Protease-42 in normal tissue, as compared to diseased tissue (particularly diseased tissue isolated from the following: brain, liver, spleen, kidney, and/or lung tissue). Significantly higher or lower levels of Protease-42 expression in the diseased tissue may suggest Protease-42 plays a role in disease progression, and antagonists against Protease-42 polypeptides would be useful therapeutically in treating, preventing, and/or ameliorating the disease. Alternatively, significantly higher or lower levels of Protease-42 expression in the diseased tissue may

suggest Protease-42 plays a defensive role against disease progression, and agonists of Protease-42 polypeptides may be useful therapeutically in treating, preventing, and/or ameliorating the disease. Also encompassed by the present invention are quantitative PCR probes corresponding to the polynucleotide sequence provided as SEQ ID NO:1 (FIGS. 1A-C).

[0173] The function of the protein may also be assessed through complementation assays in yeast. For example, in the case of the Protease-42, transforming yeast deficient in calpain activity, particularly m-calpain activity, and assessing their ability to grow would provide convincing evidence the Protease-42 polypeptide has calpain activity, and possibly m-calpain activity. Additional assay conditions and methods that may be used in assessing the function of the polynucleotides and polypeptides of the present invention are known in the art, some of which are disclosed elsewhere herein.

[0174] Alternatively, the biological function of the encoded polypeptide may be determined by disrupting a homologue of this polypeptide in Mice and/or rats and observing the resulting phenotype. Such knock-out experiments are known in the art, some of which are disclosed elsewhere herein.

[0175] Moreover, the biological function of this polypeptide may be determined by the application of antisense and/or sense methodology and the resulting generation of transgenic mice and/or rats. Expressing a particular gene in either sense or antisense orientation in a transgenic mouse or rat could lead to respectively higher or lower expression levels of that particular gene. Altering the endogenous expression levels of a gene can lead to the observation of a particular phenotype that can then be used to derive indications on the function of the gene. The gene can be either over-expressed or under expressed in every cell of the organism at all times using a strong ubiquitous promoter, or it could be expressed in one or more discrete parts of the organism using a well characterized tissue-specific promoter (e.g., a brain, liver, spleen, kidney, or lung-specific promoter), or it can be expressed at a specified time of development using an inducible and/or a developmentally regulated promoter.

[0176] In the case of Protease-42 transgenic mice or rats, if no phenotype is apparent in normal growth conditions, observing the organism under diseased conditions (neurological, hepatic, immune, hematopoietic, renal, or pulmonary diseases or disorders, cancers, etc.) may lead to understanding the function of the gene. Therefore, the application of antisense and/or sense methodology to the creation of transgenic mice or rats to refine the biological function of the polypeptide is encompassed by the present invention.

[0177] In preferred embodiments, the following N-terminal Protease-42 deletion polypeptides are encompassed by the present invention: M1-S735, A2-S735, S3-S735, S4-S735, S5-S735, G6-S735, R7-S735, V8-S735, T9-S735, 110-S735, Q11-S735, L12-S735, V13-S735, D14-S735, E15-S735, E16-S735, A17-S735, G18-S735, V19-S735, G20-S735, A21-S735, G22-S735, R23-S735, L24-S735, Q25-S735, L26-S735, F27-S735, R28-S735, G29-S735, Q30-S735, S31-S735, Y32-S735, E33-S735, A34-S735, I35-S735, R36-S735, A37-S735, A38-S735, C39-S735, L40-S735, D41-S735, S42-S735, G43-S735, I44-S735, L45-S735, F46-S735, R47-S735, D48-S735, P49-S735, Y50-S735, F51-S735, P52-S735, A53-S735, G54-S735, P55-S735, D56-S735, A57-S735, L58-S735, G59-S735, Y60-S735, D61-S735, Q62-S735, L63-S735, G64-S735, P65-S735, D66-S735, S67-S735, E68-S735, K69-S735, A70-S735, K71-S735, G72-S735, V73-S735, K74-S735, W75-S735, M76-S735, R77-S735, P78-S735, H79-S735, E80-S735, F81-S735, C82-S735, A83-S735, E84-S735, P85-S735, K86-S735, F87-S735, I88-S735, C89-S735, E90-S735, D91-S735, M92-S735, S93-S735, R94-S735, T95-S735, D96-S735, V97-S735, C98-S735, Q99-S735, G100S735, S101-S735, L102-S735, G103-S735, N104-S735, C105-S735, W106-S735, F107-S735, L108-S735, A109-S735, A110-S735, A111-S735, A112-S735, S113-S735, L114-S735, T115-S735, L116-S735, Y117-S735, P118-S735, R119-S735, L120-S735, L121-S735, R122-S735, R123-S735, V124-S735, V125-S735, P126-S735, P127-S735, G128-S735, Q129-S735, D130-S735, F131-S735, Q132-S735, H133-S735, G134-S735, Y135-S735, A136-S735, G137-S735, V138-S735, F139-S735, H140-S735, F141-S735, Q142-S735, L143-S735, W144-S735, Q145-S735, F146-S735, G147-S735, R148-S735, W149-S735, M150-S735, D151-S735, V152-S735, V153-S735, V154-S735, D155-S735, D156-S735, R157-S735, L158-S735, P159-S735, V160-S735, R161-S735, E162-S735, G163-S735, K164-S735, L165-S735, M166-S735, F167-S735, V168-S735, R169-S735, S170-S735, E171-S735, Q172-S735, R173-S735, N174-S735, E175-S735, F176-S735, W177-S735, A178-S735, P179-S735, L180-S735, L181-S735, E182-S735, K183-S735, A184-S735, Y185-S735, A186-S735, K187-S735, L188-S735, H189-S735, G190-S735, S191-S735, Y192-S735, E193-S735, V194-S735, M195-S735, R196-S735, G197-S735, G198-S735, H199-S735, M200-S735, N201-S735, E202-S735, A203-S735, F204-S735, V205-S735, D206-S735, F207-S735, T208-S735, G209-S735, G210-S735, V211-S735, G212-S735, E213-S735, V214-S735, L215-S735, Y216-S735, L217-S735, R218-S735, Q219-S735, N220-S735, S221-S735, M222-S735, G223-S735, L224-S735, F225-S735, S226-S735, A227-S735, L228-S735, R229-S735, H230-S735, A231-S735, L232-S735, A233-S735, K234-S735, E235-S735, S236-S735, L237-S735, V238-S735, G239-S735, A240-S735, T241-S735, A242-S735, L243-S735, S244-S735, D245-S735, R246-S735, G247-S735, E248-S735, Y249-S735, R250-S735, T251-S735, E252-S735, E253-S735, G254-S735, L255-S735, V256-S735, K257-S735, G258-S735, H259-S735, A260-S735, Y261-S735, S262-S735, I263-S735, T264-S735, G265-S735, T266-S735, H267-S735, K268-S735, V269-S735, F270-S735, L271-S735, G272-S735, F273-S735, T274-S735, K275-S735, V276-S735, R277-S735, L278-S735, L279-S735, R280-S735, L281-S735, R282-S735, N283-S735, P284-S735, W285-S735, G286-S735, C287-S735, V288-S735, E289-S735, W290-S735, T291-S735, G292-S735, A293-S735, W294-S735, S295-S735, D296-S735, S297-S735, C298-S735, P299-S735, R300-S735, W301-S735, D302-S735, T303-S735, L304-S735, P305-S735, T306-S735, E307-S735, C308-S735, R309-S735, D310-S735, A311-S735, L312-S735, L313-S735, V314-S735, K315-S735, K316-S735, E317-S735, D318-S735, G319-S735, E320-S735, F321-S735, W322-S735, M323-S735, E324-S735, L325-S735, R326-S735, D327-S735, F328-S735, L329-S735, L330-S735, H331-S735, F332-S735, D333-S735, T334-S735, V335-S735, Q336-S735, 1337-S735, C338-

S735, S339-S735, L340-S735, S341-S735, P342-S735, E343-S735, V344-S735, L345-S735, G346-S735, P347-S735, S348-S735, P349-S735, E350-S735, G351-S735, G352-S735, G353-S735, W354-S735, H355-S735, V356-S735, H357-S735, T358-S735, F359-S735, Q360-S735, G361-S735, R362-S735, W363-S735, V364-S735, R365-S735, G366-S735, F367-S735, N368-S735, S369-S735, G370-S735, G371-S735, S372-S735, Q373-S735, P374-S735, N375-S735, A376-S735, E377-S735, T378-S735, F379-S735, W380-S735, T381-S735, N382-S735, P383-S735, Q384-S735, F385-S735, R386-S735, L387-S735, T388-S735, L389-S735, L390-S735, E391-S735, P392-S735, D393-S735, E394-S735, E395-S735, D396-S735, D397-S735, E398-S735, D399-S735, E400-S735, E401-S735, G402-S735, P403-S735, W404-S735, G405-S735, G406-S735, W407-S735, G408-S735, A409-S735, A410-S735, G411-S735, A412-S735, R413-S735, G414-S735, P415-S735, A416-S735, R417-S735, G418-S735, G419-S735, R420-S735, T421-S735, P422-S735, K423-S735, C424-S735, T425-S735, V426-S735, L427-S735, L428-S735, S429-S735, L430-S735, I431-S735, Q432-S735, R433-S735, N434-S735, R435-S735, R436-S735, R437-S735, L438-S735, R439-S735, A440-S735, K441-S735, G442-S735, L443-S735, T444-S735, Y445-S735, L446-S735, T447-S735, V448-S735, G449-S735, F450-S735, H451-S735, V452-S735, F453-S735, Q454-S735, I455-S735, P456-S735, E457-S735, E458-S735, L459-S735, L460-S735, G461-S735, L462-S735, W463-S735, D464-S735, S465-S735, P466-S735, R467-S735, S468-S735, H469-S735, A470-S735, L471-S735, L472-S735, P473-S735, R474-S735, L475-S735, L476-S735, R477-S735, A478-S735, D479-S735, R480-S735, S481-S735, P482-S735, L483-S735, S484-S735, A485-S735, R486-S735, R487-S735, D488-S735, V489-S735, T490-S735, R491-S735, R492-S735, C493-S735, C494-S735, L495-S735, R496-S735, P497-S735, G498-S735, H499-S735, Y500-S735, L501-S735, V502-S735, V503-S735, P504-S735, S505-S735, T506-S735, A507-S735, H508-S735, A509-S735, G510-S735, D511-S735, E512-S735, A513-S735, D514-S735, F515-S735, T516-S735, L517-S735, R518-S735, V519-S735, F520-S735, S521-S735, E522-S735, R523-S735, R524-S735, H525-S735, T526-S735, A527-S735, V528-S735, E529-S735, I530-S735, D531-S735, D532-S735, V533-S735, I534-S735, S535-S735, A536-S735, D537-S735, L538-S735, Q539-S735, S540-S735, L541-S735, Q542-S735, V543-S735, G544-S735, T545-S735, V546-S735, P547-S735, G548-S735, G549-S735, A550-S735, A551-S735, W552-S735, G553-S735, G554-S735, D555-S735, L556-S735, G557-S735, Q558-S735, and/or G559-S735 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these N-terminal Protease-42 deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0178] In preferred embodiments, the following C-terminal Protease-42 deletion polypeptides are encompassed by the present invention: M1-S735, M1-F734, M1-T733, M1-A732, M1-V731, M1-E730, M1-M729, M1-W728, M1-Q727, M1-R726, M1-H725, M1-T724, M1-L723, M1-C722, M1-1721, M1-V720, M1-G719, M1-E718, M1-G717, M1-G716, M1-D715, M1-L714, M1-H713, M1-Q712, M1-S711, M1-C710, M1-H709, M1-C708, M1-F707, M1-1706, M1-C705, M1-T704, M1-L703, M1-H702, M1-A701, M1-V700, M1-C699, M1-S698, M1-V697, M1-F696, M1-R695, M1-E694, M1-F693, M1-D692, M1-V691, M1-R690, M1-L689, M1-R688, M1-S687, M1-D686, M1-R685, M1-Y684, M1-R683, M1-S682, M1-T681, M1-L680, M1-T679, M1-Q678, M1-T677, M1-L676, M1-Q675, M1-N674, M1-N673, M1-L672, M1-H671, M1-F670, M1-G669, M1-A668, M1-A667, M1-N666, M1-L665, M1-A664, M1-L663, M1-R662, M1-L661, M1-E660, M1-Y659, M1-S658, M1-N657, M1-M656, M1-T655, M1-G654, M1-S653, M1-T652, M1-D651, M1-E650, M1-D649, M1-F648, M1-K647, M1-N646, M1-F645, M1-1644, M1-A643, M1-Q642, M1-W641, M1-E640, M1-L639, M1-L638, M1-Y637, M1-G636, M1-W635, M1-L634, M1-Q633, M1-Q632, M1-F631, M1-H630, M1-H629, M1-L628, M1-A627, M1-L626, M1-S625, M1-Q624, M1-G623, M1-H622, M1-G621, M1-F620, M1-C619, M1-Q618, M1-L617, M1-L616, M1-Q615, M1-E614, M1-C613, M1-T612, M1-R611, M1-L610, M1-G609, M1-1608, M1-E607, M1-R606, M1-P605, M1-T604, M1-S603, M1-T602, M1-H601, M1-A600, M1-R599, M1-A598, M1-P597, M1-E596, M1-L595, M1-A594, M1-I593, M1-S592, M1-L591, M1-L590, M1-A589, M1-Q588, M1-L587, M1-Q586, M1-S585, M1-A584, M1-N583, M1-L582, M1-E581, M1-E580, M1-E579, M1-E578, M1-G577, M1-A576, M1-L575, M1-E574, M1-Q573, M1-F572, M1-L571, M1-Q570, M1-E569, M1-L568, M1-G567, M1-L566, M1-E565, M1-L564, M1-P563, M1-L562, M1-Y561, M1-P560, M1-G559, M1-Q558, M1-G557, M1-L556, M1-D555, M1-G554, M1-G553, M1-W552, M1-A551, M1-A550, M1-G549, M1-G548, M1-P547, M1-V546, M1-T545, M1-G544, M1-V543, M1-Q542, M1-L541, M1-S540, M1-Q539, M1-L538, M1-D537, M1-A536, M1-S535, M1-I534, M1-V533, M1-D532, M1-D531, M1-I530, M1-E529, M1-V528, M1-A527, M1-T526, M1-H525, M1-R524, M1-R523, M1-E522, M1-S521, M1-F520, M1-V519, M1-R518, M1-L517, M1-T516, M1-F515, M1-D514, M1-A513, M1-E512, M1-D511, M1-G510, M1-A509, M1-H508, M1-A507, M1-T506, M1-S505, M1-P504, M1-V503, M1-V502, M1-L501, M1-Y500, M1-H499, M1-G498, M1-P497, M1-R496, M1-L495, M1-C494, M1-C493, M1-R492, M1-R491, M1-T490, M1-V489, M1-D488, M1-R487, M1-R486, M1-A485, M1-S484, M1-L483, M1-P482, M1-S481, M1-R480, M1-D479, M1-A478, M1-R477, M1-L476, M1-L475, M1-R474, M1-P473, M1-L472, M1-L471, M1-A470, M1-H469, M1-S468, M1-R467, M1-P466, M1-S465, M1-D464, M1-W463, M1-L462, M1-G461, M1-L460, M1-L459, M1-E458, M1-E457, M1-P456, M1-I455, and/or M1-Q454 of SEQ ID NO:2. Polynucleotide sequences encoding these polypeptides are also provided. The present invention also encompasses the use of these C-terminal Protease-42 deletion polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0179] Alternatively, preferred polypeptides of the present invention may comprise polypeptide sequences corresponding to, for example, internal regions of the Protease-42 polypeptide (e.g., any combination of both N- and C-terminal Protease-42 polypeptide deletions) of SEQ ID NO:2. For example, internal regions could be defined by the equation: amino acid NX to amino acid CX, wherein NX refers to any N-terminal deletion polypeptide amino acid of Protease-42 (SEQ ID NO:2), and where CX refers to any C-terminal

deletion polypeptide amino acid of Protease-42 (SEQ ID NO:2). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these polypeptides as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0180] The Protease-42 polypeptides of the present invention were determined to comprise several phosphorylation sites based upon the Motif algorithm (Genetics Computer Group, Inc.). The phosphorylation of such sites may regulate some biological activity of the Protease-42 polypeptide. For example, phosphorylation at specific sites may be involved in regulating the proteins ability to associate or bind to other molecules (e.g., proteins, ligands, substrates, DNA, etc.).

[0181] The Protease-42 polypeptide was predicted to comprise thirteen PKC phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). In vivo, protein kinase C exhibits a preference for the phosphorylation of serine or threonine residues. The PKC phosphorylation sites have the following consensus pattern: [ST]-x-[RK], where S or T represents the site of phosphorylation and 'x' an intervening amino acid residue. Additional information regarding PKC phosphorylation sites can be found in Woodget J. R., Gould K. L., Hunter T., Eur. J. Biochem. 161:177-184(1986), and Kishimoto A., Nishiyama K., Nakanishi H., Uratsuji Y., Nomura H., Takeyama Y., Nishizuka Y., J. Biol. Chem. 260:12492-12499(1985); which are hereby incorporated by reference herein.

[0182] In preferred embodiments, the following PKC phosphorylation site polypeptides are encompassed by the present invention: MASSSGRVTIQL (SEQ ID NO:20), QLGPDSEKAKGVK (SEQ ID NO:21), GATALSDRGEYRT (SEQ ID NO:22), YSITGTHKVFLGF (SEQ ID NO:23), ARGGRTPKCTVLL (SEQ ID NO:24), LGLWDSPRSHALL (SEQ ID NO:25), DRSPLSARRDVTR (SEQ ID NO:26), ARRDVTRRCCLRP (SEQ ID NO:27), DEADFTLRVFSER (SEQ ID NO:28), TLRVFSERRHTAV (SEQ ID NO:29), RAHTSTPREIGLR (SEQ ID NO:30), LTQTLTSRYRDSR (SEQ ID NO:31), and/or GVICLTHRQWMEV (SEQ ID NO:32). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these Protease-42 PKC phosphorylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0183] The Protease-42 polypeptide was predicted to comprise four casein kinase II phosphorylation sites using the Motif algorithm (Genetics Computer Group, Inc.). Casein kinase II (CK-2) is a protein serine/threonine kinase whose activity is independent of cyclic nucleotides and calcium. CK-2 phosphorylates many different proteins. The substrate specificity [1] of this enzyme can be summarized as follows: (1) Under comparable conditions Ser is favored over Thr.; (2) An acidic residue (either Asp or Glu) must be present three residues from the C-terminal of the phosphate acceptor site; (3) Additional acidic residues in positions +1, +2, +4, and +5 increase the phosphorylation rate. Most physiological substrates have at least one acidic residue in these positions; (4) Asp is preferred to Glu as the provider of acidic determinants; and (5) A basic residue at the N-terminal of the acceptor site decreases the phosphorylation rate, while an acidic one will increase it.

[0184] A consensus pattern for casein kinase II phosphorylations site is as follows: [ST]-x(2)-[DE], wherein 'x' represents any amino acid, and S or T is the phosphorylation site.

[0185] Additional information specific to casein kinase II phosphorylation sites may be found in reference to the following publication: Pinna L. A., Biochim. Biophys. Acta 1054:267-284(1990); which is hereby incorporated herein in its entirety.

[0186] In preferred embodiments, the following casein kinase II phosphorylation site polypeptide is encompassed by the present invention: ICEDMSRTDVCQGS (SEQ ID NO:33), PQFRLTLLEPDEED (SEQ ID NO:34), SERRHTAVEIDDVI (SEQ ID NO:35), and/or RAHTSTPREIGLRT (SEQ ID NO:36). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of this casein kinase II phosphorylation site polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0187] The Protease-42 polypeptide was predicted to comprise one cAMP-and cGMP-dependent protein kinase phosphorylation site using the Motif algorithm (Genetics Computer Group, Inc.). There have been a number of studies relative to the specificity of cAMP-and cGMP-dependent protein kinases. Both types of kinases appear to share a preference for the phosphorylation of serine or threonine residues found close to at least two consecutive N-terminal basic residues.

[0188] A consensus pattern for cAMP-and cGMP-dependent protein kinase phosphorylation sites is as follows: [RK](2)-x-[ST], wherein "x" represents any amino acid, and S or T is the phosphorylation site.

[0189] Additional information specific to cAMP-and cGMP-dependent protein kinase phosphorylation sites may be found in reference to the following publication: Fremisco J. R., Glass D. B., Krebs E. G, J. Biol. Chem. 255:4240-4245(1980); Glass D. B., Smith S. B., J. Biol. Chem. 258:14797-14803(1983); and Glass D. B., El-Maghrabi M. R., Pilkis S. J., J. Biol. Chem. 261:2987-2993(1986); which is hereby incorporated herein in its entirety.

[0190] In preferred embodiments, the following cAMP-and cGMP-dependent protein kinase phosphorylation site polypeptide is encompassed by the present invention: RVFSERRHTAVEID (SEQ ID NO:37). Polynucleotides encoding this polypeptide are also provided. The present invention also encompasses the use of this cAMP-and cGMP-dependent protein kinase phosphorylation site polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0191] The Protease-42 polypeptide has been shown to comprise one glycosylation site according to the Motif algorithm (Genetics Computer Group, Inc.). As discussed more specifically herein, protein glycosylation is thought to serve a variety of functions including: augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion.

[0192] Asparagine glycosylation sites have the following consensus pattern, N-{P}-[ST]-{P}, wherein N represents

the glycosylation site. However, it is well known that that potential N-glycosylation sites are specific to the consensus sequence Asn-Xaa-Ser/Thr. However, the presence of the consensus tripeptide is not sufficient to conclude that an asparagine residue is glycosylated, due to the fact that the folding of the protein plays an important role in the regulation of N-glycosylation. It has been shown that the presence of proline between Asn and Ser/Thr will inhibit N-glycosylation; this has been confirmed by a recent statistical analysis of glycosylation sites, which also shows that about 50% of the sites that have a proline C-terminal to Ser/Thr are not glycosylated. Additional information relating to asparagine glycosylation may be found in reference to the following publications, which are hereby incorporated by reference herein: Marshall R. D., Annu. Rev. Biochem. 41:673-702(1972); Pless D. D., Lennarz W. J., Proc. Natl. Acad. Sci. U.S.A. 74:134-138(1977); Bause E., Biochem. J. 209:331-336(1983); Gavel Y., von Heijne G., Protein Eng. 3:433-442(1990); and Miletich J. P., Broze G. J. Jr., J. Biol. Chem. 265:11397-11404(1990).

[0193] In preferred embodiments, the following asparagine glycosylation site polypeptide is encompassed by the present invention: EEEELNASQLQALL (SEQ ID NO:38). Polynucleotides encoding this polypeptide are also provided. The present invention also encompasses the use of this Protease-42 asparagine glycosylation site polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0194] The Protease-42 polypeptide was predicted to comprise ten N-myristoylation sites using the Motif algorithm (Genetics Computer Group, Inc.). An appreciable number of eukaryotic proteins are acylated by the covalent addition of myristate (a C14-saturated fatty acid) to their N-terminal residue via an amide linkage. The sequence specificity of the enzyme responsible for this modification, myristoyl CoA:protein N-myristoyl transferase (NMT), has been derived from the sequence of known N-myristoylated proteins and from studies using synthetic peptides. The specificity seems to be the following: i.) The N-terminal residue must be glycine; ii.) In position 2, uncharged residues are allowed; iii.) Charged residues, proline and large hydrophobic residues are not allowed; iv.) In positions 3 and 4, most, if not all, residues are allowed; v.) In position 5, small uncharged residues are allowed (Ala, Ser, Thr, Cys, Asn and Gly). Serine is favored; and vi.) In position 6, proline is not allowed.

[0195] A consensus pattern for N-myristoylation is as follows: G-{EDRKHPFYW}-x(2)-[STAGCN]-{P}, wherein 'x' represents any amino acid, and G is the N-myristoylation site.

[0196] Additional information specific to N-myristoylation sites may be found in reference to the following publication: Towler D. A., Gordon J. T., Adams S. P., Glaser L., Annu. Rev. Biochem. 57:69-99(1988); and Grand R. J. A., Biochem. J. 258:625-638(1989); which is hereby incorporated herein in its entirety.

[0197] In preferred embodiments, the following N-myristoylation site polypeptides are encompassed by the present invention: VDEEAGVGAGRLQLFR (SEQ ID NO:39), TDVCQGSLGNCWFLAA (SEQ ID NO:40), YEVMRGGHMNEAFVDF (SEQ ID NO:41), RQNSMGLFSALRHALA (SEQ ID NO:42), YRTEEGLVKGHAYSIT (SEQ ID NO:43), GFNSGGSQPNAETFWT (SEQ ID NO:44), EEGPWGGWGAAGARGP (SEQ ID NO:45), PWGGWGAAGARGPARG (SEQ ID NO:46), QSLQVGTVPGGAAWGG (SEQ ID NO:47), GTVPGGAAWGGDLGQG (SEQ ID NO:48), GGAAWGGDLGQGPYLP (SEQ ID NO:49), TPREIGLRTCEQLLQC (SEQ ID NO:50), QCFGHGQSLALHHFQQ (SEQ ID NO:51), and/or DEDTSGTMNSYELRLA (SEQ ID NO:52). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of these N-myristoylation site polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0198] The present invention also encompasses immunogenic and/or antigenic epitopes of the Protease-42 polypeptide.

[0199] In confirmation of the Protease-42 polypeptide being a novel calpain, it has been shown to comprise one EF-hand calcium-binding domain according to the Motif algorithm (Genetics Computer Group, Inc.). Many calcium-binding proteins belong to the same evolutionary family and share a type of calcium-binding domain known as the EF-hand. This type of domain consists of a twelve residue loop flanked on both side by a twelve residue alpha-helical domain. In an EF-hand loop the calcium ion is coordinated in a pentagonal bipyramidal configuration. The six residues involved in the binding are in positions 1, 3, 5, 7, 9 and 12; these residues are denoted by X, Y, Z, -Y, -X and -Z. The invariant Glu or Asp at position 12 provides two oxygens for liganding Ca (bidentate ligand). Several representative proteins containing EF-hand regions are provided below: For each type of protein, the total number of EF-hand regions known or supposed to exist are provided in parenthesis: Aequorin and Renilla luciferin binding protein (LBP) (Ca=3); Alpha actinin (Ca=2); Calbindin (Ca=4); Calcineurin B subunit (protein phosphatase 2B regulatory subunit) (Ca=4); Calcium-binding protein from *Streptomyces erythraeus* (Ca=3?); Calcium-binding protein from *Schistosoma mansoni* (Ca=2?); Calcium-binding proteins TCBP-23 and TCBP-25 from *Tetrahymena thermophila* (Ca=4?); Calcium-dependent protein kinases (CDPK) from plants (Ca=4); Calcium vector protein from amphoxius (Ca=2); Calcyphosin (thyroid protein p24) (Ca=4?); Calmodulin (Ca=4, except in yeast where Ca=3); Calpain small and large chains (Ca=2); Calretinin (Ca=6); Calcyclin (prolactin receptor associated protein) (Ca=2); Caltractin (centrin) (Ca=2 or 4); Cell Division Control protein 31 (gene CDC31) from yeast (Ca=2?); Diacylglycerol kinase (EC 2.7.1.107) (DGK) (Ca=2); FAD-dependent glycerol-3-phosphate dehydrogenase (EC 1.1.99.5) from mammals (Ca=1); Fimbrin (plastin) (Ca=2); Flagellar calcium-binding protein (1f8) from *Trypanosoma cruzi* (Ca=1 or 2); Guanylate cyclase activating protein (GCAP) (Ca=3); Inositol phospholipid-specific phospholipase C isozymes gamma-1 and delta-1 (Ca=2) [10]; Intestinal calcium-binding protein (ICaBPs) (Ca=2); MIF related proteins 8 (MRP-8 or CFAG) and 14 (MRP-14) (Ca=2); Myosin regulatory light chains (Ca=1); Oncomodulin (Ca=2); Osteonectin (basement membrane protein BM-40) (SPARC) and proteins that contains an 'osteonectin' domain (QR1, matrix glycoprotein SC1) (Ca=1); Parvalbumins alpha and beta (Ca=2); Placental calcium-binding protein (18a2) (nerve growth factor induced protein 42a) (p9k) (Ca=2); Recoverins (visinin, hippocalcin, neurocalcin, S-modulin) (Ca=2 to 3); Reticulocalbin (Ca=4); S-100 protein, alpha and beta chains (Ca=2); Sarcoplasmic calcium-

binding protein (SCPs) (Ca=2 to 3); Sea urchin proteins Spec 1 (Ca=4), Spec 2 (Ca=4?), Lps-1 (Ca=8); Serine/threonine protein phosphatase rdgc (EC 3.1.3.16) from Drosophila (Ca=2); Sorcin V19 from hamster (Ca=2); Spectrin alpha chain (Ca=2); Squidulin (optic lobe calcium-binding protein) from squid (Ca=4); and Troponins C; from skeletal muscle (Ca=4), from cardiac muscle (Ca=3), from arthropods and molluscs (Ca=2).

[0200] A consensus pattern for EF hand calcium binding domains is the following:

```
1 2  3    4         5         6         7    8        9          10    12    13
X    Y              Z                   -Y          -X               -Z
D-x-[DNS]-{ILVFYW}-[DENSTG]-[DNQGHRK]-{GP}-[LIVMC]-[DENQSTAGC]-x(2)-[DE]-[LIVMFYW],
```

[0201] wherein X, Y, Z, -Y, -X, and -Z are as defined above, and wherein "x" represents any amino acid. Amino acid residues within the consensus at positions I (X), 3 (Y) and 12 (-Z) are the most conserved. The 6th residue in an EF-hand loop is in most cases a Gly.

[0202] Additional information relating to EF-hand calcium binding domains may be found in reference to the following publications, which are hereby incorporated by reference herein: Kawasaki H., Kretsinger R. H., Protein Prof. 2:305-490(1995); Kretsinger R. H., Cold Spring Harbor Symp. Quant. Biol. 52:499-510(1987); Moncrief N. D., Kretsinger R. H., Goodman M., J. Mol. Evol. 30:522-562(1990); Nakayama S., Moncrief N. D., Kretsinger R. H., J. Mol. Evol. 34:416-448(1992); Heizmann C. W., Hunziker W., Trends Biochem. Sci. 16:98-103(1991); Kligman D., Hilt D. C., Trends Biochem. Sci. 13:437-443(1988); Strynadka N. C. J., James M. N. G., Annu. Rev. Biochem. 58:951-98(1989); Haiech J., Sallantin J., Biochimie 67:555-560(1985); Chauvaux S., Beguin P., Aubert J.-P., Bhat K. M., Gow L. A., Wood T. M., Bairoch A., Biochem. J. 265:261-265(1990); Bairoch A., Cox J. A., FEBS Lett. 269:454-456(1990).

[0203] In preferred embodiments, the following EF-hand calcium binding domain polypeptide is encompassed by the present invention: IFNKFDEDTSGTMNSYELRLALN (SEQ ID NO:53). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of this EF-hand calcium binding domain polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0204] In further confirmation of the CAN-12 polypeptide being a calpain, it has been shown to comprise one eukaryotic thiol (cysteine) protease active site domain according to the Motif algorithm (Genetics Computer Group, Inc.). Eukaryotic thiol proteases (EC 3.4.22.-) are a family of proteolytic enzymes which contain an active site cysteine. Catalysis proceeds through a thioester intermediate and is facilitated by a nearby histidine side chain; an asparagine completes the essential catalytic triad. Non-limiting examples of proteases which are known to belong to this family are provided below: Vertebrate lysosomal cathepsins B (EC 3.4.22.1), H (EC 3.4.22.16), L (EC 3.4.22.15), and S (EC 3.4.22.27); Vertebrate lysosomal dipeptidyl peptidase I (EC 3.4.14.1) (also known as cathepsin C); Vertebrate calpains (EC 3.4.22.17) (Calpains are intracellular calcium-activated thiol protease that contain both a N-terminal catalytic domain and a C-terminal calcium-binding domain; Mammalian cathepsin K, which seems involved in osteoclastic bone resorption; Human cathepsin O; Bleomycin hydrolase (An enzyme that catalyzes the inactivation of the antitumor drug BLM (a glycopeptide); Plant enzymes: barley aleurain (EC 3.4.22.16), EP-B1/B4; kidney bean EP-C1, rice bean SH-EP; kiwi fruit actinidin (EC 3.4.22.14); papaya latex papain (EC 3.4.22.2), chymopapain (EC 3.4.22.6), caricain (EC 3.4.22.30), and proteinase IV (EC 3.4.22.25); pea turgor-responsive protein 15A; pineapple stem bromelain (EC 3.4.22.32); rape COT44; rice oryzain alpha, beta, and gamma; tomato low-temperature induced, *Arabidopsis thaliana* A494, RD19A and RD21A; House-dust mites allergens DerP1 and EurM1; Cathepsin B-like proteinases from the worms *Caenorhabditis elegans* (genes gcp-1, cpr-3, cpr-4, cpr-5 and cpr-6), *Schistosoma mansoni* (antigen SM31) and Japonica (antigen SJ31), *Haemonchus contortus* (genes AC-1 and AC-2), and *Ostertagia ostertagi* (CP-1 and CP-3); Slime mold cysteine proteinases CP1 and CP2; Cruzipain from *Trypanosoma cruzi* and *brucei;* Throphozoite cysteine proteinase (TCP) from various Plasmodium species; Proteases from *Leishmania mexicana, Theileria annulata* and *Theileria parva;* Baculoviruses cathepsin-like enzyme (v-cath); Drosophila small optic lobes protein (gene sol), a neuronal protein that contains a calpain-like domain; Yeast thiol protease BLH1/YCP1/LAP3; and *Caenorhabditis elegans* hypothetical protein $CO_6G4.2$, a calpain-like protein; Two bacterial peptidases are also part of this family—Aminopeptidase C from *Lactococcus lactis* (gene pepC), and Thiol protease tpr from *Porphyromonas gingivalis.*

[0205] A consensus pattern for eukaryotic thiol (cysteine) protease active site domains is the following: Q-x(3)-[GE]-x-C-[YW]-x(2)-[STAGC]-[STAGCV], wherein C is the active site residue, and "x" represents any amino acid. The residue in position 4 of the pattern is almost always cysteine; the only exceptions are calpains (Leu), bleomycin hydrolase (Ser) and yeast YCP1 (Ser); while the residue in position 5 of the pattern is always Gly except in papaya protease IV where it is Glu.

[0206] An additional consensus pattern for eukaryotic thiol (cysteine) protease active site domains is the following: [LIVMGSTAN]-x-H-[GSACE]-[LIVM]-x-[LIVMAT](2)-G-x-[GSADNH], wherein H is the active site residue, and "x" represents any amino acid.

[0207] An additional consensus pattern for eukaryotic thiol (cysteine) protease active site domains is the following: [FYCH]-[WI]-[LIVT]-x-[KRQAG]-N-[ST]-W-x(3)-[FYW]-G-x(2)-G-[LFYW]-[LIVMFYG]-x-[LIVMF], wherein N is the active site residue, and "x" represents any amino acid.

[0208] Additional information relating to eukaryotic thiol (cysteine) protease active site domains may be found in reference to the following publications, which are hereby incorporated by reference herein: Dufour E., Biochimie

70:1335-1342(1988); Kirschke H., Barrett A. J., Rawlings N. D., Protein Prof. 2:1587-1643(1995); Shi G.-P., Chapman H. A., Bhairi S. M., Deleeuw C., Reddy V. Y., Weiss S. J., FEBS Lett. 357:129-134(1995); Velasco G., Ferrando A. A., Puente X. S., Sanchez L. M., Lopez-Otin C., J. Biol. Chem. 269:27136-27142(1994); Chapot-Chartier M. P., Nardi M., Chopin M. C., Chopin A., Gripon J. C., Appl. Environ. Microbiol. 59:330-333(1993); Higgins D. G., McConnell D. J., Sharp P. M., Nature 340:604-604(1989); Rawlings N. D., Barrett A. J., Meth. Enzymol. 244:461-486(1994), and http://www.expasy.ch/cgi-bin/lists?peptidas.txt, which are hereby incorporated by reference in their entirety herein.

[0209] In preferred embodiments, the following eukaryotic thiol (cysteine) protease active site domain polypeptide is encompassed by the present invention: RTDVCQGSLGNCWFLAAAASLT (SEQ ID NO:54). Polynucleotides encoding these polypeptides are also provided. The present invention also encompasses the use of this EF-hand calcium binding domain polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0210] The present invention also provides a three-dimensional homology model of the Protease-42 polypeptide (see **FIG. 6**) representing amino acids M1 to L387 of Protease-42 (SEQ ID NO:2). Homology models are useful when there is no experimental information available on the protein of interest. A three-dimensional homology model can be constructed on the basis of the known structure of a homologous protein (Greer et al, 1991, Lesk, et al, 1992, Cardozo, et al, 1995, Yuan, et al, 1995). The homology model of the Protease-42 polypeptide was based upon the homologous structure of CAN2, a m-calpain family member (Strobl et al, 2000; (hCAN2; PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19) and is defined by the set of structural coordinates set forth in Table IV herein.

[0211] The Protease-42 homology model of the present invention may provide one basis for designing rational stimulators (agonists) and/or inhibitors (antagonists) of one or more of the biological functions of Protease-42, or of Protease-42 mutants having altered specificity (e.g., molecularly evolved Protease-42 polypeptides, engineered site-specific Protease-42 mutants, Protease-42 allelic variants, etc.).

[0212] Homology models are not only useful for designing rational agonists and/or antagonists, but are also useful in predicting the function of a particular polypeptide. The functional predictions from homology models are typically more accurate than the functional attributes derived from traditional polypeptide sequence homology alignments (e.g., CLUSTALW), particularly when the three dimensional structure of a related polypeptide is known (e.g., m-calpain family member hCAN2 protein; PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19). The increased prediction accuracy is based upon the fact that homology models approximate the three-dimensional structure of a protein, while homology based alignments only take into account the one dimension polypeptide sequence. Since the function of a particular polypeptide is determined not only by its primary, secondary, and tertiary structure, functional assignments derived solely upon homology alignments using the one dimensional protein sequence may be less reliable. A 3-dimensional model can be constructed on the basis of the known structure of a homologous protein (Greer et al, 1991, Lesk, et al, 1992, Cardozo, et al, 1995, Yuan, et al, 1995).

[0213] Prior to developing a homology model, those of skill in the art would appreciate that a template of a known protein, or model protein, must first be identified which will be used as a basis for constructing the homology model for the protein of unknown structure (query template). In the case of the Protease-42 polypeptide of the present invention, the model protein template used in constructing the Protease-42 homology model was the m-calpain family member hCAN2 (PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19).

[0214] Identifying a template can be accomplished using pairwise alignment of protein sequences using such programs as FASTA (Pearson, et al Methods In Enzymology 18363-98, 1990) and BLAST (Altschul, et al, J. Mol. Biol. 215, 403-10, 1990). In cases where sequence similarity is high (greater than 30%), such pairwise comparison methods may be adequate for identifying an appropriate template. Likewise, multiple sequence alignments or profile-based methods can be used to align a query sequence to an alignment of multiple (structurally and biochemically) related proteins. When the sequence similarity is low, more advanced techniques may be used. Such techniques, include, for example, protein fold recognition (protein threading; Hendlich, et al, 1990), where the compatibility of a particular polypeptide sequence with the 3-dimensional fold of a potential template protein is gauged on the basis of a knowledge-based potential.

[0215] Following the initial sequence alignment, an optional second step would be to optimally align the query template to the model template by manual manipulation and/or by the incorporation of features specific to the polypeptides (e.g., motifs, secondary structure predictions, and allowed conservations). Preferably, the incorporated features are found within both the model and query template.

[0216] The next step could be to identify structurally conserved regions that could be used to construct secondary core structure (Sali, et al, PROTEINS 23, 318-26 1995). Loops could be added using knowledge-based techniques, and by performing forcefield calculations (Sali, et al, 1995).

[0217] In order to recognize errors in a three-dimensional structure, knowledge based mean fields can be used to judge the quality of protein folds (Sippl 1993). The methods can be used to recognize misfolded structures as well as faulty parts of structural models. The technique generates an energy graph where the energy distribution for a given protein fold is displayed on the y-axis and residue position in the protein fold is displayed on the x-axis. The knowledge based mean fields compose a force field derived from a set of globular protein structures taken as a subset from the Protein Data Bank (Bernstein et. al. 1977). To analyze the quality of a model the energy distribution is plotted and compared to the energy distribution of the template from which the model was generated. **FIG. 7** shows the energy graph for the Protease-42 model (dotted line) and the template (1dkv, m-calpain) from which the model was generated. This graph supports the motif and sequence alignments in confirming that the three dimensional structure coordinates of Protease-42 are an accurate and useful representation for the polypeptide.

[0218] The term "structure coordinates" refers to Cartesian coordinates generated from the building of a homology model. In this invention, the homology model of residues M1 to L387 of Protease-42 (SEQ ID NO:2) was derived from generating a sequence alignment with m-calpain (hCAN2; PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19) using the COMPOSER suite of software within SYBYL6.7 (Tripos Associates, St. Louis, Mo.) and then generating the backbone and side chain conformations. In the original crystal structure (pdb code 1dkv) as well as the crystal structure reported elsewhere (Hosfield et al, 1999), the active site of the enzyme comprising a cysteine, a histidine and an asparagine residue was not "formed". The helix that contains the active site C105 was altered by moving the helix down one pitch so that the active site geometry could match that found in Papain (pdb code 1b4). This modified structure of human m-calpain was used as the template for construction of the homology model (illustrated in **FIG. 6** herein).

[0219] The skilled artisan would appreciate that a set of structure coordinates for a protein represents a relative set of points that define a shape in three dimensions. Thus, it is possible that an entirely different set of coordinates could define a similar or identical shape. Moreover, slight variations in the individual coordinates, as emanate from the generation of similar homology models using different alignment templates (i.e., other than the m-calpain (hCAN2; PDB code 1dkv; Genbank Accession No. gil6980465; SEQ ID NO:19), and/or using different methods in generating the homology model, will likely have minor effects on the overall shape. Variations in coordinates may also be generated because of mathematical manipulations of the structure coordinates. For example, the structure coordinates set forth in Table IV could be manipulated by fractionalization of the structure coordinates; integer additions, or integer subtractions to sets of the structure coordinates, inversion of the structure coordinates or any combination of the above.

[0220] Therefore, various computational analyses are necessary to determine whether a template molecule or a portion thereof is sufficiently similar to all or part of a query template (e.g., Protease-42) in order to be considered the same. Such analyses may be carried out in current software applications, such as SYBYL version 6.7 or INSIGHTII (Molecular Simulations Inc., San Diego, Calif.) version 2000 and as described in the accompanying User's Guides.

[0221] Using the superimposition tool in the program SYBYL, comparisons can be made between different structures and different conformations of the same structure. The procedure used in SYBYL to compare structures is divided into four steps: 1) load the structures to be compared; 2) define the atom equivalencies in these structures; 3) perform a fitting operation; and 4) analyze the results. Each structure is identified by a name. One structure is identified as the target (i.e., the fixed structure); the second structure (i.e., moving structure) is identified as the source structure. The atom equivalency within SYBYL is defined by user input. For the purpose of this invention, we will define equivalent atoms as protein backbone atoms (N, Cα, C and O) for all conserved residues between the two structures being compared. We will also consider only rigid fitting operations. When a rigid fitting method is used, the working structure is translated and rotated to obtain an optimum fit with the target structure. The fitting operation uses an algorithm that computes the optimum translation and rotation to be applied to the moving structure, such that the root mean square difference of the fit over the specified pairs of equivalent atoms is an absolute minimum. This number, given in angstroms, is reported by the SYBYL program. For the purpose of the present invention, any homology model of a Protease-42 that has a root mean square deviation of conserved residue backbone atoms (N, Cα, C, O) of less than 3.0 A when superimposed on the relevant backbone atoms described by structure coordinates listed in Table IV are considered identical. More preferably, the root mean square deviation for the Protease-42 polypeptide is less than 2.0 Å.

[0222] The homology model of the present invention is useful for the structure-based design of modulators of the Protease-42 biological function, as well as mutants with altered biological function and/or specificity.

[0223] In accordance with the structural coordinates provided in Table IV and the three dimensional homology model of Protease-42, the Protease-42 polypeptide has been shown to comprise a an active site region embodied by the following amino acids: from about amino acid S93 to about amino acid S 113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2 (FIGS. **1**A-C). In this context, the term "about" may be construed to mean 1, 2, 3, 4, 5, 6, 7, 8, 9 or 10 amino acids more in either the N- or C-terminal direction of the above referenced amino acids.

[0224] Also more preferred are polypeptides comprising all or any part of the Protease-42 active site domain, or a mutant or homologue of said polypeptide or molecular complex. By mutant or homologue of the molecule is meant a molecule that has a root mean square deviation from the backbone atoms of said Protease-42 amino acids of not more than about 4.5 Angstroms, and preferably not more than about 3.5 Angstroms.

[0225] In preferred embodiments, the following Protease-42 active site domain polypeptide is encompassed by the present invention: SRTDVCQGSLGNCWFLAAAASLTLYPRLLRRVVPPGQDFQHGYAGVFHFQLWQFGRWMDVVVDDRLPVREGKLMFVRSEQRNEFWAPLLEKAYAKLHGSYEVMRGGHMNEAF VDFTGGVGEVLYLRQNSMGLFSALRHALAKESLVGATALSDRGEYRTEEGLVKGHAYSITGTHKVFLGFTKVRLLRLRNPWGCVEWTGAWSDSCPRWDTLPTECRDALLVK EDGEFWMELRDFLLHFDTVQI (SEQ ID NO:65). Polynucleotides encoding this polypeptide are also provided. The present invention also encompasses the use of the Protease-42 active site domain polypeptide as an immunogenic and/or antigenic epitope as described elsewhere herein.

[0226] The present invention also encompasses polypeptides comprising at least a portion of the Protease-42 active site domain (SEQ ID NO:65). Such polypeptides may correspond, for example, to amino acids from about amino acid

G103 to about amino acid A110, at amino acid F167, from about amino acid T241 tio about amino acid A242, from about amino acid V256 to about amino acid S262, from about amino acid L281 to about amino acid V288, from about amino acid D318 to about amino acid G319, and at amino acid F321 of SEQ ID NO:2.

[**0227**] In preferred embodiments, the following Protease-42 active site domain amino acid substitutions are encompassed by the present invention: wherein S93 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein R94 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein T95 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein D96 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V97 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein C98 is substituted with either an A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Q99 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein G100 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S101 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein L102 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein G103 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein N104 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; wherein C105 is substituted with either an A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein W106 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein F107 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L108 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein A109 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A110 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A111 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A112 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S113 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein L114 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein T115 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein L116 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein Y117 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein P118 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein R119 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein L120 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein L121 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R122 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein R123 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein V124 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein V125 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein P126 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein P127 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein G128 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Q129 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein D130 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F131 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Q132 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein H133 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G134 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Y135 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein A136 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G137 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V138 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein F139 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein H140 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F141 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Q142 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein L143 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein W144 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein Q145 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein F146 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G147 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein R148 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein W149 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein M150 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein D151 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V152 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein V153 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein V154 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein D155 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein D156 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein R157 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein L158 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein P159 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein V160 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein R161 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein E162 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R. S, T, V, W, or Y; wherein G163 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K164 is substituted with either an A, C,

D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L165 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein M166 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein F167 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V168 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein R169 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein S170 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein E171 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Q172 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein R173 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein N174 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; wherein E175 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F176 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein W177 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein A178 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein P179 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein L180 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein L181 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein E182 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K183 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A184 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Y185 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein A186 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K187 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L188 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein H189 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G190 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S191 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein Y192 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein E193 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V194 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein M195 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein R196 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein G197 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G198 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein H199 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein M200 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein N201 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; wherein E202 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A203 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F204 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V205 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein D206 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F207 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T208 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein G209 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G210 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V211 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein G212 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E213 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V214 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein L215 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein Y216 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein L217 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R218 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein Q219 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; wherein N220 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; wherein S221 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein M222 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein G223 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L224 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein F225 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S226 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein A227 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L228 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R229 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein H230 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A231 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L232 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein A233 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K234 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E235 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S236 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein L237 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein V238 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein G239 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A240 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T,

V, W, or Y; wherein T241 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein A242 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L243 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein S244 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein D245 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein R246 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein G247 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E248 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Y249 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein R250 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein T251 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein E252 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E253 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G254 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L255 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein V256 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein K257 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G258 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein H259 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A260 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein Y261 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or W; wherein S262 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein I263 is substituted with either an A, C, D, E, F, G, H, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T264 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein G265 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T266 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein H267 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K268 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V269 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein F270 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L271 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein G272 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F273 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T274 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein K275 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V276 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein R277 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein L278 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein L279 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R280 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein L281 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R282 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein N283 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, P, Q, R, S, T, V, W, or Y; wherein P284 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein W285 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein G286 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein C287 is substituted with either an A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein V288 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein E289 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein W290 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein T291 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein G292 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A293 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein W294 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein S295 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein D296 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein S297 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, T, V, W, or Y; wherein C298 is substituted with either an A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein P299 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein R300 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein W301 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein D302 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T303 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein L304 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein P305 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W, or Y; wherein T306 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein E307 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein C308 is substituted with either an A, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein R309 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein D310 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein A311 is substituted with either a C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L312 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein L313 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein V314 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein K315 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein K316 is substituted with either an A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E317 is substituted with either an A, C, D, F, G, H, I, K, L,

M, N, P, Q, R, S, T, V, W, or Y; wherein D318 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein G319 is substituted with either an A, C, D, E, F, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein E320 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F321 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein W322 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, or Y; wherein M323 is substituted with either an A, C, D, E, F, G, H, I, K, L, N, P, Q, R, S, T, V, W, or Y; wherein E324 is substituted with either an A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L325 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein R326 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, S, T, V, W, or Y; wherein D327 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F328 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein L329 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein L330 is substituted with either an A, C, D, E, F, G, H, I, K, M, N, P, Q, R, S, T, V, W, or Y; wherein H331 is substituted with either an A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein F332 is substituted with either an A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein D333 is substituted with either an A, C, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, or Y; wherein T334 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, V, W, or Y; wherein V335 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, W, or Y; wherein Q336 is substituted with either an A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W, or Y; and/or wherein I337 is substituted with either an A, C, D, E, F, G, H, K, L, M, N, P, Q, R, S, T, V, W, or Y of SEQ ID NO:2, in addition to any combination thereof. The present invention also encompasses the use of these Protease-42 active site domain amino acid substituted polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[**0228**] In preferred embodiments, the following Protease-42 active site domain conservative amino acid substitutions are encompassed by the present invention: wherein S93 is substituted with either an A, G, M, or T; wherein R94 is substituted with either a K, or H; wherein T95 is substituted with either an A, G, M, or S; wherein D96 is substituted with an E; wherein V97 is substituted with either an A, I, or L; wherein C98 is a C; wherein Q99 is substituted with a N; wherein G100 is substituted with either an A, M, S, or T; wherein S101 is substituted with either an A, G, M, or T; wherein L102 is substituted with either an A, I, or V; wherein G103 is substituted with either an A, M, S, or T; wherein N104 is substituted with a Q; wherein C105 is a C; wherein W106 is either an F, or Y; wherein F107 is substituted with either a W, or Y; wherein L108 is substituted with either an A, I, or V; wherein A109 is substituted with either a G, I, L, M, S, T, or V; wherein A110 is substituted with either a G, I, L, M, S, T, or V; wherein A111 is substituted with either a G, I, L, M, S, T, or V; wherein A112 is substituted with either a G, I, L, M, S, T, or V; wherein S113 is substituted with either an A, G, M, or T; wherein L114 is substituted with either an A, I, or V; wherein T115 is substituted with either an A, G, M, or S; wherein L116 is substituted with either an A, I, or V; wherein Y 117 is either an F, or W; wherein P118 is a P; wherein R119 is substituted with either a K, or H; wherein L120 is substituted with either an A, I, or V; wherein L121 is substituted with either an A, I, or V; wherein R122 is substituted with either a K, or H; wherein R123 is substituted with either a K, or H; wherein V124 is substituted with either an A, I, or L; wherein V125 is substituted with either an A, I, or L; wherein P126 is a P; wherein P127 is a P; wherein G128 is substituted with either an A, M, S, or T; wherein Q129 is substituted with a N; wherein D130 is substituted with an E; wherein F131 is substituted with either a W, or Y; wherein Q132 is substituted with a N; wherein H133 is substituted with either a K, or R; wherein G134 is substituted with either an A, M, S, or T; wherein Y135 is either an F, or W; wherein A136 is substituted with either a G, I, L, M, S, T, or V; wherein G137 is substituted with either an A, M, S, or T; wherein V138 is substituted with either an A, I, or L; wherein F139 is substituted with either a W, or Y; wherein H140 is substituted with either a K, or R; wherein F141 is substituted with either a W, or Y; wherein Q142 is substituted with a N; wherein L143 is substituted with either an A, I, or V; wherein W144 is either an F, or Y; wherein Q145 is substituted with a N; wherein F146 is substituted with either a W, or Y; wherein G147 is substituted with either an A, M, S, or T; wherein R148 is substituted with either a K, or H; wherein W149 is either an F, or Y; wherein M150 is substituted with either an A, G, S, or T; wherein D151 is substituted with an E; wherein V152 is substituted with either an A, I, or L; wherein V153 is substituted with either an A, I, or L; wherein V154 is substituted with either an A, I, or L; wherein D155 is substituted with an E; wherein D156 is substituted with an E; wherein R157 is substituted with either a K, or H; wherein L158 is substituted with either an A, I, or V; wherein P159 is a P; wherein V160 is substituted with either an A, I, or L; wherein R161 is substituted with either a K, or H; wherein E162 is substituted with a D; wherein G163 is substituted with either an A, M, S, or T; wherein K164 is substituted with either a R, or H; wherein L165 is substituted with either an A, I, or V; wherein M166 is substituted with either an A, G, S, or T; wherein F167 is substituted with either a W, or Y; wherein V168 is substituted with either an A, I, or L; wherein R169 is substituted with either a K, or H; wherein S 170 is substituted with either an A, G, M, or T; wherein E171 is substituted with a D; wherein Q172 is substituted with a N; wherein R173 is substituted with either a K, or H; wherein N174 is substituted with a Q; wherein E175 is substituted with a D; wherein F176 is substituted with either a W, or Y; wherein W177 is either an F, or Y; wherein A178 is substituted with either a G, I, L, M, S, T, or V; wherein P179 is a P; wherein L180 is substituted with either an A, I, or V; wherein L181 is substituted with either an A, I, or V; wherein E182 is substituted with a D; wherein K183 is substituted with either a R, or H; wherein A184 is substituted with either a G, I, L, M, S, T, or V; wherein Y185 is either an F, or W; wherein A186 is substituted with either a G, I, L, M, S, T, or V; wherein K187 is substituted with either a R, or H; wherein L188 is substituted with either an A, I, or V; wherein H189 is substituted with either a K, or R; wherein G190 is substituted with either an A, M, S, or T; wherein S191 is substituted with either an A, G, M, or T; wherein Y192 is either an F, or W; wherein E193 is substituted with a D; wherein V194 is substituted with either an A, I, or L; wherein M195 is substituted with either an A, G, S, or T; wherein R196 is substituted with either a K, or H; wherein

G197 is substituted with either an A, M, S, or T; wherein G198 is substituted with either an A, M, S, or T; wherein H199 is substituted with either a K, or R; wherein M200 is substituted with either an A, G, S, or T; wherein N201 is substituted with a Q; wherein E202 is substituted with a D; wherein A203 is substituted with either a G, I, L, M, S, T, or V; wherein F204 is substituted with either a W, or Y; wherein V205 is substituted with either an A, I, or L; wherein D206 is substituted with an E; wherein F207 is substituted with either a W, or Y; wherein T208 is substituted with either an A, G, M, or S; wherein G209 is substituted with either an A, M, S, or T; wherein G210 is substituted with either an A, M, S, or T; wherein V211 is substituted with either an A, I, or L; wherein G212 is substituted with either an A, M, S, or T; wherein E213 is substituted with a D; wherein V214 is substituted with either an A, I, or L; wherein L215 is substituted with either an A, I, or V; wherein Y216 is either an F, or W; wherein L217 is substituted with either an A, I, or V; wherein R218 is substituted with either a K, or H; wherein Q219 is substituted with a N; wherein N220 is substituted with a Q; wherein S221 is substituted with either an A, G, M, or T; wherein M222 is substituted with either an A, G, S, or T; wherein G223 is substituted with either an A, M, S, or T; wherein L224 is substituted with either an A, I, or V; wherein F225 is substituted with either a W, or Y; wherein S226 is substituted with either an A, G, M, or T; wherein A227 is substituted with either a G, I, L, M, S, T, or V; wherein L228 is substituted with either an A, I, or V; wherein R229 is substituted with either a K, or H; wherein H230 is substituted with either a K, or R; wherein A231 is substituted with either a G, I, L, M, S, T, or V; wherein L232 is substituted with either an A, I, or V; wherein A233 is substituted with either a G, I, L, M, S, T, or V; wherein K234 is substituted with either a R, or H; wherein E235 is substituted with a D; wherein S236 is substituted with either an A, G, M, or T; wherein L237 is substituted with either an A, I, or V; wherein V238 is substituted with either an A, I, or L; wherein G239 is substituted with either an A, M, S, or T; wherein A240 is substituted with either a G, I, L, M, S, T, or V; wherein T241 is substituted with either an A, G, M, or S; wherein A242 is substituted with either a G, I, L, M, S, T, U or V; wherein L243 is substituted with either an A, I, or V; wherein S244 is substituted with either an A, G, M, or T; wherein D245 is substituted with an E; wherein R246 is substituted with either a K, or H; wherein G247 is substituted with either an A, M, S, or T; wherein E248 is substituted with a D; wherein Y249 is either an F, or W; wherein R250 is substituted with either a K, or H; wherein T251 is substituted with either an A, G, M, or S; wherein E252 is substituted with a D; wherein E253 is substituted with a D; wherein G254 is substituted with either an A, M, S, or T; wherein L255 is substituted with either an A, T, or V; wherein V256 is substituted with either an A, I, or L; wherein K257 is substituted with either a R, or H; wherein G258 is substituted with either an A, M, S, or T; wherein H259 is substituted with either a K, or R; wherein A260 is substituted with either a G, I, L, M, S, T, or V; wherein Y261 is either an F, or W; wherein S262 is substituted with either an A, G, M, or T; wherein I263 is substituted with either an A, V, or L; wherein T264 is substituted with either an A, G, M, or S; wherein G265 is substituted with either an A, M, S, or T; wherein T266 is substituted with either an A, G, M, or S; wherein H267 is substituted with either a K, or R; wherein K268 is substituted with either a R, or H; wherein V269 is substituted with either an A, I, or L; wherein F270 is substituted with either a W, or Y; wherein L271 is substituted with either an A, I, or V; wherein G272 is substituted with either an A, M, S, or T; wherein F273 is substituted with either a W, or Y; wherein T274 is substituted with either an A, G, M, or S; wherein K275 is substituted with either a R, or H; wherein V276 is substituted with either an A, I, or L; wherein R277 is substituted with either a K, or H; wherein L278 is substituted with either an A, I, or V; wherein L279 is substituted with either an A, I, or V; wherein R280 is substituted with either a K, or H; wherein L281 is substituted with either an A, I, or V; wherein R282 is substituted with either a K, or H; wherein N283 is substituted with a Q; wherein P284 is a P; wherein W285 is either an F, or Y; wherein G286 is substituted with either an A, M, S, or T; wherein C287 is a C; wherein V288 is substituted with either an A, I, or L; wherein E289 is substituted with a D; wherein W290 is either an F, or Y; wherein T291 is substituted with either an A, G, M, or S; wherein G292 is substituted with either an A, M, S, or T; wherein A293 is substituted with either a G, I, L, M, S, T, or V; wherein W294 is either an F, or Y; wherein S295 is substituted with either an A, G, M, or T; wherein D296 is substituted with an E; wherein S297 is substituted with either an A, G, M, or T; wherein C298 is a C; wherein P299 is a P; wherein R300 is substituted with either a K, or H; wherein W301 is either an F, or Y; wherein D302 is substituted with an E; wherein T303 is substituted with either an A, G, M, or S; wherein L304 is substituted with either an A, I, or V; wherein P305 is a P; wherein T306 is substituted with either an A, G, M, or S; wherein E307 is substituted with a D; wherein C308 is a C; wherein R309 is substituted with either a K, or H; wherein D310 is substituted with an E; wherein A311 is substituted with either a G, I, L, M, S, T, or V; wherein L312 is substituted with either an A, I, or V; wherein L313 is substituted with either an A, I, or V; wherein V314 is substituted with either an A, I, or L; wherein K315 is substituted with either a R, or H; wherein K316 is substituted with either a R, or H; wherein E317 is substituted with a D; wherein D318 is substituted with an E; wherein G319 is substituted with either an A, M, S, or T; wherein E320 is substituted with a D; wherein F321 is substituted with either a W, or Y; wherein W322 is either an F, or Y; wherein M323 is substituted with either an A, G, S, or T; wherein E324 is substituted with a D; wherein L325 is substituted with either an A, I, or V; wherein R326 is substituted with either a K, or H; wherein D327 is substituted with an E; wherein F328 is substituted with either a W, or Y; wherein L329 is substituted with either an A, I, or V; wherein L330 is substituted with either an A, I, or V; wherein H331 is substituted with either a K, or R; wherein F332 is substituted with either a W, or Y; wherein D333 is substituted with an E; wherein T334 is substituted with either an A, G, M, or S; wherein V335 is substituted with either an A, I, or L; wherein Q336 is substituted with a N; and/or wherein I337 is substituted with either an A, V, or L of SEQ ID NO:2 in addition to any combination thereof. Other suitable substitutions within the Protease-42 active site domain are encompassed by the present invention and are referenced elsewhere herein. The present invention also encompasses the use of these Protease-42 active site domain conservative amino acid substituted polypeptides as immunogenic and/or antigenic epitopes as described elsewhere herein.

[0229] For purposes of the present invention, by "at least a portion of" is meant all or any part of the Protease-42 active site domain defined by the structure coordinates according to Table IV (e.g., fragments thereof). More preferred are molecules comprising all or any parts of the Protease-42 active site domain, according to Table IV, or a mutant or homologue of said molecule or molecular complex. By mutant or homologue of the molecule it is meant a molecule that has a root mean square deviation from the backbone atoms of said Protease-42 amino acids of not more than 4.5 Angstroms, and preferably not more than 3.5 Angstroms.

[0230] The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations from the mean. It is a term that expresses the deviation or variation from a trend or object. For the purposes of the present invention, the "root mean square deviation" defines the variation in the backbone of a protein from the relevant portion of the backbone of the AR portion of the complex as defined by the structure coordinates described herein.

[0231] A preferred embodiment is a machine-readable data storage medium that is capable of displaying a graphical three-dimensional representation of a molecule or molecular complex that is defined by the structure coordinates of all of the amino acids in Table IV+/– a root mean square deviation from the backbone atoms of those amino acids of not more than 4.0 ANG, preferably 3.0 ANG.

[0232] The structure coordinates of a Protease-42 homology model, including portions thereof, is stored in a machine-readable storage medium. Such data may be used for a variety of purposes, such as drug discovery.

[0233] Accordingly, in one embodiment of this invention is provided a machine-readable data storage medium comprising a data storage material encoded with the structure coordinates set forth in Table IV.

[0234] One embodiment utilizes System 10 as disclosed in WO 98/11134, the disclosure of which is incorporated herein by reference in its entirety. Briefly, one version of these embodiments comprises a computer comprising a central processing unit ("CPU"), a working memory which may be, e.g, RAM (random-access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more cathode-ray tube ("CRT") display terminals, one or more keyboards, one or more input lines, and one or more output lines, all of which are interconnected by a conventional bidirectional system bus.

[0235] Input hardware, coupled to the computer by input lines, may be implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may comprise CD-ROM drives or disk drives. In conjunction with a display terminal, keyboard may also be used as an input device.

[0236] Output hardware, coupled to the computer by output lines, may similarly be implemented by conventional devices. By way of example, output hardware may include a CRT display terminal for displaying a graphical representation of a region or domain of the present invention using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

[0237] In operation, the CPU coordinates the use of the various input and output devices, coordinates data accesses from mass storage, and accesses to and from the working memory, and determines the sequence of data processing steps. A number of programs may be used to process the machine-readable data of this invention. Such programs are discussed in reference to the computational methods of drug discovery as described herein. Specific references to components of the hardware system are included as appropriate throughout the following description of the data storage medium.

[0238] For the purpose of the present invention, any magnetic data storage medium which can be encoded with machine-readable data would be sufficient for carrying out the storage requirements of the system. The medium could be a conventional floppy diskette or hard disk, having a suitable substrate, which may be conventional, and a suitable coating, which may be conventional, on one or both sides, containing magnetic domains whose polarity or orientation could be altered magnetically, for example. The medium may also have an opening for receiving the spindle of a disk drive or other data storage device.

[0239] The magnetic domains of the coating of a medium may be polarized or oriented so as to encode in a manner which may be conventional, machine readable data such as that described herein, for execution by a system such as the system described herein.

[0240] Another example of a suitable storage medium which could also be encoded with such machine-readable data, or set of instructions, which could be carried out by a system such as the system described herein, could be an optically-readable data storage medium. The medium could be a conventional compact disk read only memory (CD-ROM) or a rewritable medium such as a magneto-optical disk which is optically readable and magneto-optically writable. The medium preferably has a suitable substrate, which may be conventional, and a suitable coating, which may be conventional, usually of one side of substrate.

[0241] In the case of a CD-ROM, as is well known, the coating is reflective and is impressed with a plurality of pits to encode the machine-readable data. The arrangement of pits is read by reflecting laser light off the surface of the coating. A protective coating, which preferably is substantially transparent, is provided on top of the reflective coating.

[0242] In the case of a magneto-optical disk, as is well known, the coating has no pits, but has a plurality of magnetic domains whose polarity or orientation can be changed magnetically when heated above a certain temperature, as by a laser. The orientation of the domains can be read by measuring the polarization of laser light reflected from the coating. The arrangement of the domains encodes the data as described above.

[0243] Thus, in accordance with the present invention, data capable of displaying the three dimensional structure of the Protease-42 homology model, or portions thereof and their structurally similar homologues is stored in a machine-readable storage medium, which is capable of displaying a

graphical three-dimensional representation of the structure. Such data may be used for a variety of purposes, such as drug discovery.

[0244] For the first time, the present invention permits the use of structure-based or rational drug design techniques to design, select, and synthesize chemical entities that are capable of modulating the biological function of Protease-42.

[0245] Accordingly, the present invention is also directed to the design of small molecules which imitates the structure of the Protease-42 active site domain (SEQ ID NO:65), or a portion thereof, in accordance with the structure coordinates provided in Table IV. Alternatively, the present invention is directed to the design of small molecules which may bind to at least part of the Protease-42 active site domain (SEQ ID NO:65), or some portion thereof. For purposes of this invention, by Protease-42 active site domain, it is also meant to include mutants or homologues thereof. In a preferred embodiment, the mutants or homologues have at least 25% identity, more preferably 50% identity, more preferably 75% identity, and most preferably 90% identity to SEQ ID NO:65. In this context, the term "small molecule" may be construed to mean any molecule described known in the art or described elsewhere herein, though may include, for example, peptides, chemicals, carbohydrates, nucleic acids, PNAs, and any derivatives thereof.

[0246] The three-dimensional model structure of the Protease-42 will also provide methods for identifying modulators of biological function. Various methods or combination thereof can be used to identify these compounds.

[0247] For example, test compounds can be modeled that fit spatially into the active site domain in Protease-42 embodied by the sequence from about from about amino acid S93 to about amino acid S113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2 (corresponding to SEQ ID NO:65), in accordance with the structural coordinates of Table IV.

[0248] Structure coordinates of the active site domain in Protease-42 defined by the amino acids from about from about amino acid S93 to about amino acid S113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2, can also be used to identify structural and chemical features. Identified structural or chemical features can then be employed to design or select compounds as potential Protease-42 modulators. By structural and chemical features it is meant to include, but is not limited to, van der Waals interactions, hydrogen bonding interactions, charge interaction, hydrophobic bonding interaction, and dipole interaction. Alternatively, or in conjunction with, the three-dimensional structural model can be employed to design or select compounds as potential Protease-42 modulators. Compounds identified as potential Protease-42 modulators can then be synthesized and screened in an assay characterized by binding of a test compound to the Protease-42, or in characterizing the ability of Protease-42 to modulate a protease target in the presence of a small molecule. Examples of assays useful in screening of potential Protease-42 modulators include, but are not limited to, screening in silico, in vitro assays and high throughput assays. Finally, these methods may also involve modifying or replacing one or more amino acids at amino acid positions, Cys105, His259 and/or Asn283 of SEQ ID NO:2 in accordance with the structure coordinates of Table IV.

[0249] However, as will be understood by those of skill in the art upon this disclosure, other structure based design methods can be used. Various computational structure based design methods have been disclosed in the art.

[0250] For example, a number of computer modeling systems are available in which the sequence of the Protease-42 and the Protease-42 structure (i.e., atomic coordinates of Protease-42 and/or the atomic coordinates of the active site domain as provided in Table IV) can be input. This computer system then generates the structural details of one or more these regions in which a potential Protease-42 modulator binds so that complementary structural details of the potential modulators can be determined. Design in these modeling systems is generally based upon the compound being capable of physically and structurally associating with Protease-42. In addition, the compound must be able to assume a conformation that allows it to associate with Protease-42. Some modeling systems estimate the potential inhibitory or binding effect of a potential Protease-42 modulator prior to actual synthesis and testing.

[0251] Methods for screening chemical entities or fragments for their ability to associate with a given protein target are also well known. Often these methods begin by visual inspection of the binding site on the computer screen. Selected fragments or chemical entities are then positioned in the active site domain of Protease-42. Docking is accomplished using software such as INSIGHTII, QUANTA and SYBYL, following by energy minimization and molecular dynamics with standard molecular mechanic forcefields such as MMFF, CHARMM and AMBER. Examples of computer programs which assist in the selection of chemical fragment or chemical entities useful in the present invention include, but are not limited to, GRID (Goodford, 1985), AUTODOCK (Goodsell, 1990), and DOCK (Kuntz et al. 1982).

[0252] Upon selection of preferred chemical entities or fragments, their relationship to each other and Protease-42 can be visualized and then assembled into a single potential modulator. Programs useful in assembling the individual chemical entities include, but are not limited to CAVEAT

(Bartlett et al. 1989) and 3D Database systems (Martin 1992).

[0253] Alternatively, compounds may be designed de novo using either an empty active site or optionally including some portion of a known inhibitor. Methods of this type of design include, but are not limited to LUDI (Bohm 1992) and LeapFrog (Tripos Associates, St. Louis Mo.).

[0254] In addition, Protease-42 is overall well suited to modern methods including combinatorial chemistry.

[0255] Programs such as DOCK (Kuntz et al. 1982) can be used with the atomic coordinates from the homology model to identify potential ligands from databases or virtual databases which potentially bind Protease-42 active site domain, and which may therefore be suitable candidates for synthesis and testing.

[0256] Additionally, the three-dimensional homology model of Protease-42 will aid in the design of mutants with altered biological activity.

[0257] The following are encompassed by the present invention: a machine-readable data storage medium, comprising a data storage material encoded with machine readable data, wherein the data is defined by the structure coordinates of the model Protease-42 according to Table IV or a homologue of said model, wherein said homologue comprises backbone atoms that have a root mean square deviation from the backbone atoms of the complex of not more than 4.5 Å, preferably not more than 4.0 Å, most preferably not more than 3.5 Å, and even more preferably not more than 3.0 Å; and a machine-readable data storage medium, wherein said molecule is defined by the set of structure coordinates of the model for Protease-42 according to Table IV, or a homologue of said molecule, said homologue having a root mean square deviation from the backbone atoms of said amino acids of not more than 4.5 Å, preferably not more than 4.0 Å, most preferably not more than 3.5 Å, and even more preferably not more than 3.0 Å; a model comprising all or any part of the model defined by structure coordinates of Protease-42 according to Table IV, or a mutant or homologue of said molecule or molecular complex.

[0258] In a further embodiment, the following are encompassed by the present invention: a method for identifying a mutant of Protease-42 with altered biological properties, function, or reactivity, the method comprising any combination of steps of: use of the model or a homologue of said model according to Table IV, for the design of protein mutants with altered biological function or properties which exhibit any combination of therapeutic effects provided elsewhere herein; and use of the model or a homologue of said model, for the design of a protein with mutations in the active site domain comprised of the amino acids from about from about amino acid S93 to about amino acid S113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2 according to Table IV with altered biological function or properties which exhibit any combination of therapeutic effects provided elsewhere herein.

[0259] In further preferred embodiments, the following are encompassed by the present invention: a method for identifying modulators of Protease-42 biological properties, function, or reactivity, the method comprising any combination of steps of: modeling test compounds that overlay spatially into the active site domain defined by all or any portion of residues from about from about amino acid S93 to about amino acid S113, amino acid L121, amino acid V168, amino acid W177, amino acid E182, from about amino acid H199 to about amino acid N201, from about amino acid A203 to about amino acid F204, from about amino acid F207 to about amino acid L217, amino acid R218, from about amino acid L237 to about amino acid R246, from about amino acid L255 to about amino acid G265, from about amino acid L279 to about amino acid W290, from about amino acid K316 to about amino acid M323, and/or from about amino acid L330 to about amino acid I337 of SEQ ID NO:2 and of the three-dimensional structural model according to Table IV, or using a homologue or portion thereof.

[0260] The present invention encompasses using the structure coordinates as set forth herein to identify structural and chemical features of the Protease-42 polypeptide; employing identified structural or chemical features to design or select compounds as potential Protease-42 modulators; employing the three-dimensional structural model to design or select compounds as potential Protease-42 modulators; synthesizing the potential Protease-42 modulators; screening the potential Protease-42 modulators in an assay characterized by binding of a protein to the Protease-42; selecting the potential Protease-42 modulator from a database; designing the Protease-42 modulator de novo; and/or designing said Protease-42 modulator from a known modulator activity.

[0261] Many polynucleotide sequences, such as EST sequences, are publicly available and accessible through sequence databases. Some of these sequences are related to SEQ ID NO:1 and may have been publicly available prior to conception of the present invention. Preferably, such related polynucleotides are specifically excluded from the scope of the present invention. To list every related sequence would be cumbersome. Accordingly, preferably excluded from the present invention are one or more polynucleotides consisting of a nucleotide sequence described by the general formula of a-b, where a is any integer between 1 to 2206 of SEQ ID NO:1, b is an integer between 15 to 2220, where both a and b correspond to the positions of nucleotide residues shown in SEQ ID NO:1, and where b is greater than or equal to a+14.

TABLE I

Gene No.	CDNA CloneID	ATCC Deposit No. Z and Date	Vector	NT SEQ ID. No. X	Total NT Seq of Clone	5' NT of Start Codon of ORF	3' NT of ORF	AA Seq ID No. Y	Total AA of ORF
1.	Protease-42	PTA-3745 Oct. 01, 2001	Psport 1	1	2220	3	2207	2	735

[0262] Table I summarizes the information corresponding to each "Gene No." described above. The nucleotide sequence identified as "NT SEQ ID NO:1" was assembled from partially homologous ("overlapping") sequences obtained from the "cDNA clone ID" identified in Table 1 and, in some cases, from additional related DNA clones. The overlapping sequences were assembled into a single contiguous sequence of high redundancy (usually several overlapping sequences at each nucleotide position), resulting in a final sequence identified as SEQ ID NO:1.

[0263] The cDNA Clone ID was deposited on the date and given the corresponding deposit number listed in "ATCC Deposit No:Z and Date." "Vector" refers to the type of vector contained in the cDNA Clone ID. "Total NT Seq. Of Clone" refers to the total number of nucleotides in the clone contig identified by "Gene No." The deposited clone may contain all or most of the sequence of SEQ ID NO:1. The nucleotide position of SEQ ID NO:1 of the putative start codon (methionine) is identified as "5' NT of Start Codon of ORF."

[0264] The translated amino acid sequence, beginning with the methionine, is identified as "AA SEQ ID NO:2," although other reading frames can also be easily translated using known molecular biology techniques. The polypeptides produced by these alternative open reading frames are specifically contemplated by the present invention.

[0265] The total number of amino acids within the open reading frame of SEQ ID NO:2 is identified as "Total AA of ORF".

[0266] SEQ ID NO:1 (where X may be any of the polynucleotide sequences disclosed in the sequence listing) and the translated SEQ ID NO:2 (where Y may be any of the polypeptide sequences disclosed in the sequence listing) are sufficiently accurate and otherwise suitable for a variety of uses well known in the art and described further herein. For instance, SEQ ID NO:1 is useful for designing nucleic acid hybridization probes that will detect nucleic acid sequences contained in SEQ ID NO:1 or the cDNA contained in the deposited clone. These probes will also hybridize to nucleic acid molecules in biological samples, thereby enabling a variety of forensic and diagnostic methods of the invention. Similarly, polypeptides identified from SEQ ID NO:2 may be used, for example, to generate antibodies which bind specifically to proteins containing the polypeptides and the proteins encoded by the cDNA clones identified in Table 1.

[0267] Nevertheless, DNA sequences generated by sequencing reactions can contain sequencing errors. The errors exist as misidentified nucleotides, or as insertions or deletions of nucleotides in the generated DNA sequence. The erroneously inserted or deleted nucleotides may cause frame shifts in the reading frames of the predicted amino acid sequence. In these cases, the predicted amino acid sequence diverges from the actual amino acid sequence, even though the generated DNA sequence may be greater than 99.9% identical to the actual DNA sequence (for example, one base insertion or deletion in an open reading frame of over 1000 bases).

[0268] Accordingly, for those applications requiring precision in the nucleotide sequence or the amino acid sequence, the present invention provides not only the generated nucleotide sequence identified as SEQ ID NO:1 and the predicted translated amino acid sequence identified as SEQ ID NO:2, but also a sample of plasmid DNA containing a cDNA of the invention deposited with the ATCC, as set forth in Table 1. The nucleotide sequence of each deposited clone can readily be determined by sequencing the deposited clone in accordance with known methods. The predicted amino acid sequence can then be verified from such deposits. Moreover, the amino acid sequence of the protein encoded by a particular clone can also be directly determined by peptide sequencing or by expressing the protein in a suitable host cell containing the deposited cDNA, collecting the protein, and determining its sequence.

[0269] The present invention also relates to the genes corresponding to SEQ ID NO:1, SEQ ID NO:2, or the deposited clone. The corresponding gene can be isolated in accordance with known methods using the sequence information disclosed herein. Such methods include preparing probes or primers from the disclosed sequence and identifying or amplifying the corresponding gene from appropriate sources of genomic material.

[0270] Also provided in the present invention are species homologs, allelic variants, and/or orthologs. The skilled artisan could, using procedures well-known in the art, obtain the polynucleotide sequence corresponding to full-length genes (including, but not limited to the full-length coding region), allelic variants, splice variants, orthologs, and/or species homologues of genes corresponding to SEQ ID NO:1, SEQ ID NO:2, or a deposited clone, relying on the sequence from the sequences disclosed herein or the clones deposited with the ATCC. For example, allelic variants and/or species homologues may be isolated and identified by making suitable probes or primers which correspond to the 5', 3', or internal regions of the sequences provided herein and screening a suitable nucleic acid source for allelic variants and/or the desired homologue.

[0271] The polypeptides of the invention can be prepared in any suitable manner. Such polypeptides include isolated naturally occurring polypeptides, recombinantly produced polypeptides, synthetically produced polypeptides, or polypeptides produced by a combination of these methods. Means for preparing such polypeptides are well understood in the art.

[0272] The polypeptides may be in the form of the protein, or may be a part of a larger protein, such as a fusion protein (see below). It is often advantageous to include an additional amino acid sequence which contains secretory or leader

sequences, pro-sequences, sequences which aid in purification, such as multiple histidine residues, or an additional sequence for stability during recombinant production.

[0273] The polypeptides of the present invention are preferably provided in an isolated form, and preferably are substantially purified. A recombinantly produced version of a polypeptide, can be substantially purified using techniques described herein or otherwise known in the art, such as, for example, by the one-step method described in Smith and Johnson, Gene 67:31-40 (1988). Polypeptides of the invention also can be purified from natural, synthetic or recombinant sources using protocols described herein or otherwise known in the art, such as, for example, antibodies of the invention raised against the full-length form of the protein.

[0274] The present invention provides a polynucleotide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:1, and/or a cDNA provided in ATCC Deposit No:Z. The present invention also provides a polypeptide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:2, and/or a polypeptide encoded by the cDNA provided in ATCC Deposit NO:Z. The present invention also provides polynucleotides encoding a polypeptide comprising, or alternatively consisting of the polypeptide sequence of SEQ ID NO:2, and/or a polypeptide sequence encoded by the cDNA contained in ATCC Deposit No:Z.

[0275] Preferably, the present invention is directed to a polynucleotide comprising, or alternatively consisting of, the sequence identified as SEQ ID NO:1, and/or a cDNA provided in ATCC Deposit No:Z that is less than, or equal to, a polynucleotide sequence that is 5 mega basepairs, 1 mega basepairs, 0.5 mega basepairs, 0.1 mega basepairs, 50,000 basepairs, 20,000 basepairs, or 10,000 basepairs in length.

[0276] The present invention encompasses polynucleotides with sequences complementary to those of the polynucleotides of the present invention disclosed herein. Such sequences may be complementary to the sequence disclosed as SEQ ID NO:1, the sequence contained in a deposit, and/or the nucleic acid sequence encoding the sequence disclosed as SEQ ID NO:2.

[0277] The present invention also encompasses polynucleotides capable of hybridizing, preferably under reduced stringency conditions, more preferably, under stringent conditions, and most preferably under highly stingent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in Table II below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

TABLE II

Stringency Condition	Polynucleotide Hybrid±	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	> or equal to 50	65° C.; 1xSSC - or- 42° C.; 1xSSC, 50% formamide	65° C.; 0.3xSSC
B	DNA:DNA	<50	Tb*; 1xSSC	Tb*; 1xSSC

TABLE II-continued

Stringency Condition	Polynucleotide Hybrid±	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
C	DNA:RNA	> or equal to 50	67° C.; 1xSSC - or- 45° C.; 1xSSC, 50% formamide	67° C.; 0.3xSSC
D	DNA:RNA	<50	Td*; 1xSSC	Td*; 1xSSC
E	RNA:RNA	> or equal to 50	70° C.; 1xSSC - or- 50° C.; 1xSSC, 50% formamide	70° C.; 0.3xSSC
F	RNA:RNA	<50	Tf*; 1xSSC	Tf*; 1xSSC
G	DNA:DNA	> or equal to 50	65° C.; 4xSSC - or- 45° C.; 4xSSC, 50% formamide	65° C.; 1xSSC
H	DNA:DNA	<50	Th*; 4xSSC	Th*; 4xSSC
I	DNA:RNA	> or equal to 50	67° C.; 4xSSC - or- 45° C.; 4xSSC, 50% formamide	67° C.; 1xSSC
J	DNA:RNA	<50	Tj*; 4xSSC	Tj*; 4xSSC
K	RNA:RNA	> or equal to 50	70° C.; 4xSSC - or- 40° C.; 6xSSC, 50% formamide	67° C.; 1xSSC
L	RNA:RNA	<50	Tl*; 2xSSC	Tl*; 2xSSC
M	DNA:DNA	> or equal to 50	50° C.; 4xSSC - or- 40° C. 6xSSC, 50% formamide	50° C.; 2xSSC
N	DNA:DNA	<50	Tn*; 6xSSC	Tn*; 6xSSC
O	DNA:RNA	> or equal to 50	55° C.; 4xSSC - or- 42° C.; 6xSSC, 50% formamide	55° C.; 2xSSC
P	DNA:RNA	<50	Tp*; 6xSSC	Tp*; 6xSSC
Q	RNA:RNA	> or equal to 50	60° C.; 4xSSC - or- 45° C.; 6xSSC, 50% formamide	60° C.; 2xSSC
R	RNA:RNA	<50	Tr*; 4xSSC	Tr*; 4xSSC

‡The "hybrid length" is the anticipated length for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide of unknown sequence, the hybrid is assumed to be that of the hybridizing polynucleotide of the present invention. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity. Methods of aligning two or more polynucleotide sequences and/or determining the percent identity between two polynucleotide sequences are well known in the art (e.g., MegAlign program of the DNA*Star suite of programs, etc).

†SSPE (1xSSPE is 0.15M NaCl, 10 mM NaH2PO4, and 1.25 mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15 mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete. The hydridizations and washes may additionally include 5X Denhardt's reagent, .5–1.0% SDS, 100 ug/ml denatured, fragmented salmon sperm DNA, 0.5% sodium pyrophosphate, and up to 50% formamide.

*Tb–Tr: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5–10° C. less than the melting temperature Tm of the hybrids there Tm is determined according to the following equations. For hybrids less than 18 base pairs in length, Tm(° C.) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, Tm(° C.) =81.5 + 16.6($\log_{10}$[Na+]) + 0.41(% G + C) – (600/N), where N is the number of bases in the hybrid, and [Na+] is the concentration of sodium ions in the hybridization buffer ([NA+] for 1xSSC = .165 M).

±The present invention encompasses the substitution of any one, or more DNA or RNA hybrid partners with either a PNA, or a modified polynucleotide. Such modified polynucleotides are known in the art and are more particularly described elsewhere herein.

[0278] Additional examples of stringency conditions for polynucleotide hybridization are provided, for example, in Sambrook, J., E. F. Fritsch, and T. Maniatis, 1989, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., chapters 9 and 11, and Current Protocols in Molecular Biology, 1995, F. M., Ausubel et al., eds, John Wiley and Sons, Inc., sections 2.10 and 6.3-6.4, which are hereby incorporated by reference herein.

[0279] Preferably, such hybridizing polynucleotides have at least 70% sequence identity (more preferably, at least 80% identity; and most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which they hybridize, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. The determination of identity is well known in the art, and discussed more specifically elsewhere herein.

[0280] The invention encompasses the application of PCR methodology to the polynucleotide sequences of the present invention, the clone deposited with the ATCC, and/or the cDNA encoding the polypeptides of the present invention. PCR techniques for the amplification of nucleic acids are described in U.S. Pat. No. 4,683,195 and Saiki et al., Science, 239:487-491 (1988). PCR, for example, may include the following steps, of denaturation of template nucleic acid (if double-stranded), annealing of primer to target, and polymerization. The nucleic acid probed or used as a template in the amplification reaction may be genomic DNA, cDNA, RNA, or a PNA. PCR may be used to amplify specific sequences from genomic DNA, specific RNA sequence, and/or cDNA transcribed from mRNA. References for the general use of PCR techniques, including specific method parameters, include Mullis et al., Cold Spring Harbor Symp. Quant. Biol., 51:263, (1987), Ehrlich (ed), PCR Technology, Stockton Press, NY, 1989; Ehrlich et al., Science, 252:1643-1650, (1991); and "PCR Protocols, A Guide to Methods and Applications", Eds., Innis et al., Academic Press, New York, (1990).

Polynucleotide and Polypeptide Variants

[0281] The present invention also encompases variants (e.g., allelic variants, orthologs, etc.) of the polynucleotide sequence disclosed herein in SEQ ID NO:1, the complementary strand thereto, and/or the cDNA sequence contained in the deposited clone.

[0282] The present invention also encompasses variants of the polypeptide sequence, and/or fragments therein, disclosed in SEQ ID NO:2, a polypeptide encoded by the polunucleotide sequence in SEQ ID NO:1, and/or a polypeptide encoded by a cDNA in the deposited clone.

[0283] "Variant" refers to a polynucleotide or polypeptide differing from the polynucleotide or polypeptide of the present invention, but retaining essential properties thereof. Generally, variants are overall closely similar, and, in many regions, identical to the polynucleotide or polypeptide of the present invention.

[0284] Thus, one aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a Protease-42 related polypeptide having an amino acid sequence as shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (b) a nucleotide sequence encoding a mature Protease-42 related polypeptide having the amino acid sequence as shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (c) a nucleotide sequence encoding a biologically active fragment of a Protease-42 related polypeptide having an amino acid sequence shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (d) a nucleotide sequence encoding an antigenic fragment of a Protease-42 related polypeptide having an amino acid sequence shown in the sequence listing and described in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (e) a nucleotide sequence encoding a Protease-42 related polypeptide comprising the complete amino acid sequence encoded by a human cDNA plasmid containined in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (f) a nucleotide sequence encoding a mature Protease-42 realted polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (g) a nucleotide sequence encoding a biologically active fragement of a Protease-42 related polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (h) a nucleotide sequence encoding an antigenic fragment of a Protease-42 related polypeptide having an amino acid sequence encoded by a human cDNA plasmid contained in SEQ ID NO:1 or the cDNA contained in ATCC deposit No:PTA-3745; (I) a nucleotide sequence complimentary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

[0285] The present invention is also directed to polynucleotide sequences which comprise, or alternatively consist of, a polynucleotide sequence which is at least about 80%, 85%, 90%, 91%, 92%, 92.5%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides encoded by these nucleic acid molecules are also encompassed by the invention. In another embodiment, the invention encompasses nucleic acid molecule which comprise, or alternatively, consist of a polynucleotide which hybridizes under stringent conditions, or alternatively, under lower stringency conditions, to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention, as are polypeptides encoded by these polypeptides.

[0286] Another aspect of the invention provides an isolated nucleic acid molecule comprising, or alternatively, consisting of, a polynucleotide having a nucleotide sequence selected from the group consisting of: (a) a nucleotide sequence encoding a Protease-42 related polypeptide having an amino acid sequence as shown in the sequence listing and described in Table 1; (b) a nucleotide sequence encoding a mature Protease-42 related polypeptide having the amino acid sequence as shown in the sequence listing and

described in Table 1; (c) a nucleotide sequence encoding a biologically active fragment of a Protease-42 related polypeptide having an amino acid sequence as shown in the sequence listing and described in Table 1; (d) a nucleotide sequence encoding an antigenic fragment of a Protease-42 related polypeptide having an amino acid sequence as shown in the sequence listing and described in Table 1; (e) a nucleotide sequence encoding a Protease-42 related polypeptide comprising the complete amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in Table 1; (f) a nucleotide sequence encoding a mature Protease-42 related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in Table 1: (g) a nucleotide sequence encoding a biologically active fragment of a Protease-42 related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC Deposit and described in Table 1; (h) a nucleotide sequence encoding an antigenic fragment of a Protease-42 related polypeptide having an amino acid sequence encoded by a human cDNA in a cDNA plasmid contained in the ATCC deposit and described in Table 1; (i) a nucleotide sequence complimentary to any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h) above.

[0287] The present invention is also directed to nucleic acid molecules which comprise, or alternatively, consist of, a nucleotide sequence which is at least about 80%, 85%, 90%, 91%, 92%, 92.5%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, any of the nucleotide sequences in (a), (b), (c), (d), (e), (f), (g), or (h), above.

[0288] The present invention encompasses polypeptide sequences which comprise, or alternatively consist of, an amino acid sequence which is at least about 80%, 85%, 88.4%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, the following non-limited examples, the polypeptide sequence identified as SEQ ID NO:2, the polypeptide sequence encoded by a cDNA provided in the deposited clone, and/or polypeptide fragments of any of the polypeptides provided herein. Polynucleotides encoded by these nucleic acid molecules are also encompassed by the invention. In another embodiment, the invention encompasses nucleic acid molecule which comprise, or alternatively, consist of a polynucleotide which hybridizes under stringent conditions, or alternatively, under lower stringency conditions, to a polynucleotide in (a), (b), (c), (d), (e), (f), (g), or (h), above. Polynucleotides which hybridize to the complement of these nucleic acid molecules under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompassed by the invention, as are polypeptides encoded by these polypeptides.

[0289] The present invention is also directed to polypeptides which comprise, or alternatively consist of, an amino acid sequence which is at least about 80%, 85%, 88.5%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for example, the polypeptide sequence shown in SEQ ID NO:2, a polypeptide sequence encoded by the nucleotide sequence in SEQ ID NO:1, a polypeptide sequence encoded by the cDNA in cDNA plasmid:Z, and/or polypeptide fragments of any of these polypeptides (e.g., those fragments described herein). Polynucleotides which hybridize to the complement of the nucleic acid molecules encoding these polypeptides under stringent hybridization conditions or alternatively, under lower stringency conditions, are also encompasses by the present invention, as are the polypeptides encoded by these polynucleotides.

[0290] By a nucleic acid having a nucleotide sequence at least, for example, 95% "identical" to a reference nucleotide sequence of the present invention, it is intended that the nucleotide sequence of the nucleic acid is identical to the reference sequence except that the nucleotide sequence may include up to five point mutations per each 100 nucleotides of the reference nucleotide sequence encoding the polypeptide. In other words, to obtain a nucleic acid having a nucleotide sequence at least 95% identical to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence. The query sequence may be an entire sequence referenced in Table 1, the ORF (open reading frame), or any fragment specified as described herein.

[0291] As a practical matter, whether any particular nucleic acid molecule or polypeptide is at least about 80%, 85%, 88.5%, 90%, 91%, 92%, 92.5%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to a nucleotide sequence of the present invention can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the CLUSTALW computer program (Thompson, J. D., et al., Nucleic Acids Research, 2(22):4673-4680, (1994)), which is based on the algorithm of Higgins, D. G., et al., Computer Applications in the Biosciences (CABIOS), 8(2):189-191, (1992). In a sequence alignment the query and subject sequences are both DNA sequences. An RNA sequence can be compared by converting U's to T's. However, the CLUSTALW algorithm automatically converts U's to T's when comparing RNA sequences to DNA sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a CLUSTALW alignment of DNA sequences to calculate percent identity via pairwise alignments are: Matrix=IUB, k-tuple=1, Number of Top Diagonals=5, Gap Penalty=3, Gap Open Penalty 10, Gap Extension Penalty=0.1, Scoring Method=Percent, Window Size=5 or the length of the subject nucleotide sequence, whichever is shorter. For multiple alignments, the following CLUSTALW parameters are preferred: Gap Opening Penalty=10; Gap Extension Parameter=0.05; Gap Separation Penalty Range=8; End Gap Separation Penalty=Off; % Identity for Alignment Delay=40%; Residue Specific Gaps:Off; Hydrophilic Residue Gap=Off; and Transition Weighting=0. The pairwise and multple alignment parameters provided for CLUSTALW above represent the default parameters as provided with the AlignX software program (Vector NTI suite of programs, version 6.0).

[0292] The present invention encompasses the application of a manual correction to the percent identity results, in the instance where the subject sequence is shorter than the query sequence because of 5' or 3' deletions, not because of internal deletions. If only the local pairwise percent identity is required, no manual correction is needed. However, a manual correction may be applied to determine the global percent identity from a global polynucleotide alignment. Percent identity calculations based upon global polynucleotide alignments are often preferred since they reflect the percent identity between the polynucleotide molecules as a whole (i.e., including any polynucleotide overhangs, not just overlapping regions), as opposed to, only local matching polynucleotides. Manual corrections for global percent identity determinations are required since the CLUSTALW program does not account for 5' and 3' truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the 5' or 3' ends, relative to the query sequence, the percent identity is corrected by calculating the number of bases of the query sequence that are 5' and 3' of the subject sequence, which are not matched/aligned, as a percent of the total bases of the query sequence. Whether a nucleotide is matched/aligned is determined by results of the CLUSTALW sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above CLUSTALW program using the specified parameters, to arrive at a final percent identity score. This corrected score may be used for the purposes of the present invention. Only bases outside the 5' and 3' bases of the subject sequence, as displayed by the CLUSTALW alignment, which are not matched/aligned with the query sequence, are calculated for the purposes of manually adjusting the percent identity score.

[0293] For example, a 90 base subject sequence is aligned to a 100 base query sequence to determine percent identity. The deletions occur at the 5' end of the subject sequence and therefore, the CLUSTALW alignment does not show a matched/alignment of the first 10 bases at 5' end. The 10 unpaired bases represent 10% of the sequence (number of bases at the 5' and 3' ends not matched/total number of bases in the query sequence) so 10% is subtracted from the percent identity score calculated by the CLUSTALW program. If the remaining 90 bases were perfectly matched the final percent identity would be 90%. In another example, a 90 base subject sequence is compared with a 100 base query sequence. This time the deletions are internal deletions so that there are no bases on the 5' or 3' of the subject sequence which are not matched/aligned with the query. In this case the percent identity calculated by CLUSTALW is not manually corrected. Once again, only bases 5' and 3' of the subject sequence which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are required for the purposes of the present invention.

[0294] By a polypeptide having an amino acid sequence at least, for example, 95% "identical" to a query amino acid sequence of the present invention, it is intended that the amino acid sequence of the subject polypeptide is identical to the query sequence except that the subject polypeptide sequence may include up to five amino acid alterations per each 100 amino acids of the query amino acid sequence. In other words, to obtain a polypeptide having an amino acid sequence at least 95% identical to a query amino acid sequence, up to 5% of the amino acid residues in the subject sequence may be inserted, deleted, or substituted with another amino acid. These alterations of the reference sequence may occur at the amino- or carboxy-terminal positions of the reference amino acid sequence or anywhere between those terminal positions, interspersed either individually among residues in the reference sequence or in one or more contiguous groups within the reference sequence.

[0295] As a practical matter, whether any particular polypeptide is at least about 80%, 85%, 88.5%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, or 99.9% identical to, for instance, an amino acid sequence referenced in Table 1 (SEQ ID NO:2) or to the amino acid sequence encoded by cDNA contained in a deposited clone, can be determined conventionally using known computer programs. A preferred method for determining the best overall match between a query sequence (a sequence of the present invention) and a subject sequence, also referred to as a global sequence alignment, can be determined using the CLUSTALW computer program (Thompson, J. D., et al., Nucleic Acids Research, 2(22):4673-4680, (1994)), which is based on the algorithm of Higgins, D. G., et al., Computer Applications in the Biosciences (CABIOS), 8(2):189-191, (1992). In a sequence alignment the query and subject sequences are both amino acid sequences. The result of said global sequence alignment is in percent identity. Preferred parameters used in a CLUSTALW alignment of DNA sequences to calculate percent identity via pairwise alignments are: Matrix=BLOSUM, k-tuple=1, Number of Top Diagonals=5, Gap Penalty=3, Gap Open Penalty 10, Gap Extension Penalty=0.1, Scoring Method=Percent, Window Size=5 or the length of the subject nucleotide sequence, whichever is shorter. For multiple alignments, the following CLUSTALW parameters are preferred: Gap Opening Penalty=10; Gap Extension Parameter=0.05; Gap Separation Penalty Range=8; End Gap Separation Penalty=Off; % Identity for Alignment Delay=40%; Residue Specific Gaps:Off; Hydrophilic Residue Gap=Off; and Transition Weighting=0. The pairwise and multple alignment parameters provided for CLUSTALW above represent the default parameters as provided with the AlignX software program (Vector NTI suite of programs, version 6.0).

[0296] The present invention encompasses the application of a manual correction to the percent identity results, in the instance where the subject sequence is shorter than the query sequence because of N- or C-terminal deletions, not because of internal deletions. If only the local pairwise percent identity is required, no manual correction is needed. However, a manual correction may be applied to determine the global percent identity from a global polypeptide alignment. Percent identity calculations based upon global polypeptide alignments are often preferred since they reflect the percent identity between the polypeptide molecules as a whole (i.e., including any polypeptide overhangs, not just overlapping regions), as opposed to, only local matching polypeptides. Manual corrections for global percent identity determinations are required since the CLUSTALW program does not account for N- and C-terminal truncations of the subject sequence when calculating percent identity. For subject sequences truncated at the N- and C-termini, relative to the query sequence, the percent identity is corrected by calculating the number of residues of the query sequence that are N- and C-terminal of the subject sequence, which are not matched/aligned with a corresponding subject residue, as a

percent of the total bases of the query sequence. Whether a residue is matched/aligned is determined by results of the CLUSTALW sequence alignment. This percentage is then subtracted from the percent identity, calculated by the above CLUSTALW program using the specified parameters, to arrive at a final percent identity score. This final percent identity score is what may be used for the purposes of the present invention. Only residues to the N- and C-termini of the subject sequence, which are not matched/aligned with the query sequence, are considered for the purposes of manually adjusting the percent identity score. That is, only query residue positions outside the farthest N- and C-terminal residues of the subject sequence.

[0297] For example, a 90 amino acid residue subject sequence is aligned with a 100 residue query sequence to determine percent identity. The deletion occurs at the N-terminus of the subject sequence and therefore, the CLUSTALW alignment does not show a matching/alignment of the first 10 residues at the N-terminus. The 10 unpaired residues represent 10% of the sequence (number of residues at the N- and C-termini not matched/total number of residues in the query sequence) so 10% is subtracted from the percent identity score calculated by the CLUSTALW program. If the remaining 90 residues were perfectly matched the final percent identity would be 90%. In another example, a 90 residue subject sequence is compared with a 100 residue query sequence. This time the deletions are internal deletions so there are no residues at the N- or C-termini of the subject sequence, which are not matched/aligned with the query. In this case the percent identity calculated by CLUSTALW is not manually corrected. Once again, only residue positions outside the N- and C-terminal ends of the subject sequence, as displayed in the CLUSTALW alignment, which are not matched/aligned with the query sequence are manually corrected for. No other manual corrections are required for the purposes of the present invention.

[0298] In addition to the above method of aligning two or more polynucleotide or polypeptide sequences to arrive at a percent identity value for the aligned sequences, it may be desirable in some circumstances to use a modified version of the CLUSTALW algorithm which takes into account known structural features of the sequences to be aligned, such as for example, the SWISS-PROT designations for each sequence. The result of such a modifed CLUSTALW algorithm may provide a more accurate value of the percent identity for two polynucleotide or polypeptide sequences. Support for such a modified version of CLUSTALW is provided within the CLUSTALW algorithm and would be readily appreciated to one of skill in the art of bioinformatics.

[0299] The variants may contain alterations in the coding regions, non-coding regions, or both. Especially preferred are polynucleotide variants containing alterations which produce silent substitutions, additions, or deletions, but do not alter the properties or activities of the encoded polypeptide. Nucleotide variants produced by silent substitutions due to the degeneracy of the genetic code are preferred. Moreover, variants in which 5-10, 1-5, or 1-2 amino acids are substituted, deleted, or added in any combination are also preferred. Polynucleotide variants can be produced for a variety of reasons, e.g., to optimize codon expression for a particular host (change codons in the mRNA to those preferred by a bacterial host such as *E. coli*).

[0300] Naturally occurring variants are called "allelic variants," and refer to one of several alternate forms of a gene occupying a given locus on a chromosome of an organism. (Genes II, Lewin, B., ed., John Wiley & Sons, New York (1985).) These allelic variants can vary at either the polynucleotide and/or polypeptide level and are included in the present invention. Alternatively, non-naturally occurring variants may be produced by mutagenesis techniques or by direct synthesis.

[0301] Using known methods of protein engineering and recombinant DNA technology, variants may be generated to improve or alter the characteristics of the polypeptides of the present invention. For instance, one or more amino acids can be deleted from the N-terminus or C-terminus of the protein without substantial loss of biological function. The authors of Ron et al., J. Biol. Chem. 268: 2984-2988 (1993), reported variant KGF proteins having heparin binding activity even after deleting 3, 8, or 27 amino-terminal amino acid residues. Similarly, Interferon gamma exhibited up to ten times higher activity after deleting 8-10 amino acid residues from the carboxy terminus of this protein (Dobeli et al., J. Biotechnology 7:199-216 (1988)).

[0302] Moreover, ample evidence demonstrates that variants often retain a biological activity similar to that of the naturally occurring protein. For example, Gayle and coworkers (J. Biol. Chem. 268:22105-22111 (1993)) conducted extensive mutational analysis of human cytokine IL-1a. They used random mutagenesis to generate over 3,500 individual IL-1a mutants that averaged 2.5 amino acid changes per variant over the entire length of the molecule. Multiple mutations were examined at every possible amino acid position. The investigators found that "[m]ost of the molecule could be altered with little effect on either [binding or biological activity]." In fact, only 23 unique amino acid sequences, out of more than 3,500 nucleotide sequences examined, produced a protein that significantly differed in activity from wild-type.

[0303] Furthermore, even if deleting one or more amino acids from the N-terminus or C-terminus of a polypeptide results in modification or loss of one or more biological functions, other biological activities may still be retained. For example, the ability of a deletion variant to induce and/or to bind antibodies which recognize the protein will likely be retained when less than the majority of the residues of the protein are removed from the N-terminus or C-terminus. Whether a particular polypeptide lacking N- or C-terminal residues of a protein retains such immunogenic activities can readily be determined by routine methods described herein and otherwise known in the art.

[0304] Alternatively, such N-terminus or C-terminus deletions of a polypeptide of the present invention may, in fact, result in a significant increase in one or more of the biological activities of the polypeptide(s). For example, biological activity of many polypeptides are governed by the presence of regulatory domains at either one or both termini. Such regulatory domains effectively inhibit the biological activity of such polypeptides in lieu of an activation event (e.g., binding to a cognate ligand or receptor, phosphorylation, proteolytic processing, etc.). Thus, by eliminating the regulatory domain of a polypeptide, the polypeptide may effectively be rendered biologically active in the absence of an activation event.

[0305] Thus, the invention further includes polypeptide variants that show substantial biological activity. Such variants include deletions, insertions, inversions, repeats, and substitutions selected according to general rules known in the art so as have little effect on activity. For example, guidance concerning how to make phenotypically silent amino acid substitutions is provided in Bowie et al., Science 247:1306-1310 (1990), wherein the authors indicate that there are two main strategies for studying the tolerance of an amino acid sequence to change.

[0306] The first strategy exploits the tolerance of amino acid substitutions by natural selection during the process of evolution. By comparing amino acid sequences in different species, conserved amino acids can be identified. These conserved amino acids are likely important for protein function. In contrast, the amino acid positions where substitutions have been tolerated by natural selection indicates that these positions are not critical for protein function. Thus, positions tolerating amino acid substitution could be modified while still maintaining biological activity of the protein.

[0307] The second strategy uses genetic engineering to introduce amino acid changes at specific positions of a cloned gene to identify regions critical for protein function. For example, site directed mutagenesis or alanine-scanning mutagenesis (introduction of single alanine mutations at every residue in the molecule) can be used. (Cunningham and Wells, Science 244:1081-1085 (1989).) The resulting mutant molecules can then be tested for biological activity.

[0308] As the authors state, these two strategies have revealed that proteins are surprisingly tolerant of amino acid substitutions. The authors further indicate which amino acid changes are likely to be permissive at certain amino acid positions in the protein. For example, most buried (within the tertiary structure of the protein) amino acid residues require nonpolar side chains, whereas few features of surface side chains are generally conserved.

[0309] The invention encompasses polypeptides having a lower degree of identity but having sufficient similarity so as to perform one or more of the same functions performed by the polypeptide of the present invention. Similarity is determined by conserved amino acid substitution. Such substitutions are those that substitute a given amino acid in a polypeptide by another amino acid of like characteristics (e.g., chemical properties). According to Cunningham et al above, such conservative substitutions are likely to be phenotypically silent. Additional guidance concerning which amino acid changes are likely to be phenotypically silent are found in Bowie et al., Science 247:1306-1310 (1990).

[0310] Tolerated conservative amino acid substitutions of the present invention involve replacement of the aliphatic or hydrophobic amino acids Ala, Val, Leu and Ile; replacement of the hydroxyl residues Ser and Thr; replacement of the acidic residues Asp and Glu; replacement of the amide residues Asn and Gin, replacement of the basic residues Lys, Arg, and His; replacement of the aromatic residues Phe, Tyr, and Trp, and replacement of the small-sized amino acids Ala, Ser, Thr, Met, and Gly.

[0311] In addition, the present invention also encompasses the conservative substitutions provided in Table III below.

TABLE III

For Amino Acid	Code	Replace with any of:
Alanine	A	D-Ala, Gly, beta-Ala, L-Cys, D-Cys
Arginine	R	D-Arg, Lys, D-Lys, homo-Arg, D-homo-Arg, Met, Ile, D-Met, D-Ile, Orn, D-Orn
Asparagine	N	D-Asn, Asp, D-Asp, Glu, D-Glu, Gln, D-Gln
Aspartic Acid	D	D-Asp, D-Asn, Asn, Glu, D-Glu, Gln, D-Gln
Cysteine	C	D-Cys, S-Me-Cys, Met, D-Met, Thr, D-Thr
Glutamine	Q	D-Gln, Asn, D-Asn, Glu, D-Glu, Asp, D-Asp
Glutamic Acid	E	D-Glu, D-Asp, Asp, Asn, D-Asn, Gln, D-Gln
Glycine	G	Ala, D-Ala, Pro, D-Pro, β-Ala, Acp
Isoleucine	I	D-Ile, Val, D-Val, Leu, D-Leu, Met, D-Met
Leucine	L	D-Leu, Val, D-Val, Met, D-Met
Lysine	K	D-Lys, Arg, D-Arg, homo-Arg, D-homo-Arg, Met, D-Met, Ile, D-Ile, Orn, D-Orn
Methionine	M	D-Met, S-Me-Cys, Ile, D-Ile, Leu, D-Leu, Val, D-Val
Phenylalanine	F	D-Phe, Tyr, D-Thr, L-Dopa, His, D-His, Trp, D-Trp, Trans-3,4, or 5-phenylproline, cis-3,4, or 5-phenylproline
Proline	P	D-Pro, L-1-thioazolidine-4-carboxylic acid, D- or L-1-oxazolidine-4-carboxylic acid
Serine	S	D-Ser, Thr, D-Thr, allo-Thr, Met, D-Met, Met(O), D-Met(O), L-Cys, D-Cys
Threonine	T	D-Thr, Ser, D-Ser, allo-Thr, Met, D-Met, Met(O), D-Met(O), Val, D-Val
Tyrosine	Y	D-Tyr, Phe, D-Phe, L-Dopa, His, D-His
Valine	V	D-Val, Leu, D-Leu, Ile, D-Ile, Met, D-Met

[0312] Aside from the uses described above, such amino acid substitutions may also increase protein or peptide stability. The invention encompasses amino acid substitutions that contain, for example, one or more non-peptide bonds (which replace the peptide bonds) in the protein or peptide sequence. Also included are substitutions that include amino acid residues other than naturally occurring L-amino acids, e.g., D-amino acids or non-naturally occurring or synthetic amino acids, e.g., β or γ amino acids.

[0313] Both identity and similarity can be readily calculated by reference to the following publications: Computational Molecular Biology, Lesk, A.M., ed., Oxford University Press, New York, 1988; Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, New York, 1993; Informatics Computer Analysis of Sequence Data, Part 1, Griffin, A. M., and Griffin, H. G., eds., Humana Press, New Jersey, 1994; Sequence Analysis in Molecular Biology, von Heinje, G., Academic Press, 1987; and Sequence Analysis Primer, Gribskov, M. and Devereux, J., eds., M Stockton Press, New York, 1991.

[0314] In addition, the present invention also encompasses substitution of amino acids based upon the probability of an amino acid substitution resulting in conservation of function. Such probabilities are determined by aligning multiple genes with related function and assessing the relative penalty of each substitution to proper gene function. Such probabilities are often described in a matrix and are used by some algorithms (e.g., BLAST, CLUSTALW, GAP, etc.) in calculating percent similarity wherein similarity refers to the degree by which one amino acid may substitute for another amino acid without lose of function. An example of such a matrix is the PAM250 or BLOSUM62 matrix.

[0315] Aside from the canonical chemically conservative substitutions referenced above, the invention also encom-

passes substitutions which are typically not classified as conservative, but that may be chemically conservative under certain circumstances. Analysis of enzymatic catalysis for proteases, for example, has shown that certain amino acids within the active site of some enzymes may have highly perturbed pKa's due to the unique microenvironment of the active site. Such perturbed pKa's could enable some amino acids to substitute for other amino acids while conserving enzymatic structure and function. Examples of amino acids that are known to have amino acids with perturbed pKa's are the Glu-35 residue of Lysozyme, the Ile-16 residue of Chymotrypsin, the His-159 residue of Papain, etc. The conservation of function relates to either anomalous protonation or anomalous deprotonation of such amino acids, relative to their canonical, non-perturbed pKa. The pKa perturbation may enable these amino acids to actively participate in general acid-base catalysis due to the unique ionization environment within the enzyme active site. Thus, substituting an amino acid capable of serving as either a general acid or general base within the microenvironment of an enzyme active site or cavity, as may be the case, in the same or similar capacity as the wild-type amino acid, would effectively serve as a conservative amino substitution.

[0316] Besides conservative amino acid substitution, variants of the present invention include, but are not limited to, the following: (i) substitutions with one or more of the non-conserved amino acid residues, where the substituted amino acid residues may or may not be one encoded by the genetic code, or (ii) substitution with one or more of amino acid residues having a substituent group, or (iii) fusion of the mature polypeptide with another compound, such as a compound to increase the stability and/or solubility of the polypeptide (for example, polyethylene glycol), or (iv) fusion of the polypeptide with additional amino acids, such as, for example, an IgG Fc fusion region peptide, or leader or secretory sequence, or a sequence facilitating purification. Such variant polypeptides are deemed to be within the scope of those skilled in the art from the teachings herein.

[0317] For example, polypeptide variants containing amino acid substitutions of charged amino acids with other charged or neutral amino acids may produce proteins with improved characteristics, such as less aggregation. Aggregation of pharmaceutical formulations both reduces activity and increases clearance due to the aggregate's immunogenic activity. (Pinckard et al., Clin. Exp. Immunol. 2:331-340 (1967); Robbins et al., Diabetes 36: 838-845 (1987); Cleland et al., Crit. Rev. Therapeutic Drug Carrier Systems 10:307-377 (1993).)

[0318] Moreover, the invention further includes polypeptide variants created through the application of molecular evolution ("DNA Shuffling") methodology to the polynucleotide disclosed as SEQ ID NO:1, the sequence of the clone submitted in a deposit, and/or the cDNA encoding the polypeptide disclosed as SEQ ID NO:2. Such DNA Shuffling technology is known in the art and more particularly described elsewhere herein (e.g., WPC, Stemmer, PNAS, 91:10747, (1994)), and in the Examples provided herein).

[0319] A further embodiment of the invention relates to a polypeptide which comprises the amino acid sequence of the present invention having an amino acid sequence which contains at least one amino acid substitution, but not more than 50 amino acid substitutions, even more preferably, not more than 40 amino acid substitutions, still more preferably, not more than 30 amino acid substitutions, and still even more preferably, not more than 20 amino acid substitutions. Of course, in order of ever-increasing preference, it is highly preferable for a peptide or polypeptide to have an amino acid sequence which comprises the amino acid sequence of the present invention, which contains at least one, but not more than 10, 9, 8, 7, 6, 5, 4, 3, 2 or 1 amino acid substitutions. In specific embodiments, the number of additions, substitutions, and/or deletions in the amino acid sequence of the present invention or fragments thereof (e.g., the mature form and/or other fragments described herein), is 1-5,5-10, 5-25, 5-50, 10-50 or 50-150, conservative amino acid substitutions are preferable.

Polynucleotide and Polypeptide Fragments

[0320] The present invention is directed to polynucleotide fragments of the polynucleotides of the invention, in addition to polypeptides encoded therein by said polynucleotides and/or fragments.

[0321] In the present invention, a "polynucleotide fragment" refers to a short polynucleotide having a nucleic acid sequence which: is a portion of that contained in a deposited clone, or encoding the polypeptide encoded by the cDNA in a deposited clone; is a portion of that shown in SEQ ID NO:1 or the complementary strand thereto, or is a portion of a polynucleotide sequence encoding the polypeptide of SEQ ID NO:2. The nucleotide fragments of the invention are preferably at least about 15 nt, and more preferably at least about 20 nt, still more preferably at least about 30 nt, and even more preferably, at least about 40 nt, at least about 50 nt, at least about 75 nt, or at least about 150 nt in length. A fragment "at least 20 nt in length," for example, is intended to include 20 or more contiguous bases from the cDNA sequence contained in a deposited clone or the nucleotide sequence shown in SEQ ID NO:1. In this context "about" includes the particularly recited value, a value larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus, or at both termini. These nucleotide fragments have uses that include, but are not limited to, as diagnostic probes and primers as discussed herein. Of course, larger fragments (e.g., 50, 150, 500, 600, 2000 nucleotides) are preferred.

[0322] Moreover, representative examples of polynucleotide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, a sequence from about nucleotide number 1-50, 51-100, 101-150, 151-200, 201-250, 251-300, 301-350, 351-400, 401-450, 451-500, 501-550, 551-600, 651-700, 701-750, 751-800, 800-850, 851-900, 901-950, 951-1000, 1001-1050, 1051-1100, 1101-1150, 1151-1200, 1201-1250, 1251-1300, 1301-1350, 1351-1400, 1401-1450, 1451-1500, 1501-1550, 1551-1600, 1601-1650, 1651-1700, 1701-1750, 1751-1800, 1801-1850, 1851-1900, 1901-1950, 1951-2000, or 2001 to the end of SEQ ID NO:1, or the complementary strand thereto, or the cDNA contained in a deposited clone. In this context "about" includes the particularly recited ranges, and ranges larger or smaller by several (5, 4, 3, 2, or 1) nucleotides, at either terminus or at both termini. Preferably, these fragments encode a polypeptide which has biological activity. More preferably, these polynucleotides can be used as probes or primers as discussed herein. Also encompassed by the present invention are polynucleotides which hybrid-

ize to these nucleic acid molecules under stringent hybridization conditions or lower stringency conditions, as are the polypeptides encoded by these polynucleotides.

[0323] In the present invention, a "polypeptide fragment" refers to an amino acid sequence which is a portion of that contained in SEQ ID NO:2 or encoded by the cDNA contained in a deposited clone. Protein (polypeptide) fragments may be "free-standing," or comprised within a larger polypeptide of which the fragment forms a part or region, most preferably as a single continuous region. Representative examples of polypeptide fragments of the invention, include, for example, fragments comprising, or alternatively consisting of, from about amino acid number 1-20, 21-40, 41-60, 61-80, 81-100, 102-120, 121-140, 141-160, or 161 to the end of the coding region. Moreover, polypeptide fragments can be about 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, or 150 amino acids in length. In this context "about" includes the particularly recited ranges or values, and ranges or values larger or smaller by several (5, 4, 3, 2, or 1) amino acids, at either extreme or at both extremes. Polynucleotides encoding these polypeptides are also encompassed by the invention.

[0324] Preferred polypeptide fragments include the full-length protein. Further preferred polypeptide fragments include the full-length protein having a continuous series of deleted residues from the amino or the carboxy terminus, or both. For example, any number of amino acids, ranging from 1-60, can be deleted from the amino terminus of the full-length polypeptide. Similarly, any number of amino acids, ranging from 1-30, can be deleted from the carboxy terminus of the full-length protein. Furthermore, any combination of the above amino and carboxy terminus deletions are preferred. Similarly, polynucleotides encoding these polypeptide fragments are also preferred.

[0325] Also preferred are polypeptide and polynucleotide fragments characterized by structural or functional domains, such as fragments that comprise alpha-helix and alpha-helix forming regions, beta-sheet and beta-sheet-forming regions, turn and turn-forming regions, coil and coil-forming regions, hydrophilic regions, hydrophobic regions, alpha amphipathic regions, beta amphipathic regions, flexible regions, surface-forming regions, substrate binding region, and high antigenic index regions. Polypeptide fragments of SEQ ID NO:2 falling within conserved domains are specifically contemplated by the present invention. Moreover, polynucleotides encoding these domains are also contemplated.

[0326] Other preferred polypeptide fragments are biologically active fragments. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide of the present invention. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity. Polynucleotides encoding these polypeptide fragments are also encompassed by the invention.

[0327] In a preferred embodiment, the functional activity displayed by a polypeptide encoded by a polynucleotide fragment of the invention may be one or more biological activities typically associated with the full-length polypeptide of the invention. Illustrative of these biological activities includes the fragments ability to bind to at least one of the same antibodies which bind to the full-length protein, the fragments ability to interact with at lease one of the same proteins which bind to the full-length, the fragments ability to elicit at least one of the same immune responses as the full-length protein (i.e., to cause the immune system to create antibodies specific to the same epitope, etc.), the fragments ability to bind to at least one of the same polynucleotides as the full-length protein, the fragments ability to bind to a receptor of the full-length protein, the fragments ability to bind to a ligand of the full-length protein, and the fragments ability to multimerize with the full-length protein. However, the skilled artisan would appreciate that some fragments may have biological activities which are desirable and directly inapposite to the biological activity of the full-length protein. The functional activity of polypeptides of the invention, including fragments, variants, derivatives, and analogs thereof can be determined by numerous methods available to the skilled artisan, some of which are described elsewhere herein.

[0328] The present invention encompasses polypeptides comprising, or alternatively consisting of, an epitope of the polypeptide having an amino acid sequence of SEQ ID NO:2, or an epitope of the polypeptide sequence encoded by a polynucleotide sequence contained in ATCC Deposit No:Z or encoded by a polynucleotide that hybridizes to the complement of the sequence of SEQ ID NO:1 or contained in ATCC Deposit No:Z under stringent hybridization conditions or lower stringency hybridization conditions as defined supra. The present invention further encompasses polynucleotide sequences encoding an epitope of a polypeptide sequence of the invention (such as, for example, the sequence disclosed in SEQ ID NO:1), polynucleotide sequences of the complementary strand of a polynucleotide sequence encoding an epitope of the invention, and polynucleotide sequences which hybridize to the complementary strand under stringent hybridization conditions or lower stringency hybridization conditions defined supra.

[0329] The term "epitopes," as used herein, refers to portions of a polypeptide having antigenic or immunogenic activity in an animal, preferably a mammal, and most preferably in a human. In a preferred embodiment, the present invention encompasses a polypeptide comprising an epitope, as well as the polynucleotide encoding this polypeptide. An "immunogenic epitope," as used herein, is defined as a portion of a protein that elicits an antibody response in an animal, as determined by any method known in the art, for example, by the methods for generating antibodies described infra. (See, for example, Geysen et al., Proc. Natl. Acad. Sci. USA 81:3998-4002 (1983)). The term "antigenic epitope," as used herein, is defined as a portion of a protein to which an antibody can immunospecifically bind its antigen as determined by any method well known in the art, for example, by the immunoassays described herein. Immunospecific binding excludes non-specific binding but does not necessarily exclude cross-reactivity with other antigens. Antigenic epitopes need not necessarily be immunogenic.

[0330] Fragments which function as epitopes may be produced by any conventional means. (See, e.g., Houghten, Proc. Natl. Acad. Sci. USA 82:5131-5135 (1985), further described in U.S. Pat. No. 4,631,211).

[0331] In the present invention, antigenic epitopes preferably contain a sequence of at least 4, at least 5, at least 6, at least 7, more preferably at least 8, at least 9, at least 10, at

least 11, at least 12, at least 13, at least 14, at least 15, at least 20, at least 25, at least 30, at least 40, at least 50, and, most preferably, between about 15 to about 30 amino acids. Preferred polypeptides comprising immunogenic or antigenic epitopes are at least 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, or 100 amino acid residues in length, or longer. Additional non-exclusive preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as portions thereof. Antigenic epitopes are useful, for example, to raise antibodies, including monoclonal antibodies, that specifically bind the epitope. Preferred antigenic epitopes include the antigenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these antigenic epitopes. Antigenic epitopes can be used as the target molecules in immunoassays. (See, for instance, Wilson et al., Cell 37:767-778 (1984); Sutcliffe et al., Science 219:660-666 (1983)).

[0332] Similarly, immunogenic epitopes can be used, for example, to induce antibodies according to methods well known in the art. (See, for instance, Sutcliffe et al., supra; Wilson et al., supra; Chow et al., Proc. Natl. Acad. Sci. USA 82:910-914; and Bittle et al., J. Gen. Virol. 66:2347-2354 (1985). Preferred immunogenic epitopes include the immunogenic epitopes disclosed herein, as well as any combination of two, three, four, five or more of these immunogenic epitopes. The polypeptides comprising one or more immunogenic epitopes may be presented for eliciting an antibody response together with a carrier protein, such as an albumin, to an animal system (such as rabbit or mouse), or, if the polypeptide is of sufficient length (at least about 25 amino acids), the polypeptide may be presented without a carrier. However, immunogenic epitopes comprising as few as 8 to 10 amino acids have been shown to be sufficient to raise antibodies capable of binding to, at the very least, linear epitopes in a denatured polypeptide (e.g., in Western blotting).

[0333] Epitope-bearing polypeptides of the present invention may be used to induce antibodies according to methods well known in the art including, but not limited to, in vivo immunization, in vitro immunization, and phage display methods. See, e.g., Sutcliffe et al., supra; Wilson et al., supra, and Bittle et al., J. Gen. Virol., 66:2347-2354 (1985). If in vivo immunization is used, animals may be immunized with free peptide; however, anti-peptide antibody titer may be boosted by coupling the peptide to a macromolecular carrier, such as keyhole limpet hemacyanin (KLH) or tetanus toxoid. For instance, peptides containing cysteine residues may be coupled to a carrier using a linker such as maleimidobenzoyl-N-hydroxysuccinimide ester (MBS), while other peptides may be coupled to carriers using a more general linking agent such as glutaraldehyde. Animals such as rabbits, rats and mice are immunized with either free or carrier-coupled peptides, for instance, by intraperitoneal and/or intradermal injection of emulsions containing about 100 μg of peptide or carrier protein and Freund's adjuvant or any other adjuvant known for stimulating an immune response. Several booster injections may be needed, for instance, at intervals of about two weeks, to provide a useful titer of anti-peptide antibody which can be detected, for example, by ELISA assay using free peptide adsorbed to a solid surface. The titer of anti-peptide antibodies in serum from an immunized animal may be increased by selection of anti-peptide antibodies, for instance, by adsorption to the peptide on a solid support and elution of the selected antibodies according to methods well known in the art.

[0334] As one of skill in the art will appreciate, and as discussed above, the polypeptides of the present invention comprising an immunogenic or antigenic epitope can be fused to other polypeptide sequences. For example, the polypeptides of the present invention may be fused with the constant domain of immunoglobulins (IgA, IgE, IgG, IgM), or portions thereof (CH1, CH2, CH3, or any combination thereof and portions thereof) resulting in chimeric polypeptides. Such fusion proteins may facilitate purification and may increase half-life in vivo. This has been shown for chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. See, e.g., EP 394,827; Traunecker et al., Nature, 331:84-86 (1988). Enhanced delivery of an antigen across the epithelial barrier to the immune system has been demonstrated for antigens (e.g., insulin) conjugated to an FcRn binding partner such as IgG or Fe fragments (see, e.g., PCT Publications WO 96/22024 and WO 99/04813). IgG Fusion proteins that have a disulfide-linked dimeric structure due to the IgG portion disulfide bonds have also been found to be more efficient in binding and neutralizing other molecules than monomeric polypeptides or fragments thereof alone. See, e.g., Fountoulakis et al., J. Biochem., 270:3958-3964 (1995). Nucleic acids encoding the above epitopes can also be recombined with a gene of interest as an epitope tag (e.g., the hemagglutinin ("HA") tag or flag tag) to aid in detection and purification of the expressed polypeptide. For example, a system described by Janknecht et al. allows for the ready purification of non-denatured fusion proteins expressed in human cell lines (Janknecht et al., 1991, Proc. Natl. Acad. Sci. USA 88:8972-897). In this system, the gene of interest is subcloned into a vaccinia recombination plasmid such that the open reading frame of the gene is translationally fused to an amino-terminal tag consisting of six histidine residues. The tag serves as a matrix binding domain for the fusion protein. Extracts from cells infected with the recombinant vaccinia virus are loaded onto Ni2+ nitriloacetic acid-agarose column and histidine-tagged proteins can be selectively eluted with imidazole-containing buffers.

[0335] Additional fusion proteins of the invention may be generated through the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling"). DNA shuffling may be employed to modulate the activities of polypeptides of the invention, such methods can be used to generate polypeptides with altered activity, as well as agonists and antagonists of the polypeptides. See, generally, U.S. Pat. Nos. 5,605,793; 5,811,238; 5,830,721; 5,834,252; and 5,837,458, and Patten et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, Trends Biotechnol. 16(2):76-82 (1998); Hansson, et al., J. Mol. Biol. 287:265-76 (1999); and Lorenzo and Balls, Biotechniques 24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference in its entirety). In one embodiment, alteration of polynucleotides corresponding to SEQ ID NO:1 and the polypeptides encoded by these polynucleotides may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments by homologous or site-specific recombination to generate variation in the polynucleotide sequence. In another embodiment, polynucleotides of the invention, or the encoded polypeptides, may be

altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of a polynucleotide encoding a polypeptide of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules.

Antibodies

[0336] Further polypeptides of the invention relate to antibodies and T-cell antigen receptors (TCR) which immunospecifically bind a polypeptide, polypeptide fragment, or variant of SEQ ID NO:2, and/or an epitope, of the present invention (as determined by immunoassays well known in the art for assaying specific antibody-antigen binding). Antibodies of the invention include, but are not limited to, polyclonal, monoclonal, monovalent, bispecific, heteroconjugate, multispecific, human, humanized or chimeric antibodies, single chain antibodies, Fab fragments, F(ab') fragments, fragments produced by a Fab expression library, anti-idiotypic (anti-Id) antibodies (including, e.g., anti-Id antibodies to antibodies of the invention), and epitope-binding fragments of any of the above. The term "antibody," as used herein, refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site that immunospecifically binds an antigen. The immunoglobulin molecules of the invention can be of any type (e.g., IgG, IgE, IgM, IgD, IgA and IgY), class (e.g., IgG1, IgG2, IgG3, IgG4, IgA1 and IgA2) or subclass of immunoglobulin molecule. Moreover, the term "antibody" (Ab) or "monoclonal antibody" (Mab) is meant to include intact molecules, as well as, antibody fragments (such as, for example, Fab and F(ab')2 fragments) which are capable of specifically binding to protein. Fab and F(ab')2 fragments lack the Fc fragment of intact antibody, clear more rapidly from the circulation of the animal or plant, and may have less non-specific tissue binding than an intact antibody (Wahl et al., J. Nucl. Med. 24:316-325 (1983)). Thus, these fragments are preferred, as well as the products of a FAB or other immunoglobulin expression library. Moreover, antibodies of the present invention include chimeric, single chain, and humanized antibodies.

[0337] Most preferably the antibodies are human antigen-binding antibody fragments of the present invention and include, but are not limited to, Fab, Fab' and F(ab')2, Fd, single-chain Fvs (scFv), single-chain antibodies, disulfide-linked Fvs (sdFv) and fragments comprising either a VL or VH domain. Antigen-binding antibody fragments, including single-chain antibodies, may comprise the variable region(s) alone or in combination with the entirety or a portion of the following: hinge region, CH1, CH2, and CH3 domains. Also included in the invention are antigen-binding fragments also comprising any combination of variable region(s) with a hinge region, CH1, CH2, and CH3 domains. The antibodies of the invention may be from any animal origin including birds and mammals. Preferably, the antibodies are human, murine (e.g., mouse and rat), donkey, ship rabbit, goat, guinea pig, camel, horse, or chicken. As used herein, "human" antibodies include antibodies having the amino acid sequence of a human immunoglobulin and include antibodies isolated from human immunoglobulin libraries or from animals transgenic for one or more human immunoglobulin and that do not express endogenous immunoglobulins, as described infra and, for example in, U.S. Pat. No. 5,939,598 by Kucherlapati et al.

[0338] The antibodies of the present invention may be monospecific, bispecific, trispecific or of greater multispecificity. Multispecific antibodies may be specific for different epitopes of a polypeptide of the present invention or may be specific for both a polypeptide of the present invention as well as for a heterologous epitope, such as a heterologous polypeptide or solid support material. See, e.g., PCT publications WO 93/17715; WO 92/08802; WO 91/00360; WO 92/05793; Tutt, et al., J. Immunol. 147:60-69 (1991); U.S. Pat. Nos. 4,474,893; 4,714,681; 4,925,648; 5,573,920; 5,601,819; Kostelny et al., J. Immunol. 148:1547-1553 (1992).

[0339] Antibodies of the present invention may be described or specified in terms of the epitope(s) or portion(s) of a polypeptide of the present invention which they recognize or specifically bind. The epitope(s) or polypeptide portion(s) may be specified as described herein, e.g., by N-terminal and C-terminal positions, by size in contiguous amino acid residues, or listed in the Tables and Figures. Antibodies which specifically bind any epitope or polypeptide of the present invention may also be excluded. Therefore, the present invention includes antibodies that specifically bind polypeptides of the present invention, and allows for the exclusion of the same.

[0340] Antibodies of the present invention may also be described or specified in terms of their cross-reactivity. Antibodies that do not bind any other analog, ortholog, or homologue of a polypeptide of the present invention are included. Antibodies that bind polypeptides with at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 65%, at least 60%, at least 55%, and at least 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In specific embodiments, antibodies of the present invention cross-react with murine, rat and/or rabbit homologues of human proteins and the corresponding epitopes thereof. Antibodies that do not bind polypeptides with less than 95%, less than 90%, less than 85%, less than 80%, less than 75%, less than 70%, less than 65%, less than 60%, less than 55%, and less than 50% identity (as calculated using methods known in the art and described herein) to a polypeptide of the present invention are also included in the present invention. In a specific embodiment, the above-described cross-reactivity is with respect to any single specific antigenic or immunogenic polypeptide, or combination(s) of 2, 3, 4, 5, or more of the specific antigenic and/or immunogenic polypeptides disclosed herein. Further included in the present invention are antibodies which bind polypeptides encoded by polynucleotides which hybridize to a polynucleotide of the present invention under stringent hybridization conditions (as described herein). Antibodies of the present invention may also be described or specified in terms of their binding affinity to a polypeptide of the invention. Preferred binding affinities include those with a dissociation constant or Kd less than 5×10–2 M, 10–2 M, 5×10–3 M, 10–3 M, 5×10–4 M, 10–4 M, **5×10^{-5}** M, 10–5 M, 5×10–6 M, 10–6M, 5×10–7 M, 107 M, 5×10–8 M, 10–8 M, 5×10–9 M, 10–9 M,

5×10−10 M, 10−10 M, 5×10−11 M, 10−11 M, 5×10−12 M, 10−12 M, 5×10−13 M, 10−13 M, 5×10−14 M, 10−14 M, 5×10−15 M, or 10−15 M.

[0341] The invention also provides antibodies that competitively inhibit binding of an antibody to an epitope of the invention as determined by any method known in the art for determining competitive binding, for example, the immunoassays described herein. In preferred embodiments, the antibody competitively inhibits binding to the epitope by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50%.

[0342] Antibodies of the present invention may act as agonists or antagonists of the polypeptides of the present invention. For example, the present invention includes antibodies which disrupt the receptor/ligand interactions with the polypeptides of the invention either partially or fully. Preferably, antibodies of the present invention bind an antigenic epitope disclosed herein, or a portion thereof. The invention features both receptor-specific antibodies and ligand-specific antibodies. The invention also features receptor-specific antibodies which do not prevent ligand binding but prevent receptor activation. Receptor activation (i.e., signaling) may be determined by techniques described herein or otherwise known in the art. For example, receptor activation can be determined by detecting the phosphorylation (e.g., tyrosine or serine/threonine) of the receptor or its substrate by immunoprecipitation followed by western blot analysis (for example, as described supra). In specific embodiments, antibodies are provided that inhibit ligand activity or receptor activity by at least 95%, at least 90%, at least 85%, at least 80%, at least 75%, at least 70%, at least 60%, or at least 50% of the activity in absence of the antibody.

[0343] The invention also features receptor-specific antibodies which both prevent ligand binding and receptor activation as well as antibodies that recognize the receptor-ligand complex, and, preferably, do not specifically recognize the unbound receptor or the unbound ligand. Likewise, included in the invention are neutralizing antibodies which bind the ligand and prevent binding of the ligand to the receptor, as well as antibodies which bind the ligand, thereby preventing receptor activation, but do not prevent the ligand from binding the receptor. Further included in the invention are antibodies which activate the receptor. These antibodies may act as receptor agonists, i.e., potentiate or activate either all or a subset of the biological activities of the ligand-mediated receptor activation, for example, by inducing dimerization of the receptor. The antibodies may be specified as agonists, antagonists or inverse agonists for biological activities comprising the specific biological activities of the peptides of the invention disclosed herein. The above antibody agonists can be made using methods known in the art. See, e.g., PCT publication WO 96/40281; U.S. Pat. No. 5,811,097; Deng et al., Blood 92(6):1981-1988 (1998); Chen et al., Cancer Res. 58(16):3668-3678 (1998); Harrop et al., J. Immunol. 161(4):1786-1794 (1998); Zhu et al., Cancer Res. 58(15):3209-3214 (1998); Yoon et al., J. Immunol. 160(7):3170-3179 (1998); Prat et al., J. Cell. Sci. 111 (Pt2):237-247 (1998); Pitard et al., J. Immunol. Methods 205(2):177-190 (1997); Liautard et al., Cytokine 9(4):233-241 (1997); Carlson et al., J. Biol. Chem. 272(17):11295-11301 (1997); Taryman et al., Neuron 14(4):755-762 (1995); Muller et al., Structure 6(9):1153-1167 (1998); Bartunek et al., Cytokine 8(1):14-20 (1996) (which are all incorporated by reference herein in their entireties).

[0344] Antibodies of the present invention may be used, for example, but not limited to, to purify, detect, and target the polypeptides of the present invention, including both in vitro and in vivo diagnostic and therapeutic methods. For example, the antibodies have use in immunoassays for qualitatively and quantitatively measuring levels of the polypeptides of the present invention in biological samples. See, e.g., Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988) (incorporated by reference herein in its entirety).

[0345] As discussed in more detail below, the antibodies of the present invention may be used either alone or in combination with other compositions. The antibodies may further be recombinantly fused to a heterologous polypeptide at the N- or C-terminus or chemically conjugated (including covalently and non-covalently conjugations) to polypeptides or other compositions. For example, antibodies of the present invention may be recombinantly fused or conjugated to molecules useful as labels in detection assays and effector molecules such as heterologous polypeptides, drugs, radionucleotides, or toxins. See, e.g., PCT publications WO 92/08495; WO 91/14438; WO 89/12624; U.S. Pat. No. 5,314,995; and EP 396,387.

[0346] The antibodies of the invention include derivatives that are modified, i.e., by the covalent attachment of any type of molecule to the antibody such that covalent attachment does not prevent the antibody from generating an anti-idiotypic response. For example, but not by way of limitation, the antibody derivatives include antibodies that have been modified, e.g., by glycosylation, acetylation, pegylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to a cellular ligand or other protein, etc. Any of numerous chemical modifications may be carried out by known techniques, including, but not limited to specific chemical cleavage, acetylation, formylation, metabolic synthesis of tunicamycin, etc. Additionally, the derivative may contain one or more non-classical amino acids.

[0347] The antibodies of the present invention may be generated by any suitable method known in the art.

[0348] The antibodies of the present invention may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan (Harlow, et al., Antibodies: A Laboratory Manual, (Cold spring Harbor Laboratory Press, 2nd ed. (1988); and Current Protocols, Chapter 2; which are hereby incorporated herein by reference in its entirety). In a preferred method, a preparation of the Protease-19 protein is prepared and purified to render it substantially free of natural contaminants. Such a preparation is then introduced into an animal in order to produce polyclonal antisera of greater specific activity. For example, a polypeptide of the invention can be administered to various host animals including, but not limited to, rabbits, mice, rats, etc. to induce the production of sera containing polyclonal antibodies specific for the antigen. The administration of the polypeptides of the present invention may entail one or more injections of an immunizing agent and, if desired, an adjuvant. Various adjuvants may be used to increase the immunological response, depending on the host species, and include but are not limited to, Freund's (com-

plete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanins, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and corynebacterium parvum. Such adjuvants are also well known in the art. For the purposes of the invention, "immunizing agent" may be defined as a polypeptide of the invention, including fragments, variants, and/or derivatives thereof, in addition to fusions with heterologous polypeptides and other forms of the polypeptides described herein.

[0349] Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections, though they may also be given intramuscularly, and/or through IV). The immunizing agent may include polypeptides of the present invention or a fusion protein or variants thereof. Depending upon the nature of the polypeptides (i.e., percent hydrophobicity, percent hydrophilicity, stability, net charge, isoelectric point etc.), it may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Such conjugation includes either chemical conjugation by derivitizing active chemical functional groups to both the polypeptide of the present invention and the immunogenic protein such that a covalent bond is formed, or through fusion-protein based methodology, or other methods known to the skilled artisan. Examples of such immunogenic proteins include, but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacille Calmette-Guerin) and *Corynebacterium parvum*. Additional examples of adjuvants which may be employed includes the MPL-TDM adjuvant (monophosphoryl lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

[0350] The antibodies of the present invention may comprise monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975) and U.S. Pat. No. 4,376,110, by Harlow, et al., Antibodies: A Laboratory Manual, (Cold spring Harbor Laboratory Press, 2nded. (1988), by Hammerling, et al., Monoclonal Antibodies and T-Cell Hybridomas (Elsevier, N.Y., pp. 563-681 (1981); Köhler et al., Eur. J. Immunol. 6:511 (1976); Köhler et al., Eur. J. Immunol. 6:292 (1976), or other methods known to the artisan. Other examples of methods which may be employed for producing monoclonal antibodies includes, but are not limited to, the human B-cell hybridoma technique (Kosbor et al., 1983, Immunology Today 4:72; Cole et al., 1983, Proc. Natl. Acad. Sci. USA 80:2026-2030), and the EBV-hybridoma technique (Cole et al., 1985, Monoclonal Antibodies And Cancer Therapy, Alan R. Liss, Inc., pp. 77-96). Such antibodies may be of any immunoglobulin class including IgG, IgM, IgE, IgA, IgD and any subclass thereof. The hybridoma producing the mAb of this invention may be cultivated in vitro or in vivo. Production of high titers of mAbs in vivo makes this the presently preferred method of production.

[0351] In a hybridoma method, a mouse, a humanized mouse, a mouse with a human immune system, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized in vitro.

[0352] The immunizing agent will typically include polypeptides of the present invention or a fusion protein thereof. Preferably, the immunizing agent consists of an Protease-19 polypeptide or, more preferably, with a Protease-19 polypeptide-expressing cell. Such cells may be cultured in any suitable tissue culture medium; however, it is preferable to culture cells in Earle's modified Eagle's medium supplemented with 10% fetal bovine serum (inactivated at about 56 degrees C.), and supplemented with about 10 g/l of nonessential amino acids, about 1,000 U/ml of penicillin, and about 100 ug/ml of streptomycin. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell (Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986), pp. 59-103). Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

[0353] Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, Calif. and the American Type Culture Collection, Manassas, Va. More preferred are the parent myeloma cell line (SP20) as provided by the ATCC. As inferred throughout the specification, human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies (Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63).

[0354] The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against the polypeptides of the present invention. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an in vitro binding

assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbant assay (ELISA). Such techniques are known in the art and within the skill of the artisan. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollart, Anal. Biochem., 107:220 (1980).

[0355] After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods (Goding, supra, and/or according to Wands et al. (Gastroenterology 80:225-232 (1981)). Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640. Alternatively, the hybridoma cells may be grown in vivo as ascites in a mammal.

[0356] The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-sepharose, hydroxyapatite chromatography, gel exclusion chromatography, gel electrophoresis, dialysis, or affinity chromatography.

[0357] The skilled artisan would acknowledge that a variety of methods exist in the art for the production of monoclonal antibodies and thus, the invention is not limited to their sole production in hydridomas. For example, the monoclonal antibodies may be made by recombinant DNA methods, such as those described in U.S. Pat. No. 4,816,567. In this context, the term "monoclonal antibody" refers to an antibody derived from a single eukaryotic, phage, or prokaryotic clone. The DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies, or such chains from human, humanized, or other sources). The hydridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transformed into host cells such as Simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding sequence for human heavy and light chain constant domains in place of the homologous murine sequences (U.S. Pat. No. 4,816,567; Morrison et al, supra) or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

[0358] The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

[0359] In vitro methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art. Monoclonal antibodies can be prepared using a wide variety of techniques known in the art including the use of hybridoma, recombinant, and phage display technologies, or a combination thereof. For example, monoclonal antibodies can be produced using hybridoma techniques including those known in the art and taught, for example, in Harlow et al., Antibodies: A Laboratory Manual, (Cold Spring Harbor Laboratory Press, 2nd ed. 1988); Hammerling, et al., in: Monoclonal Antibodies and T-Cell Hybridomas 563-681 (Elsevier, N.Y., 1981) (said references incorporated by reference in their entireties). The term "monoclonal antibody" as used herein is not limited to antibodies produced through hybridoma technology. The term "monoclonal antibody" refers to an antibody that is derived from a single clone, including any eukaryotic, prokaryotic, or phage clone, and not the method by which it is produced.

[0360] Methods for producing and screening for specific antibodies using hybridoma technology are routine and well known in the art and are discussed in detail in the Examples described herein. In a non-limiting example, mice can be immunized with a polypeptide of the invention or a cell expressing such peptide. Once an immune response is detected, e.g., antibodies specific for the antigen are detected in the mouse serum, the mouse spleen is harvested and splenocytes isolated. The splenocytes are then fused by well known techniques to any suitable myeloma cells, for example cells from cell line SP20 available from the ATCC. Hybridomas are selected and cloned by limited dilution. The hybridoma clones are then assayed by methods known in the art for cells that secrete antibodies capable of binding a polypeptide of the invention. Ascites fluid, which generally contains high levels of antibodies, can be generated by immunizing mice with positive hybridoma clones.

[0361] Accordingly, the present invention provides methods of generating monoclonal antibodies as well as antibodies produced by the method comprising culturing a hybridoma cell secreting an antibody of the invention wherein, preferably, the hybridoma is generated by fusing splenocytes isolated from a mouse immunized with an antigen of the invention with myeloma cells and then screening the hybridomas resulting from the fusion for hybridoma clones that secrete an antibody able to bind a polypeptide of the invention.

[0362] Antibody fragments which recognize specific epitopes may be generated by known techniques. For example, Fab and F(ab')2 fragments of the invention may be produced by proteolytic cleavage of immunoglobulin molecules, using enzymes such as papain (to produce Fab fragments) or pepsin (to produce F(ab')2 fragments). F(ab')2 fragments contain the variable region, the light chain constant region and the CH1 domain of the heavy chain.

[0363] For example, the antibodies of the present invention can also be generated using various phage display methods known in the art. In phage display methods, functional antibody domains are displayed on the surface of phage particles which carry the polynucleotide sequences encoding them. In a particular embodiment, such phage can be utilized to display antigen binding domains expressed

from a repertoire or combinatorial antibody library (e.g., human or murine). Phage expressing an antigen binding domain that binds the antigen of interest can be selected or identified with antigen, e.g., using labeled antigen or antigen bound or captured to a solid surface or bead. Phage used in these methods are typically filamentous phage including fd and M13 binding domains expressed from phage with Fab, Fv or disulfide stabilized Fv antibody domains recombinantly fused to either the phage gene III or gene VIII protein. Examples of phage display methods that can be used to make the antibodies of the present invention include those disclosed in Brinkman et al., J. Immunol. Methods 182:41-50 (1995); Ames et al., J. Immunol. Methods 184:177-186 (1995); Kettleborough et al., Eur. J. Immunol. 24:952-958 (1994); Persic et al., Gene 187 9-18 (1997); Burton et al., Advances in Immunology 57:191-280 (1994); PCT application No. PCT/GB91/01134; PCT publications WO 90/02809; WO 91/10737; WO 92/01047; WO 92/18619; WO 93/11236; WO 95/15982; WO 95/20401; and U.S. Pat. Nos. 5,698,426; 5,223,409; 5,403,484; 5,580,717; 5,427,908; 5,750,753; 5,821,047; 5,571,698; 5,427,908; 5,516,637; 5,780,225; 5,658,727; 5,733,743 and 5,969,108; each of which is incorporated herein by reference in its entirety.

[0364] As described in the above references, after phage selection, the antibody coding regions from the phage can be isolated and used to generate whole antibodies, including human antibodies, or any other desired antigen binding fragment, and expressed in any desired host, including mammalian cells, insect cells, plant cells, yeast, and bacteria, e.g., as described in detail below. For example, techniques to recombinantly produce Fab, Fab' and F(ab')2 fragments can also be employed using methods known in the art such as those disclosed in PCT publication WO 92/22324; Mullinax et al., BioTechniques 12(6):864-869 (1992); and Sawai et al., AJR134:26-34 (1995); and Better et al., Science 240:1041-1043 (1988) (said references incorporated by reference in their entireties). Examples of techniques which can be used to produce single-chain Fvs and antibodies include those described in U.S. Pat. Nos. 4,946,778 and 5,258,498; Huston et al., Methods in Enzymology 203:46-88 (1991); Shu et al., PNAS 90:7995-7999 (1993); and Skerra et al., Science 240:1038-1040 (1988).

[0365] For some uses, including in vivo use of antibodies in humans and in vitro detection assays, it may be preferable to use chimeric, humanized, or human antibodies. A chimeric antibody is a molecule in which different portions of the antibody are derived from different animal species, such as antibodies having a variable region derived from a murine monoclonal antibody and a human immunoglobulin constant region. Methods for producing chimeric antibodies are known in the art. See e.g., Morrison, Science 229:1202 (1985); Oi et al., BioTechniques 4:214 (1986); Gillies et al., (1989) J. Immunol. Methods 125:191-202; Cabilly et al., Taniguchi et al., EP 171496; Morrison et al., EP 173494; Neuberger et al., WO 8601533; Robinson et al., WO 8702671; Boulianne et al., Nature 312:643 (1984); Neuberger et al., Nature 314:268 (1985); U.S. Pat. Nos. 5,807,715; 4,816,567; and 4,816,397, which are incorporated herein by reference in their entirety. Humanized antibodies are antibody molecules from non-human species antibody that binds the desired antigen having one or more complementarity determining regions (CDRs) from the non-human species and a framework regions from a human immunoglobulin molecule. Often, framework residues in the human framework regions will be substituted with the corresponding residue from the CDR donor antibody to alter, preferably improve, antigen binding. These framework substitutions are identified by methods well known in the art, e.g., by modeling of the interactions of the CDR and framework residues to identify framework residues important for antigen binding and sequence comparison to identify unusual framework residues at particular positions. (See, e.g., Queen et al., U.S. Pat. No. 5,585,089; Riechmann et al., Nature 332:323 (1988), which are incorporated herein by reference in their entireties.) Antibodies can be humanized using a variety of techniques known in the art including, for example, CDR-grafting (EP 239,400; PCT publication WO 91/09967; U.S. Pat. Nos. 5,225,539; 5,530,101; and 5,585,089), veneering or resurfacing (EP 592,106; EP 519,596; Padlan, Molecular Immunology 28(4/5):489-498 (1991); Studnicka et al., Protein Engineering 7(6):805-814 (1994); Roguska. et al., PNAS 91:969-973 (1994)), and chain shuffling (U.S. Pat. No. 5,565,332). Generally, a humanized antibody has one or more amino acid residues introduced into it from a source that is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the methods of Winter and co-workers (Jones et al., Nature, 321:522-525 (1986); Reichmann et al., Nature, 332:323-327 (1988); Verhoeyen et al., Science, 239:1534-1536 (1988), by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Pat. No. 4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possible some FR residues are substituted from analogous sites in rodent antibodies.

[0366] In general, the humanized antibody will comprise substantially all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin (Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature 332:323-329 (1988)1 and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992).

[0367] Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Human antibodies can be made by a variety of methods known in the art including phage display methods described above using antibody libraries derived from human immunoglobulin sequences. See also, U.S. Pat. Nos. 4,444,887 and 4,716,111; and PCT publications WO 98/46645, WO 98/50433, WO 98/24893, WO 98/16654, WO 96/34096, WO 96/33735, and WO 91/10741; each of which is incorporated herein by reference in its entirety. The techniques of cole et al., and Boerder et al., are also available for the preparation of human monoclonal antibodies (cole et al., Monoclonal Antibodies and Cancer Therapy, Alan R. Riss, (1985); and Boerner et al., J. Immunol., 147(1):86-95, (1991)).

[0368] Human antibodies can also be produced using transgenic mice which are incapable of expressing functional endogenous immunoglobulins, but which can express human immunoglobulin genes. For example, the human heavy and light chain immunoglobulin gene complexes may be introduced randomly or by homologous recombination into mouse embryonic stem cells. Alternatively, the human variable region, constant region, and diversity region may be introduced into mouse embryonic stem cells in addition to the human heavy and light chain genes. The mouse heavy and light chain immunoglobulin genes may be rendered non-functional separately or simultaneously with the introduction of human immunoglobulin loci by homologous recombination. In particular, homozygous deletion of the JH region prevents endogenous antibody production. The modified embryonic stem cells are expanded and microinjected into blastocysts to produce chimeric mice. The chimeric mice are then bred to produce homozygous offspring which express human antibodies. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained from the immunized, transgenic mice using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA, IgM and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar, Int. Rev. Immunol. 13:65-93 (1995). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, see, e.g., PCT publications WO 98/24893; WO 92/01047; WO 96/34096; WO 96/33735; European Patent No. 0 598 877; U.S. Pat. Nos. 5,413,923; 5,625,126; 5,633,425; 5,569,825; 5,661,016; 5,545,806; 5,814,318; 5,885,793; 5,916,771; and 5,939,598, which are incorporated by reference herein in their entirety. In addition, companies such as Abgenix, Inc. (Freemont, Calif.), Genpharm (San Jose, Calif.), and Medarex, Inc. (Princeton, N.J.) can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

[0369] Similarly, human antibodies can be made by introducing human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and creation of an antibody repertoire. This approach is described, for example, in U.S. Pat. Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,106, and in the following scientific publications: Marks et al., Biotechnol., 10:779-783 (1992); Lonberg et al., Nature 368:856-859 (1994); Fishwild et al., Nature Biotechnol., 14:845-51 (1996); Neuberger, Nature Biotechnol., 14:826 (1996); Lonberg and Huszer, Intern. Rev. Immunol., 13:65-93 (1995).

[0370] Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a mouse antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al., Bio/technology 12:899-903 (1988)).

[0371] Further, antibodies to the polypeptides of the invention can, in turn, be utilized to generate anti-idiotype antibodies that "mimic" polypeptides of the invention using techniques well known to those skilled in the art. (See, e.g., Greenspan & Bona, FASEB J. 7(5):437-444; (1989) and Nissinoff, J. Immunol. 147(8):2429-2438 (1991)). For example, antibodies which bind to and competitively inhibit polypeptide multimerization and/or binding of a polypeptide of the invention to a ligand can be used to generate anti-idiotypes that "mimic" the polypeptide multimerization and/or binding domain and, as a consequence, bind to and neutralize polypeptide and/or its ligand. Such neutralizing anti-idiotypes or Fab fragments of such anti-idiotypes can be used in therapeutic regimens to neutralize polypeptide ligand. For example, such anti-idiotypic antibodies can be used to bind a polypeptide of the invention and/or to bind its ligands/receptors, and thereby block its biological activity.

[0372] Such anti-idiotypic antibodies capable of binding to the Protease-19 polypeptide can be produced in a two-step procedure. Such a method makes use of the fact that antibodies are themselves antigens, and therefore, it is possible to obtain an antibody that binds to a second antibody. In accordance with this method, protein specific antibodies are used to immunize an animal, preferably a mouse. The splenocytes of such an animal are then used to produce hybridoma cells, and the hybridoma cells are screened to identify clones that produce an antibody whose ability to bind to the protein-specific antibody can be blocked by the polypeptide. Such antibodies comprise anti-idiotypic antibodies to the protein-specific antibody and can be used to immunize an animal to induce formation of further protein-specific antibodies.

[0373] The antibodies of the present invention may be bispecific antibodies. Bispecific antibodies are monoclonal, Preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present invention, one of the binding specificities may be directed towards a polypeptide of the present invention, the other may be for any other antigen, and preferably for a cell-surface protein, receptor, receptor subunit, tissue-specific antigen, virally derived protein, virally encoded envelope protein, bacterially derived protein, or bacterial surface protein, etc.

[0374] Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities (Milstein and Cuello, Nature, 305:537-539 (1983). Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published May 13, 1993, and in Traunecker et al., EMBO J., 10:3655-3659 (1991).

[0375] Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be

fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transformed into a suitable host organism. For further details of generating bispecific antibodies see, for example Suresh et al., Meth. In Enzym., 121:210 (1986).

[0376] Heteroconjugate antibodies are also contemplated by the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells (U.S. Pat. No. 4,676,980), and for the treatment of HIV infection (WO 91/00360; WO 92/20373; and EP03089). It is contemplated that the antibodies may be prepared in vitro using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example, immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioester bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Pat. No. 4,676,980.

Polynucleotides Encoding Antibodies

[0377] The invention further provides polynucleotides comprising a nucleotide sequence encoding an antibody of the invention and fragments thereof. The invention also encompasses polynucleotides that hybridize under stringent or lower stringency hybridization conditions, e.g., as defined supra, to polynucleotides that encode an antibody, preferably, that specifically binds to a polypeptide of the invention, preferably, an antibody that binds to a polypeptide having the amino acid sequence of SEQ ID NO:2.

[0378] The polynucleotides may be obtained, and the nucleotide sequence of the polynucleotides determined, by any method known in the art. For example, if the nucleotide sequence of the antibody is known, a polynucleotide encoding the antibody may be assembled from chemically synthesized oligonucleotides (e.g., as described in Kutmeier et al., BioTechniques 17:242 (1994)), which, briefly, involves the synthesis of overlapping oligonucleotides containing portions of the sequence encoding the antibody, annealing and ligating of those oligonucleotides, and then amplification of the ligated oligonucleotides by PCR.

[0379] Alternatively, a polynucleotide encoding an antibody may be generated from nucleic acid from a suitable source. If a clone containing a nucleic acid encoding a particular antibody is not available, but the sequence of the antibody molecule is known, a nucleic acid encoding the immunoglobulin may be chemically synthesized or obtained from a suitable source (e.g., an antibody cDNA library, or a cDNA library generated from, or nucleic acid, preferably poly A+ RNA, isolated from, any tissue or cells expressing the antibody, such as hybridoma cells selected to express an antibody of the invention) by PCR amplification using synthetic primers hybridizable to the 3' and 5' ends of the sequence or by cloning using an oligonucleotide probe specific for the particular gene sequence to identify, e.g., a cDNA clone from a cDNA library that encodes the antibody. Amplified nucleic acids generated by PCR may then be cloned into replicable cloning vectors using any method well known in the art.

[0380] Once the nucleotide sequence and corresponding amino acid sequence of the antibody is determined, the nucleotide sequence of the antibody may be manipulated using methods well known in the art for the manipulation of nucleotide sequences, e.g., recombinant DNA techniques, site directed mutagenesis, PCR, etc. (see, for example, the techniques described in Sambrook et al., 1990, Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. and Ausubel et al., eds., 1998, Current Protocols in Molecular Biology, John Wiley & Sons, NY, which are both incorporated by reference herein in their entireties), to generate antibodies having a different amino acid sequence, for example to create amino acid substitutions, deletions, and/or insertions.

[0381] In a specific embodiment, the amino acid sequence of the heavy and/or light chain variable domains may be inspected to identify the sequences of the complementarity determining regions (CDRs) by methods that are well know in the art, e.g., by comparison to known amino acid sequences of other heavy and light chain variable regions to determine the regions of sequence hypervariability. Using routine recombinant DNA techniques, one or more of the CDRs may be inserted within framework regions, e.g., into human framework regions to humanize a non-human antibody, as described supra. The framework regions may be naturally occurring or consensus framework regions, and preferably human framework regions (see, e.g., Chothia et al., J. Mol. Biol. 278: 457-479 (1998) for a listing of human framework regions). Preferably, the polynucleotide generated by the combination of the framework regions and CDRs encodes an antibody that specifically binds a polypeptide of the invention. Preferably, as discussed supra, one or more amino acid substitutions may be made within the framework regions, and, preferably, the amino acid substitutions improve binding of the antibody to its antigen. Additionally, such methods may be used to make amino acid substitutions or deletions of one or more variable region cysteine residues participating in an intrachain disulfide bond to generate antibody molecules lacking one or more intrachain disulfide bonds. Other alterations to the polynucleotide are encompassed by the present invention and within the skill of the art.

[0382] In addition, techniques developed for the production of "chimeric antibodies" (Morrison et al., Proc. Natl. Acad. Sci. 81:851-855 (1984); Neuberger et al., Nature 312:604-608 (1984); Takeda et al., Nature 314:452-454 (1985)) by splicing genes from a mouse antibody molecule of appropriate antigen specificity together with genes from a human antibody molecule of appropriate biological activity can be used. As described supra, a chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region, e.g., humanized antibodies.

[0383] Alternatively, techniques described for the production of single chain antibodies (U.S. Pat. No. 4,946,778; Bird, Science 242:423-42 (1988); Huston et al., Proc. Natl.

Acad. Sci. USA 85:5879-5883 (1988); and Ward et al., Nature 334:544-54 (1989)) can be adapted to produce single chain antibodies. Single chain antibodies are formed by linking the heavy and light chain fragments of the Fv region via an amino acid bridge, resulting in a single chain polypeptide. Techniques for the assembly of functional Fv fragments in *E. coli* may also be used (Skerra et al., Science 242:1038-1041 (1988)).

[0384] More preferably, a clone encoding an antibody of the present invention may be obtained according to the method described in the Example section herein.

Methods of Producing Antibodies

[0385] The antibodies of the invention can be produced by any method known in the art for the synthesis of antibodies, in particular, by chemical synthesis or preferably, by recombinant expression techniques.

[0386] Recombinant expression of an antibody of the invention, or fragment, derivative or analog thereof, (e.g., a heavy or light chain of an antibody of the invention or a single chain antibody of the invention), requires construction of an expression vector containing a polynucleotide that encodes the antibody. Once a polynucleotide encoding an antibody molecule or a heavy or light chain of an antibody, or portion thereof (preferably containing the heavy or light chain variable domain), of the invention has been obtained, the vector for the production of the antibody molecule may be produced by recombinant DNA technology using techniques well known in the art. Thus, methods for preparing a protein by expressing a polynucleotide containing an antibody encoding nucleotide sequence are described herein. Methods which are well known to those skilled in the art can be used to construct expression vectors containing antibody coding sequences and appropriate transcriptional and translational control signals. These methods include, for example, in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. The invention, thus, provides replicable vectors comprising a nucleotide sequence encoding an antibody molecule of the invention, or a heavy or light chain thereof, or a heavy or light chain variable domain, operably linked to a promoter. Such vectors may include the nucleotide sequence encoding the constant region of the antibody molecule (see, e.g., PCT Publication WO 86/05807; PCT Publication WO 89/01036; and U.S. Pat. No. 5,122,464) and the variable domain of the antibody may be cloned into such a vector for expression of the entire heavy or light chain.

[0387] The expression vector is transferred to a host cell by conventional techniques and the transfected cells are then cultured by conventional techniques to produce an antibody of the invention. Thus, the invention includes host cells containing a polynucleotide encoding an antibody of the invention, or a heavy or light chain thereof, or a single chain antibody of the invention, operably linked to a heterologous promoter. In preferred embodiments for the expression of double-chained antibodies, vectors encoding both the heavy and light chains may be co-expressed in the host cell for expression of the entire immunoglobulin molecule, as detailed below.

[0388] A variety of host-expression vector systems may be utilized to express the antibody molecules of the invention. Such host-expression systems represent vehicles by which the coding sequences of interest may be produced and subsequently purified, but also represent cells which may, when transformed or transfected with the appropriate nucleotide coding sequences, express an antibody molecule of the invention in situ. These include but are not limited to microorganisms such as bacteria (e.g., *E. coli, B. subtilis*) transformed with recombinant bacteriophage DNA, plasmid DNA or cosmid DNA expression vectors containing antibody coding sequences; yeast (e.g., Saccharomyces, Pichia) transformed with recombinant yeast expression vectors containing antibody coding sequences; insect cell systems infected with recombinant virus expression vectors (e.g., baculovirus) containing antibody coding sequences; plant cell systems infected with recombinant virus expression vectors (e.g., cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or transformed with recombinant plasmid expression vectors (e.g., Ti plasmid) containing antibody coding sequences; or mammalian cell systems (e.g., COS, CHO, BHK, 293, 3T3 cells) harboring recombinant expression constructs containing promoters derived from the genome of mammalian cells (e.g., metallothionein promoter) or from mammalian viruses (e.g., the adenovirus late promoter; the vaccinia virus 7.5K promoter). Preferably, bacterial cells such as *Escherichia coli*, and more preferably, eukaryotic cells, especially for the expression of whole recombinant antibody molecule, are used for the expression of a recombinant antibody molecule. For example, mammalian cells such as Chinese hamster ovary cells (CHO), in conjunction with a vector such as the major intermediate early gene promoter element from human cytomegalovirus is an effective expression system for antibodies (Foecking et al., Gene 45:101 (1986); Cockett et al., Bio/Technology 8:2 (1990)).

[0389] In bacterial systems, a number of expression vectors may be advantageously selected depending upon the use intended for the antibody molecule being expressed. For example, when a large quantity of such a protein is to be produced, for the generation of pharmaceutical compositions of an antibody molecule, vectors which direct the expression of high levels of fusion protein products that are readily purified may be desirable. Such vectors include, but are not limited, to the *E. coli* expression vector pUR278 (Ruther et al., EMBO J. 2:1791 (1983)), in which the antibody coding sequence may be ligated individually into the vector in frame with the lac Z coding region so that a fusion protein is produced; pIN vectors (Inouye & Inouye, Nucleic Acids Res. 13:3101-3109 (1985); Van Heeke & Schuster, J. Biol. Chem. 24:5503-5509 (1989)); and the like. pGEX vectors may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption and binding to matrix glutathione-agarose beads followed by elution in the presence of free glutathione. The pGEX vectors are designed to include thrombin or factor Xa protease cleavage sites so that the cloned target gene product can be released from the GST moiety.

[0390] In an insect system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes. The virus grows in *Spodoptera frugiperda* cells. The antibody coding sequence may be cloned individually into non-essential regions (for example the polyhedrin gene) of the virus and placed under control of an AcNPV promoter (for example the polyhedrin promoter).

[0391] In mammalian host cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, the antibody coding sequence of interest may be ligated to an adenovirus transcription/translation control complex, e.g., the late promoter and tripartite leader sequence. This chimeric gene may then be inserted in the adenovirus genome by in vitro or in vivo recombination. Insertion in a non-essential region of the viral genome (e.g., region E1 or E3) will result in a recombinant virus that is viable and capable of expressing the antibody molecule in infected hosts. (e.g., see Logan & Shenk, Proc. Natl. Acad. Sci. USA 81:355-359 (1984)). Specific initiation signals may also be required for efficient translation of inserted antibody coding sequences. These signals include the ATG initiation codon and adjacent sequences. Furthermore, the initiation codon must be in phase with the reading frame of the desired coding sequence to ensure translation of the entire insert. These exogenous translational control signals and initiation codons can be of a variety of origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of appropriate transcription enhancer elements, transcription terminators, etc. (see Bittner et al., Methods in Enzymol. 153:51-544 (1987)).

[0392] In addition, a host cell strain may be chosen which modulates the expression of the inserted sequences, or modifies and processes the gene product in the specific fashion desired. Such modifications (e.g., glycosylation) and processing (e.g., cleavage) of protein products may be important for the function of the protein. Different host cells have characteristic and specific mechanisms for the post-translational processing and modification of proteins and gene products. Appropriate cell lines or host systems can be chosen to ensure the correct modification and processing of the foreign protein expressed. To this end, eukaryotic host cells which possess the cellular machinery for proper processing of the primary transcript, glycosylation, and phosphorylation of the gene product may be used. Such mammalian host cells include but are not limited to CHO, VERY, BHK, Hela, COS, MDCK, 293, 3T3, WI38, and in particular, breast cancer cell lines such as, for example, BT483, Hs578T, HTB2, BT20 and T47D, and normal mammary gland cell line such as, for example, CRL7030 and Hs578Bst.

[0393] For long-term, high-yield production of recombinant proteins, stable expression is preferred. For example, cell lines which stably express the antibody molecule may be engineered. Rather than using expression vectors which contain viral origins of replication, host cells can be transformed with DNA controlled by appropriate expression control elements (e.g., promoter, enhancer, sequences, transcription terminators, polyadenylation sites, etc.), and a selectable marker. Following the introduction of the foreign DNA, engineered cells may be allowed to grow for 1-2 days in an enriched media, and then are switched to a selective media. The selectable marker in the recombinant plasmid confers resistance to the selection and allows cells to stably integrate the plasmid into their chromosomes and grow to form foci which in turn can be cloned and expanded into cell lines. This method may advantageously be used to engineer cell lines which express the antibody molecule. Such engineered cell lines may be particularly useful in screening and evaluation of compounds that interact directly or indirectly with the antibody molecule.

[0394] A number of selection systems may be used, including but not limited to the herpes simplex virus thymidine kinase (Wigler et al., Cell 11:223 (1977)), hypoxanthine-guanine phosphoribosyltransferase (Szybalska & Szybalski, Proc. Natl. Acad. Sci. USA 48:202 (1992)), and adenine phosphoribosyltransferase (Lowy et al., Cell 22:817 (1980)) genes can be employed in tk-, hgprt- or aprt-cells, respectively. Also, antimetabolite resistance can be used as the basis of selection for the following genes: dhfr, which confers resistance to methotrexate (Wigler et al., Natl. Acad. Sci. USA 77:357 (1980); O'Hare et al., Proc. Natl. Acad. Sci. USA 78:1527 (1981)); gpt, which confers resistance to mycophenolic acid (Mulligan & Berg, Proc. Natl. Acad. Sci. USA 78:2072 (1981)); neo, which confers resistance to the aminoglycoside G-418 Clinical Pharmacy 12:488-505; Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science 260:926-932 (1993); and Morgan and Anderson, Ann. Rev. Biochem. 62:191-217 (1993); May, 1993, TIB TECH 11(5):155-215); and hygro, which confers resistance to hygromycin (Santerre et al., Gene 30:147 (1984)). Methods commonly known in the art of recombinant DNA technology may be routinely applied to select the desired recombinant clone, and such methods are described, for example, in Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, NY (1993); Kriegler, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY (1990); and in Chapters 12 and 13, Dracopoli et al. (eds), Current Protocols in Human Genetics, John Wiley & Sons, NY (1994); Colberre-Garapin et al., J. Mol. Biol. 150:1 (1981), which are incorporated by reference herein in their entireties.

[0395] The expression levels of an antibody molecule can be increased by vector amplification (for a review, see Bebbington and Hentschel, The use of vectors based on gene amplification for the expression of cloned genes in mammalian cells in DNA cloning, Vol.3. (Academic Press, New York, 1987)). When a marker in the vector system expressing antibody is amplifiable, increase in the level of inhibitor present in culture of host cell will increase the number of copies of the marker gene. Since the amplified region is associated with the antibody gene, production of the antibody will also increase (Crouse et al., Mol. Cell. Biol. 3:257 (1983)).

[0396] The host cell may be co-transfected with two expression vectors of the invention, the first vector encoding a heavy chain derived polypeptide and the second vector encoding a light chain derived polypeptide. The two vectors may contain identical selectable markers which enable equal expression of heavy and light chain polypeptides. Alternatively, a single vector may be used which encodes, and is capable of expressing, both heavy and light chain polypeptides. In such situations, the light chain should be placed before the heavy chain to avoid an excess of toxic free heavy chain (Proudfoot, Nature 322:52 (1986); Kohler, Proc. Natl. Acad. Sci. USA 77:2197 (1980)). The coding sequences for the heavy and light chains may comprise cDNA or genomic DNA.

[0397] Once an antibody molecule of the invention has been produced by an animal, chemically synthesized, or recombinantly expressed, it may be purified by any method known in the art for purification of an immunoglobulin molecule, for example, by chromatography (e.g., ion

exchange, affinity, particularly by affinity for the specific antigen after Protein A, and sizing column chromatography), centrifugation, differential solubility, or by any other standard technique for the purification of proteins. In addition, the antibodies of the present invention or fragments thereof can be fused to heterologous polypeptide sequences described herein or otherwise known in the art, to facilitate purification.

[0398] The present invention encompasses antibodies recombinantly fused or chemically conjugated (including both covalently and non-covalently conjugations) to a polypeptide (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention to generate fusion proteins. The fusion does not necessarily need to be direct, but may occur through linker sequences. The antibodies may be specific for antigens other than polypeptides (or portion thereof, preferably at least 10, 20, 30, 40, 50, 60, 70, 80, 90 or 100 amino acids of the polypeptide) of the present invention. For example, antibodies may be used to target the polypeptides of the present invention to particular cell types, either in vitro or in vivo, by fusing or conjugating the polypeptides of the present invention to antibodies specific for particular cell surface receptors. Antibodies fused or conjugated to the polypeptides of the present invention may also be used in in vitro immunoassays and purification methods using methods known in the art. See e.g., Harbor et al., supra, and PCT publication WO 93/21232; EP 439,095; Naramura et al., Immunol. Lett. 39:91-99 (1994); U.S. Pat. No. 5,474,981; Gillies et al., PNAS 89:1428-1432 (1992); Fell et al., J. Immunol. 146:2446-2452(1991), which are incorporated by reference in their entireties.

[0399] The present invention further includes compositions comprising the polypeptides of the present invention fused or conjugated to antibody domains other than the variable regions. For example, the polypeptides of the present invention may be fused or conjugated to an antibody Fc region, or portion thereof. The antibody portion fused to a polypeptide of the present invention may comprise the constant region, hinge region, CH1 domain, CH2 domain, and CH3 domain or any combination of whole domains or portions thereof. The polypeptides may also be fused or conjugated to the above antibody portions to form multimers. For example, Fc portions fused to the polypeptides of the present invention can form dimers through disulfide bonding between the Fc portions. Higher multimeric forms can be made by fusing the polypeptides to portions of IgA and IgM. Methods for fusing or conjugating the polypeptides of the present invention to antibody portions are known in the art. See, e.g., U.S. Pat. Nos. 5,336,603; 5,622,929; 5,359,046; 5,349,053; 5,447,851; 5,112,946; EP 307,434; EP 367,166; PCT publications WO 96/04388; WO 91/06570; Ashkenazi et al., Proc. Natl. Acad. Sci. USA 88:10535-10539 (1991); Zheng et al., J. Immunol. 154:5590-5600 (1995); and Vil et al., Proc. Natl. Acad. Sci. USA 89:11337-11341(1992) (said references incorporated by reference in their entireties).

[0400] As discussed, supra, the polypeptides corresponding to a polypeptide, polypeptide fragment, or a variant of SEQ ID NO:2 may be fused or conjugated to the above antibody portions to increase the in vivo half life of the polypeptides or for use in immunoassays using methods known in the art. Further, the polypeptides corresponding to SEQ ID NO:2 may be fused or conjugated to the above antibody portions to facilitate purification. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP 394,827; Traunecker et al., Nature 331:84-86 (1988). The polypeptides of the present invention fused or conjugated to an antibody having disulfide-linked dimeric structures (due to the IgG) may also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995)). In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP A 232,262). Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, Bennett et al., J. Molecular Recognition 8:52-58 (1995); Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).

[0401] Moreover, the antibodies or fragments thereof of the present invention can be fused to marker sequences, such as a peptide to facilitate purification. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, Calif., 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Other peptide tags useful for purification include, but are not limited to, the "HA" tag, which corresponds to an epitope derived from the influenza hemagglutinin protein (Wilson et al., Cell 37:767 (1984)) and the "flag" tag.

[0402] The present invention further encompasses antibodies or fragments thereof conjugated to a diagnostic or therapeutic agent. The antibodies can be used diagnostically to, for example, monitor the development or progression of a tumor as part of a clinical testing procedure to, e.g., determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, radioactive materials, positron emitting metals using various positron emission tomographies, and nonradioactive paramagnetic metal ions. The detectable substance may be coupled or conjugated either directly to the antibody (or fragment thereof) or indirectly, through an intermediate (such as, for example, a linker known in the art) using techniques known in the art. See, for example, U.S. Pat. No. 4,741,900 for metal ions which can be conjugated to antibodies for use as diagnostics according to the present invention. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, beta-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine,

dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin; and examples of suitable radioactive material include 125I, 131I, 111In or 99Tc.

[0403] Further, an antibody or fragment thereof may be conjugated to a therapeutic moiety such as a cytotoxin, e.g., a cytostatic or cytocidal agent, a therapeutic agent or a radioactive metal ion, e.g., alpha-emitters such as, for example, 213Bi. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include paclitaxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, coichicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologues thereof. Therapeutic agents include, but are not limited to, antimetabolites (e.g., methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (e.g., mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (e.g., daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (e.g., dactinomycin (formerly actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (e.g., vincristine and vinblastine).

[0404] The conjugates of the invention can be used for modifying a given biological response, the therapeutic agent or drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, a-interferon, β-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator, an apoptotic agent, e.g., TNF-alpha, TNF-beta, AIM I (See, International Publication No. WO 97/33899), AIM II (See, International Publication No. WO 97/34911), Fas Ligand (Takahashi et al., Int. Immunol., 6:1567-1574 (1994)), VEGI (See, International Publication No. WO 99/23105), a thrombotic agent or an anti-angiogenic agent, e.g., angiostatin or endostatin; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophage colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

[0405] Antibodies may also be attached to solid supports, which are particularly useful for immunoassays or purification of the target antigen. Such solid supports include, but are not limited to, glass, cellulose, polyacrylamide, nylon, polystyrene, polyvinyl chloride or polypropylene.

[0406] Techniques for conjugating such therapeutic moiety to antibodies are well known, see, e.g., Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev. 62:119-58 (1982).

[0407] Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Pat. No. 4,676,980, which is incorporated herein by reference in its entirety.

[0408] An antibody, with or without a therapeutic moiety conjugated to it, administered alone or in combination with cytotoxic factor(s) and/or cytokine(s) can be used as a therapeutic.

[0409] The present invention also encompasses the creation of synthetic antibodies directed against the polypeptides of the present invention. One example of synthetic antibodies is described in Radrizzani, M., et al., Medicina, (Aires), 59(6):753-8, (1999)). Recently, a new class of synthetic antibodies has been described and are referred to as molecularly imprinted polymers (MIPs) (Semorex, Inc.). Antibodies, peptides, and enzymes are often used as molecular recognition elements in chemical and biological sensors. However, their lack of stability and signal transduction mechanisms limits their use as sensing devices. Molecularly imprinted polymers (MIPs) are capable of mimicking the function of biological receptors but with less stability constraints. Such polymers provide high sensitivity and selectivity while maintaining excellent thermal and mechanical stability. MIPs have the ability to bind to small molecules and to target molecules such as organics and proteins' with equal or greater potency than that of natural antibodies. These "super" MIPs have higher affinities for their target and thus require lower concentrations for efficacious binding.

[0410] During synthesis, the MIPs are imprinted so as to have complementary size, shape, charge and functional groups of the selected target by using the target molecule itself (such as a polypeptide, antibody, etc.), or a substance having a very similar structure, as its "print" or "template." MIPs can be derivatized with the same reagents afforded to antibodies. For example, fluorescent 'super' MIPs can be coated onto beads or wells for use in highly sensitive separations or assays, or for use in high throughput screening of proteins.

[0411] Moreover, MIPs based upon the structure of the polypeptide(s) of the present invention may be useful in screening for compounds that bind to the polypeptide(s) of the invention. Such a MIP would serve the role of a synthetic "receptor" by minimicking the native architecture of the polypeptide. In fact, the ability of a MIP to serve the role of a synthetic receptor has already been demonstrated for the estrogen receptor (Ye, L., Yu, Y., Mosbach, K, Analyst., 126(6):760-5, (2001); Dickert, F, L., Hayden, O., Halikias, K, P, Analyst., 126(6):766-71, (2001)). A synthetic receptor may either be mimicked in its entirety (e.g., as the entire protein), or mimicked as a series of short peptides corresponding to the protein (Rachkov, A., Minoura, N, Biochim,

Biophys, Acta., 1544(1-2):255-66, (2001)). Such a synthetic receptor MIPs may be employed in any one or more of the screening methods described elsewhere herein.

[0412] MIPs have also been shown to be useful in "sensing" the presence of its mimicked molecule (Cheng, Z., Wang, E., Yang, X, Biosens, Bioelectron., 16(3):179-85, (2001); Jenkins, A, L., Yin, R., Jensen, J. L, Analyst., 126(6):798-802, (2001); Jenkins, A, L., Yin, R., Jensen, J. L, Analyst., 126(6):798-802, (2001)). For example, a MIP designed using a polypeptide of the present invention may be used in assays designed to identify, and potentially quantitate, the level of said polypeptide in a sample. Such a MIP may be used as a substitute for any component described in the assays, or kits, provided herein (e.g., ELISA, etc.).

[0413] A number of methods may be employed to create MIPs to a specific receptor, ligand, polypeptide, peptide, organic molecule. Several preferred methods are described by Esteban et al in J. Anal, Chem., 370(7):795-802, (2001), which is hereby incorporated herein by reference in its entirety in addition to any references cited therein. Additional methods are known in the art and are encompassed by the present invention, such as for example, Hart, B, R., Shea, K, J. J. Am. Chem, Soc., 123(9):2072-3, (2001); and Quaglia, M., Chenon, K., Hall, A, J., De, Lorenzi, E., Sellergren, B, J. Am. Chem, Soc., 123(10):2146-54, (2001); which are hereby incorporated by reference in their entirety herein.

Uses for Antibodies Directed Against Polypeptides of the Invention

[0414] The antibodies of the present invention have various utilities. For example, such antibodies may be used in diagnostic assays to detect the presence or quantification of the polypeptides of the invention in a sample. Such a diagnostic assay may be comprised of at least two steps. The first, subjecting a sample with the antibody, wherein the sample is a tissue (e.g., human, animal, etc.), biological fluid (e.g., blood, urine, sputum, semen, amniotic fluid, saliva, etc.), biological extract (e.g., tissue or cellular homogenate, etc.), a protein microchip (e.g., See Arenkov P, et al., Anal Biochem., 278(2):123-131 (2000)), or a chromatography column, etc. And a second step involving the quantification of antibody bound to the substrate. Alternatively, the method may additionally involve a first step of attaching the antibody, either covalently, electrostatically, or reversibly, to a solid support, and a second step of subjecting the bound antibody to the sample, as defined above and elsewhere herein.

[0415] Various diagnostic assay techniques are known in the art, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogenous phases (Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc., (1987), pp147-158). The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as 2H, 14C, 32P, or 125I, a florescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase, green fluorescent protein, or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); Dafvid et al., Biochem., 13:1014 (1974); Pain et al., J. Immunol. Metho., 40:219(1981); and Nygren, J. Histochem. And Cytochem., 30:407 (1982).

[0416] Antibodies directed against the polypeptides of the present invention are useful for the affinity purification of such polypeptides from recombinant cell culture or natural sources. In this process, the antibodies against a particular polypeptide are immobilized on a suitable support, such as a Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the polypeptides to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except for the desired polypeptides, which are bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the desired polypeptide from the antibody.

Immunophenotyping

[0417] The antibodies of the invention may be utilized for immunophenotyping of cell lines and biological samples. The translation product of the gene of the present invention may be useful as a cell specific marker, or more specifically as a cellular marker that is differentially expressed at various stages of differentiation and/or maturation of particular cell types. Monoclonal antibodies directed against a specific epitope, or combination of epitopes, will allow for the screening of cellular populations expressing the marker. Various techniques can be utilized using monoclonal antibodies to screen for cellular populations expressing the marker(s), and include magnetic separation using antibody-coated magnetic beads, "panning" with antibody attached to a solid matrix (i.e., plate), and flow cytometry (See, e.g., U.S. Pat. No. 5,985,660; and Morrison et al., Cell, 96:737-49 (1999)).

[0418] These techniques allow for the screening of particular populations of cells, such as might be found with hematological malignancies (i.e. minimal residual disease (MRD) in acute leukemic patients) and "non-self" cells in transplantations to prevent Graft-versus-Host Disease (GVHD). Alternatively, these techniques allow for the screening of hematopoietic stem and progenitor cells capable of undergoing proliferation and/or differentiation, as might be found in human umbilical cord blood.

Assays for Antibody Binding

[0419] The antibodies of the invention may be assayed for immunospecific binding by any method known in the art. The immunoassays which can be used include but are not limited to competitive and non-competitive assay systems using techniques such as western blots, radioimmunoassays, ELISA (enzyme linked immunosorbent assay), "sandwich" immunoassays, immunoprecipitation assays, precipitin reactions, gel diffusion precipitin reactions, immunodiffusion assays, agglutination assays, complement-fixation assays, immunoradiometric assays, fluorescent immunoassays, protein A immunoassays, to name but a few. Such assays are routine and well known in the art (see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1,

John Wiley & Sons, Inc., New York, which is incorporated by reference herein in its entirety). Exemplary immunoassays are described briefly below (but are not intended by way of limitation).

[0420] Immunoprecipitation protocols generally comprise lysing a population of cells in a lysis buffer such as RIPA buffer (1% NP-40 or Triton X-100, 1% sodium deoxycholate, 0.1% SDS, 0.15 M NaCl, 0.01 M sodium phosphate at pH 7.2, 1% Trasylol) supplemented with protein phosphatase and/or protease inhibitors (e.g., EDTA, PMSF, aprotinin, sodium vanadate), adding the antibody of interest to the cell lysate, incubating for a period of time (e.g., 1-4 hours) at 4° C., adding protein A and/or protein G sepharose beads to the cell lysate, incubating for about an hour or more at 4° C., washing the beads in lysis buffer and resuspending the beads in SDS/sample buffer. The ability of the antibody of interest to immunoprecipitate a particular antigen can be assessed by, e.g., western blot analysis. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the binding of the antibody to an antigen and decrease the background (e.g., pre-clearing the cell lysate with sepharose beads). For further discussion regarding immunoprecipitation protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.16.1.

[0421] Western blot analysis generally comprises preparing protein samples, electrophoresis of the protein samples in a polyacrylamide gel (e.g., 8%-20% SDS-PAGE depending on the molecular weight of the antigen), transferring the protein sample from the polyacrylamide gel to a membrane such as nitrocellulose, PVDF or nylon, blocking the membrane in blocking solution (e.g., PBS with 3% BSA or non-fat milk), washing the membrane in washing buffer (e.g., PBS-Tween 20), blocking the membrane with primary antibody (the antibody of interest) diluted in blocking buffer, washing the membrane in washing buffer, blocking the membrane with a secondary antibody (which recognizes the primary antibody, e.g., an anti-human antibody) conjugated to an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) or radioactive molecule (e.g., 32P or 125I) diluted in blocking buffer, washing the membrane in wash buffer, and detecting the presence of the antigen. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected and to reduce the background noise. For further discussion regarding western blot protocols see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 10.8.1.

[0422] ELISAs comprise preparing antigen, coating the well of a 96 well microtiter plate with the antigen, adding the antibody of interest conjugated to a detectable compound such as an enzymatic substrate (e.g., horseradish peroxidase or alkaline phosphatase) to the well and incubating for a period of time, and detecting the presence of the antigen. In ELISAs the antibody of interest does not have to be conjugated to a detectable compound; instead, a second antibody (which recognizes the antibody of interest) conjugated to a detectable compound may be added to the well. Further, instead of coating the well with the antigen, the antibody may be coated to the well. In this case, a second antibody conjugated to a detectable compound may be added following the addition of the antigen of interest to the coated well. One of skill in the art would be knowledgeable as to the parameters that can be modified to increase the signal detected as well as other variations of ELISAs known in the art. For further discussion regarding ELISAs see, e.g., Ausubel et al, eds, 1994, Current Protocols in Molecular Biology, Vol. 1, John Wiley & Sons, Inc., New York at 11.2.1.

[0423] The binding affinity of an antibody to an antigen and the off-rate of an antibody-antigen interaction can be determined by competitive binding assays. One example of a competitive binding assay is a radioimmunoassay comprising the incubation of labeled antigen (e.g., 3H or 125I) with the antibody of interest in the presence of increasing amounts of unlabeled antigen, and the detection of the antibody bound to the labeled antigen. The affinity of the antibody of interest for a particular antigen and the binding off-rates can be determined from the data by scatchard plot analysis. Competition with a second antibody can also be determined using radioimmunoassays. In this case, the antigen is incubated with antibody of interest conjugated to a labeled compound (e.g., 3H or 125I) in the presence of increasing amounts of an unlabeled second antibody.

Therapeutic Uses of Antibodies

[0424] The present invention is further directed to antibody-based therapies which involve administering antibodies of the invention to an animal, preferably a mammal, and most preferably a human, patient for treating one or more of the disclosed diseases, disorders, or conditions. Therapeutic compounds of the invention include, but are not limited to, antibodies of the invention (including fragments, analogs and derivatives thereof as described herein) and nucleic acids encoding antibodies of the invention (including fragments, analogs and derivatives thereof and anti-idiotypic antibodies as described herein). The antibodies of the invention can be used to treat, inhibit or prevent diseases, disorders or conditions associated with aberrant expression and/or activity of a polypeptide of the invention, including, but not limited to, any one or more of the diseases, disorders, or conditions described herein. The treatment and/or prevention of diseases, disorders, or conditions associated with aberrant expression and/or activity of a polypeptide of the invention includes, but is not limited to, alleviating symptoms associated with those diseases, disorders or conditions. Antibodies of the invention may be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

[0425] A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

[0426] The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors (such as, e.g., IL-2, IL-3 and IL-7), for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

[0427] The antibodies of the invention may be administered alone or in combination with other types of treatments (e.g., radiation therapy, chemotherapy, hormonal therapy, immunotherapy and anti-tumor agents). Generally, administration of products of a species origin or species reactivity (in the case of antibodies) that is the same species as that of the patient is preferred. Thus, in a preferred embodiment, human antibodies, fragments derivatives, analogs, or nucleic acids, are administered to a human patient for therapy or prophylaxis.

[0428] It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of disorders related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides of the invention, including fragments thereof. Preferred binding affinities include those with a dissociation constant or Kd less than 5×10−2 M, 10−2 M, 5×10−3 M, 10−3 M, 5×10−4 M, 10−4 M, 5×10−5 M, 10−5 M, 5×10−6 M, 10−6 M, 5×10−7 M, 10−7 M, 5×10−8 M, 10−8 M, 5×10−9 M, 10−9 M, 5×10−10 M, 10−10 M, 5×10−11 M, 10−11 M, 5×10−12 M, 10−12 M, 5×10−13 M, 10−13 M, 5×10−14 M, 10−14 M, 5×10−15 M, and 10−15 M.

[0429] Antibodies directed against polypeptides of the present invention are useful for inhibiting allergic reactions in animals. For example, by administering a therapeutically acceptable dose of an antibody, or antibodies, of the present invention, or a cocktail of the present antibodies, or in combination with other antibodies of varying sources, the animal may not elicit an allergic response to antigens.

[0430] Likewise, one could envision cloning the gene encoding an antibody directed against a polypeptide of the present invention, said polypeptide having the potential to elicit an allergic and/or immune response in an organism, and transforming the organism with said antibody gene such that it is expressed (e.g., constitutively, inducibly, etc.) in the organism. Thus, the organism would effectively become resistant to an allergic response resulting from the ingestion or presence of such an immune/allergic reactive polypeptide. Moreover, such a use of the antibodies of the present invention may have particular utility in preventing and/or ameliorating autoimmune diseases and/or disorders, as such conditions are typically a result of antibodies being directed against endogenous proteins. For example, in the instance where the polypeptide of the present invention is responsible for modulating the immune response to auto-antigens, transforming the organism and/or individual with a construct comprising any of the promoters disclosed herein or otherwise known in the art, in addition, to a polynucleotide encoding the antibody directed against the polypeptide of the present invention could effective inhibit the organisms immune system from eliciting an immune response to the auto-antigen(s). Detailed descriptions of therapeutic and/or gene therapy applications of the present invention are provided elsewhere herein.

[0431] Alternatively, antibodies of the present invention could be produced in a plant (e.g., cloning the gene of the antibody directed against a polypeptide of the present invention, and transforming a plant with a suitable vector comprising said gene for constitutive expression of the antibody within the plant), and the plant subsequently ingested by an animal, thereby conferring temporary immunity to the animal for the specific antigen the antibody is directed towards (See, for example, U.S. Pat. Nos. 5,914,123 and 6,034,298).

[0432] In another embodiment, antibodies of the present invention, preferably polyclonal antibodies, more preferably monoclonal antibodies, and most preferably single-chain antibodies, can be used as a means of inhibiting gene expression of a particular gene, or genes, in a human, mammal, and/or other organism. See, for example, International Publication Number WO 00/05391, published Feb. 3, 2000, to Dow Agrosciences LLC. The application of such methods for the antibodies of the present invention are known in the art, and are more particularly described elsewhere herein.

[0433] In yet another embodiment, antibodies of the present invention may be useful for multimerizing the polypeptides of the present invention. For example, certain proteins may confer enhanced biological activity when present in a multimeric state (i.e., such enhanced activity may be due to the increased effective concentration of such proteins whereby more protein is available in a localized location).

Antibody-Based Gene Therapy

[0434] In a specific embodiment, nucleic acids comprising sequences encoding antibodies or functional derivatives thereof, are administered to treat, inhibit or prevent a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention, by way of gene therapy. Gene therapy refers to therapy performed by the administration to a subject of an expressed or expressible nucleic acid. In this embodiment of the invention, the nucleic acids produce their encoded protein that mediates a therapeutic effect.

[0435] Any of the methods for gene therapy available in the art can be used according to the present invention. Exemplary methods are described below.

[0436] For general reviews of the methods of gene therapy, see Goldspiel et al., Clinical Pharmacy 12:488-505 (1993); Wu and Wu, Biotherapy 3:87-95 (1991); Tolstoshev, Ann. Rev. Pharmacol. Toxicol. 32:573-596 (1993); Mulligan, Science 260:926-932 (1993); and Morgan and Anderson, Ann. Rev. Biochem. 62:191-217 (1993); May, TIBTECH 11(5):155-215 (1993). Methods commonly known in the art of recombinant DNA technology which can be used are described in Ausubel et al. (eds.), Current Protocols in Molecular Biology, John Wiley & Sons, NY (1993); and Kriegler, Gene Transfer and Expression, A Laboratory Manual, Stockton Press, NY (1990).

[0437] In a preferred aspect, the compound comprises nucleic acid sequences encoding an antibody, said nucleic acid sequences being part of expression vectors that express the antibody or fragments or chimeric proteins or heavy or light chains thereof in a suitable host. In particular, such nucleic acid sequences have promoters operably linked to the antibody coding region, said promoter being inducible or constitutive, and, optionally, tissue-specific. In another particular embodiment, nucleic acid molecules are used in which the antibody coding sequences and any other desired

sequences are flanked by regions that promote homologous recombination at a desired site in the genome, thus providing for intrachromosomal expression of the antibody encoding nucleic acids (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al., Nature 342:435-438 (1989). In specific embodiments, the expressed antibody molecule is a single chain antibody; alternatively, the nucleic acid sequences include sequences encoding both the heavy and light chains, or fragments thereof, of the antibody.

[0438] Delivery of the nucleic acids into a patient may be either direct, in which case the patient is directly exposed to the nucleic acid or nucleic acid-carrying vectors, or indirect, in which case, cells are first transformed with the nucleic acids in vitro, then transplanted into the patient. These two approaches are known, respectively, as in vivo or ex vivo gene therapy.

[0439] In a specific embodiment, the nucleic acid sequences are directly administered in vivo, where it is expressed to produce the encoded product. This can be accomplished by any of numerous methods known in the art, e.g., by constructing them as part of an appropriate nucleic acid expression vector and administering it so that they become intracellular, e.g., by infection using defective or attenuated retrovirals or other viral vectors (see U.S. Pat. No. 4,980,286), or by direct injection of naked DNA, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, encapsulation in liposomes, microparticles, or microcapsules, or by administering them in linkage to a peptide which is known to enter the nucleus, by administering it in linkage to a ligand subject to receptor-mediated endocytosis (see, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)) (which can be used to target cell types specifically expressing the receptors), etc. In another embodiment, nucleic acid-ligand complexes can be formed in which the ligand comprises a fusogenic viral peptide to disrupt endosomes, allowing the nucleic acid to avoid lysosomal degradation. In yet another embodiment, the nucleic acid can be targeted in vivo for cell specific uptake and expression, by targeting a specific receptor (see, e.g., PCT Publications WO 92/06180; WO 92/22635; WO92/20316; WO93/14188, WO 93/20221). Alternatively, the nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination (Koller and Smithies, Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); Zijlstra et al., Nature 342:435-438 (1989)).

[0440] In a specific embodiment, viral vectors that contains nucleic acid sequences encoding an antibody of the invention are used. For example, a retroviral vector can be used (see Miller et al., Meth. Enzymol. 217:581-599 (1993)). These retroviral vectors contain the components necessary for the correct packaging of the viral genome and integration into the host cell DNA. The nucleic acid sequences encoding the antibody to be used in gene therapy are cloned into one or more vectors, which facilitates delivery of the gene into a patient. More detail about retroviral vectors can be found in Boesen et al., Biotherapy 6:291-302 (1994), which describes the use of a retroviral vector to deliver the mdr1 gene to hematopoietic stem cells in order to make the stem cells more resistant to chemotherapy. Other references illustrating the use of retroviral vectors in gene therapy are: Clowes et al., J. Clin. Invest. 93:644-651 (1994); Kiem et al., Blood 83:1467-1473 (1994); Salmons and Gunzberg, Human Gene Therapy 4:129-141 (1993); and Grossman and Wilson, Curr. Opin. in Genetics and Devel. 3:110-114 (1993).

[0441] Adenoviruses are other viral vectors that can be used in gene therapy. Adenoviruses are especially attractive vehicles for delivering genes to respiratory epithelia. Adenoviruses naturally infect respiratory epithelia where they cause a mild disease. Other targets for adenovirus-based delivery systems are liver, the central nervous system, endothelial cells, and muscle. Adenoviruses have the advantage of being capable of infecting non-dividing cells. Kozarsky and Wilson, Current Opinion in Genetics and Development 3:499-503 (1993) present a review of adenovirus-based gene therapy. Bout et al., Human Gene Therapy 5:3-10 (1994) demonstrated the use of adenovirus vectors to transfer genes to the respiratory epithelia of rhesus monkeys. Other instances of the use of adenoviruses in gene therapy can be found in Rosenfeld et al., Science 252:431-434 (1991); Rosenfeld et al., Cell 68:143-155 (1992); Mastrangeli et al., J. Clin. Invest. 91:225-234 (1993); PCT Publication WO94/12649; and Wang, et al., Gene Therapy 2:775-783 (1995). In a preferred embodiment, adenovirus vectors are used.

[0442] Adeno-associated virus (AAV) has also been proposed for use in gene therapy (Walsh et al., Proc. Soc. Exp. Biol. Med. 204:289-300 (1993); U.S. Pat. No. 5,436,146).

[0443] Another approach to gene therapy involves transferring a gene to cells in tissue culture by such methods as electroporation, lipofection, calcium phosphate mediated transfection, or viral infection. Usually, the method of transfer includes the transfer of a selectable marker to the cells. The cells are then placed under selection to isolate those cells that have taken up and are expressing the transferred gene. Those cells are then delivered to a patient.

[0444] In this embodiment, the nucleic acid is introduced into a cell prior to administration in vivo of the resulting recombinant cell. Such introduction can be carried out by any method known in the art, including but not limited to transfection, electroporation, microinjection, infection with a viral or bacteriophage vector containing the nucleic acid sequences, cell fusion, chromosome-mediated gene transfer, microcell-mediated gene transfer, spheroplast fusion, etc. Numerous techniques are known in the art for the introduction of foreign genes into cells (see, e.g., Loeffler and Behr, Meth. Enzymol. 217:599-618 (1993); Cohen et al., Meth. Enzymol. 217:618-644 (1993); Cline, Pharmac. Ther. 29:69-92m (1985) and may be used in accordance with the present invention, provided that the necessary developmental and physiological functions of the recipient cells are not disrupted. The technique should provide for the stable transfer of the nucleic acid to the cell, so that the nucleic acid is expressible by the cell and preferably heritable and expressible by its cell progeny.

[0445] The resulting recombinant cells can be delivered to a patient by various methods known in the art. Recombinant blood cells (e.g., hematopoietic stem or progenitor cells) are preferably administered intravenously. The amount of cells envisioned for use depends on the desired effect, patient state, etc., and can be determined by one skilled in the art.

[0446] Cells into which a nucleic acid can be introduced for purposes of gene therapy encompass any desired, avail-

able cell type, and include but are not limited to epithelial cells, endothelial cells, keratinocytes, fibroblasts, muscle cells, hepatocytes; blood cells such as Tlymphocytes, Blymphocytes, monocytes, macrophages, neutrophils, eosinophils, megakaryocytes, granulocytes; various stem or progenitor cells, in particular hematopoietic stem or progenitor cells, e.g., as obtained from bone marrow, umbilical cord blood, peripheral blood, fetal liver, etc.

[0447] In a preferred embodiment, the cell used for gene therapy is autologous to the patient.

[0448] In an embodiment in which recombinant cells are used in gene therapy, nucleic acid sequences encoding an antibody are introduced into the cells such that they are expressible by the cells or their progeny, and the recombinant cells are then administered in vivo for therapeutic effect. In a specific embodiment, stem or progenitor cells are used. Any stem and/or progenitor cells which can be isolated and maintained in vitro can potentially be used in accordance with this embodiment of the present invention (see e.g. PCT Publication WO 94/08598; Stemple and Anderson, Cell 71:973-985 (1992); Rheinwald, Meth. Cell Bio. 21A:229 (1980); and Pittelkow and Scott, Mayo Clinic Proc. 61:771 (1986)).

[0449] In a specific embodiment, the nucleic acid to be introduced for purposes of gene therapy comprises an inducible promoter operably linked to the coding region, such that expression of the nucleic acid is controllable by controlling the presence or absence of the appropriate inducer of transcription. Demonstration of Therapeutic or Prophylactic Activity

[0450] The compounds or pharmaceutical compositions of the invention are preferably tested in vitro, and then in vivo for the desired therapeutic or prophylactic activity, prior to use in humans. For example, in vitro assays to demonstrate the therapeutic or prophylactic utility of a compound or pharmaceutical composition include, the effect of a compound on a cell line or a patient tissue sample. The effect of the compound or composition on the cell line and/or tissue sample can be determined utilizing techniques known to those of skill in the art including, but not limited to, rosette formation assays and cell lysis assays. In accordance with the invention, in vitro assays which can be used to determine whether administration of a specific compound is indicated, include in vitro cell culture assays in which a patient tissue sample is grown in culture, and exposed to or otherwise administered a compound, and the effect of such compound upon the tissue sample is observed.

Therapeutic/Prophylactic Administration and Compositions

[0451] The invention provides methods of treatment, inhibition and prophylaxis by administration to a subject of an effective amount of a compound or pharmaceutical composition of the invention, preferably an antibody of the invention. In a preferred aspect, the compound is substantially purified (e.g., substantially free from substances that limit its effect or produce undesired side-effects). The subject is preferably an animal, including but not limited to animals such as cows, pigs, horses, chickens, cats, dogs, etc., and is preferably a mammal, and most preferably human.

[0452] Formulations and methods of administration that can be employed when the compound comprises a nucleic acid or an immunoglobulin are described above; additional appropriate formulations and routes of administration can be selected from among those described herein below.

[0453] Various delivery systems are known and can be used to administer a compound of the invention, e.g., encapsulation in liposomes, microparticles, microcapsules, recombinant cells capable of expressing the compound, receptor-mediated endocytosis (see, e.g., Wu and Wu, J. Biol. Chem. 262:4429-4432 (1987)), construction of a nucleic acid as part of a retroviral or other vector, etc. Methods of introduction include but are not limited to intradermal, intramuscular, intraperitoneal, intravenous, subcutaneous, intranasal, epidural, and oral routes. The compounds or compositions may be administered by any convenient route, for example by infusion or bolus injection, by absorption through epithelial or mucocutaneous linings (e.g., oral mucosa, rectal and intestinal mucosa, etc.) and may be administered together with other biologically active agents. Administration can be systemic or local. In addition, it may be desirable to introduce the pharmaceutical compounds or compositions of the invention into the central nervous system by any suitable route, including intraventricular and intrathecal injection; intraventricular injection may be facilitated by an intraventricular catheter, for example, attached to a reservoir, such as an Ommaya reservoir. Pulmonary administration can also be employed, e.g., by use of an inhaler or nebulizer, and formulation with an aerosolizing agent.

[0454] In a specific embodiment, it may be desirable to administer the pharmaceutical compounds or compositions of the invention locally to the area in need of treatment; this may be achieved by, for example, and not by way of limitation, local infusion during surgery, topical application, e.g., in conjunction with a wound dressing after surgery, by injection, by means of a catheter, by means of a suppository, or by means of an implant, said implant being of a porous, non-porous, or gelatinous material, including membranes, such as sialastic membranes, or fibers. Preferably, when administering a protein, including an antibody, of the invention, care must be taken to use materials to which the protein does not absorb.

[0455] In another embodiment, the compound or composition can be delivered in a vesicle, in particular a liposome (see Langer, Science 249:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 353-365 (1989); Lopez-Berestein, ibid., pp. 317-327; see generally ibid.)

[0456] In yet another embodiment, the compound or composition can be delivered in a controlled release system. In one embodiment, a pump may be used (see Langer, supra; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)). In another embodiment, polymeric materials can be used (see Medical Applications of Controlled Release, Langer and Wise (eds.), CRC Pres., Boca Raton, Fla. (1974); Controlled Drug Bioavailability, Drug Product Design and Performance, Smolen and Ball (eds.), Wiley, New York (1984); Ranger and Peppas, J., Macromol. Sci. Rev. Macromol. Chem. 23:61 (1983); see also Levy et al., Science 228:190 (1985); During et al., Ann. Neurol. 25:351 (1989); Howard et al., J. Neurosurg. 71:105 (1989)).

In yet another embodiment, a controlled release system can be placed in proximity of the therapeutic target, i.e., the brain, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138 (1984)).

[0457] Other controlled release systems are discussed in the review by Langer (Science 249:1527-1533 (1990)).

[0458] In a specific embodiment where the compound of the invention is a nucleic acid encoding a protein, the nucleic acid can be administered in vivo to promote expression of its encoded protein, by constructing it as part of an appropriate nucleic acid expression vector and administering it so that it becomes intracellular, e.g., by use of a retroviral vector (see U.S. Pat. No. 4,980,286), or by direct injection, or by use of microparticle bombardment (e.g., a gene gun; Biolistic, Dupont), or coating with lipids or cell-surface receptors or transfecting agents, or by administering it in linkage to a homeobox-like peptide which is known to enter the nucleus (see e.g., Joliot et al., Proc. Natl. Acad. Sci. USA 88:1864-1868 (1991)), etc. Alternatively, a nucleic acid can be introduced intracellularly and incorporated within host cell DNA for expression, by homologous recombination.

[0459] The present invention also provides pharmaceutical compositions. Such compositions comprise a therapeutically effective amount of a compound, and a pharmaceutically acceptable carrier. In a specific embodiment, the term "pharmaceutically acceptable" means approved by a regulatory agency of the Federal or a state government or listed in the U.S. Pharmacopeia or other generally recognized pharmacopeia for use in animals, and more particularly in humans. The term "carrier" refers to a diluent, adjuvant, excipient, or vehicle with which the therapeutic is administered. Such pharmaceutical carriers can be sterile liquids, such as water and oils, including those of petroleum, animal, vegetable or synthetic origin, such as peanut oil, soybean oil, mineral oil, sesame oil and the like. Water is a preferred carrier when the pharmaceutical composition is administered intravenously. Saline solutions and aqueous dextrose and glycerol solutions can also be employed as liquid carriers, particularly for injectable solutions. Suitable pharmaceutical excipients include starch, glucose, lactose, sucrose, gelatin, malt, rice, flour, chalk, silica gel, sodium stearate, glycerol monostearate, talc, sodium chloride, dried skim milk, glycerol, propylene, glycol, water, ethanol and the like. The composition, if desired, can also contain minor amounts of wetting or emulsifying agents, or pH buffering agents. These compositions can take the form of solutions, suspensions, emulsion, tablets, pills, capsules, powders, sustained-release formulations and the like. The composition can be formulated as a suppository, with traditional binders and carriers such as triglycerides. Oral formulation can include standard carriers such as pharmaceutical grades of mannitol, lactose, starch, magnesium stearate, sodium saccharine, cellulose, magnesium carbonate, etc. Examples of suitable pharmaceutical carriers are described in "Remington's Pharmaceutical Sciences" by E. W. Martin. Such compositions will contain a therapeutically effective amount of the compound, preferably in purified form, together with a suitable amount of carrier so as to provide the form for proper administration to the patient. The formulation should suit the mode of administration.

[0460] In a preferred embodiment, the composition is formulated in accordance with routine procedures as a pharmaceutical composition adapted for intravenous administration to human beings. Typically, compositions for intravenous administration are solutions in sterile isotonic aqueous buffer. Where necessary, the composition may also include a solubilizing agent and a local anesthetic such as lignocaine to ease pain at the site of the injection. Generally, the ingredients are supplied either separately or mixed together in unit dosage form, for example, as a dry lyophilized powder or water free concentrate in a hermetically sealed container such as an ampoule or sachette indicating the quantity of active agent. Where the composition is to be administered by infusion, it can be dispensed with an infusion bottle containing sterile pharmaceutical grade water or saline. Where the composition is administered by injection, an ampoule of sterile water for injection or saline can be provided so that the ingredients may be mixed prior to administration.

[0461] The compounds of the invention can be formulated as neutral or salt forms. Pharmaceutically acceptable salts include those formed with anions such as those derived from hydrochloric, phosphoric, acetic, oxalic, tartaric acids, etc., and those formed with cations such as those derived from sodium, potassium, ammonium, calcium, ferric hydroxides, isopropylamine, triethylamine, 2-ethylamino ethanol, histidine, procaine, etc.

[0462] The amount of the compound of the invention which will be effective in the treatment, inhibition and prevention of a disease or disorder associated with aberrant expression and/or activity of a polypeptide of the invention can be determined by standard clinical techniques. In addition, in vitro assays may optionally be employed to help identify optimal dosage ranges. The precise dose to be employed in the formulation will also depend on the route of administration, and the seriousness of the disease or disorder, and should be decided according to the judgment of the practitioner and each patient's circumstances. Effective doses may be extrapolated from dose-response curves derived from in vitro or animal model test systems.

[0463] For antibodies, the dosage administered to a patient is typically 0.1 mg/kg to 100 mg/kg of the patient's body weight. Preferably, the dosage administered to a patient is between 0.1 mg/kg and 20 mg/kg of the patient's body weight, more preferably 1 mg/kg to 10 mg/kg of the patient's body weight. Generally, human antibodies have a longer half-life within the human body than antibodies from other species due to the immune response to the foreign polypeptides. Thus, lower dosages of human antibodies and less frequent administration is often possible. Further, the dosage and frequency of administration of antibodies of the invention may be reduced by enhancing uptake and tissue penetration (e.g., into the brain) of the antibodies by modifications such as, for example, lipidation.

[0464] The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the pharmaceutical compositions of the invention. Optionally associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration.

Diagnosis and Imaging with Antibodies

[0465] Labeled antibodies, and derivatives and analogs thereof, which specifically bind to a polypeptide of interest can be used for diagnostic purposes to detect, diagnose, or monitor diseases, disorders, and/or conditions associated with the aberrant expression and/or activity of a polypeptide of the invention. The invention provides for the detection of aberrant expression of a polypeptide of interest, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of aberrant expression.

[0466] The invention provides a diagnostic assay for diagnosing a disorder, comprising (a) assaying the expression of the polypeptide of interest in cells or body fluid of an individual using one or more antibodies specific to the polypeptide interest and (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a particular disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

[0467] Antibodies of the invention can be used to assay protein levels in a biological sample using classical immunohistological methods known to those of skill in the art (e.g., see Jalkanen, et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, et al., J. Cell. Biol. 105:3087-3096 (1987)). Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase; radioisotopes, such as iodine (1251, 1211), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99Tc); luminescent labels, such as luminol; and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0468] One aspect of the invention is the detection and diagnosis of a disease or disorder associated with aberrant expression of a polypeptide of interest in an animal, preferably a mammal and most preferably a human. In one embodiment, diagnosis comprises: a) administering (for example, parenterally, subcutaneously, or intraperitoneally) to a subject an effective amount of a labeled molecule which specifically binds to the polypeptide of interest; b) waiting for a time interval following the administering for permitting the labeled molecule to preferentially concentrate at sites in the subject where the polypeptide is expressed (and for unbound labeled molecule to be cleared to background level); c) determining background level; and d) detecting the labeled molecule in the subject, such that detection of labeled molecule above the background level indicates that the subject has a particular disease or disorder associated with aberrant expression of the polypeptide of interest. Background level can be determined by various methods including, comparing the amount of labeled molecule detected to a standard value previously determined for a particular system.

[0469] It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S. W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S. W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).

[0470] Depending on several variables, including the type of label used and the mode of administration, the time interval following the administration for permitting the labeled molecule to preferentially concentrate at sites in the subject and for unbound labeled molecule to be cleared to background level is 6 to 48 hours or 6 to 24 hours or 6 to 12 hours. In another embodiment the time interval following administration is 5 to 20 days or 5 to 10 days.

[0471] In an embodiment, monitoring of the disease or disorder is carried out by repeating the method for diagnosing the disease or disease, for example, one month after initial diagnosis, six months after initial diagnosis, one year after initial diagnosis, etc.

[0472] Presence of the labeled molecule can be detected in the patient using methods known in the art for in vivo scanning. These methods depend upon the type of label used. Skilled artisans will be able to determine the appropriate method for detecting a particular label. Methods and devices that may be used in the diagnostic methods of the invention include, but are not limited to, computed tomography (CT), whole body scan such as position emission tomography (PET), magnetic resonance imaging (MRI), and sonography.

[0473] In a specific embodiment, the molecule is labeled with a radioisotope and is detected in the patient using a radiation responsive surgical instrument (Thurston et al., U.S. Pat. No. 5,441,050). In another embodiment, the molecule is labeled with a fluorescent compound and is detected in the patient using a fluorescence responsive scanning instrument. In another embodiment, the molecule is labeled with a positron emitting metal and is detected in the patent using positron emission-tomography. In yet another embodiment, the molecule is labeled with a paramagnetic label and is detected in a patient using magnetic resonance imaging (MRI).

Kits

[0474] The present invention provides kits that can be used in the above methods. In one embodiment, a kit comprises an antibody of the invention, preferably a purified antibody, in one or more containers. In a specific embodiment, the kits of the present invention contain a substantially

isolated polypeptide comprising an epitope which is specifically immunoreactive with an antibody included in the kit. Preferably, the kits of the present invention further comprise a control antibody which does not react with the polypeptide of interest. In another specific embodiment, the kits of the present invention contain a means for detecting the binding of an antibody to a polypeptide of interest (e.g., the antibody may be conjugated to a detectable substrate such as a fluorescent compound, an enzymatic substrate, a radioactive compound or a luminescent compound, or a second antibody which recognizes the first antibody may be conjugated to a detectable substrate).

[0475] In another specific embodiment of the present invention, the kit is a diagnostic kit for use in screening serum containing antibodies specific against proliferative and/or cancerous polynucleotides and polypeptides. Such a kit may include a control antibody that does not react with the polypeptide of interest. Such a kit may include a substantially isolated polypeptide antigen comprising an epitope which is specifically immunoreactive with at least one anti-polypeptide antigen antibody. Further, such a kit includes means for detecting the binding of said antibody to the antigen (e.g., the antibody may be conjugated to a fluorescent compound such as fluorescein or rhodamine which can be detected by flow cytometry). In specific embodiments, the kit may include a recombinantly produced or chemically synthesized polypeptide antigen. The polypeptide antigen of the kit may also be attached to a solid support.

[0476] In a more specific embodiment the detecting means of the above-described kit includes a solid support to which said polypeptide antigen is attached. Such a kit may also include a non-attached reporter-labeled anti-human antibody. In this embodiment, binding of the antibody to the polypeptide antigen can be detected by binding of the said reporter-labeled antibody.

[0477] In an additional embodiment, the invention includes a diagnostic kit for use in screening serum containing antigens of the polypeptide of the invention. The diagnostic kit includes a substantially isolated antibody specifically immunoreactive with polypeptide or polynucleotide antigens, and means for detecting the binding of the polynucleotide or polypeptide antigen to the antibody. In one embodiment, the antibody is attached to a solid support. In a specific embodiment, the antibody may be a monoclonal antibody. The detecting means of the kit may include a second, labeled monoclonal antibody. Alternatively, or in addition, the detecting means may include a labeled, competing antigen.

[0478] In one diagnostic configuration, test serum is reacted with a solid phase reagent having a surface-bound antigen obtained by the methods of the present invention. After binding with specific antigen antibody to the reagent and removing unbound serum components by washing, the reagent is reacted with reporter-labeled anti-human antibody to bind reporter to the reagent in proportion to the amount of bound anti-antigen antibody on the solid support. The reagent is again washed to remove unbound labeled antibody, and the amount of reporter associated with the reagent is determined. Typically, the reporter is an enzyme which is detected by incubating the solid phase in the presence of a suitable fluorometric, luminescent or calorimetric substrate (Sigma, St. Louis, Mo.).

[0479] The solid surface reagent in the above assay is prepared by known techniques for attaching protein material to solid support material, such as polymeric beads, dip sticks, 96-well plate or filter material. These attachment methods generally include non-specific adsorption of the protein to the support or covalent attachment of the protein, typically through a free amine group, to a chemically reactive group on the solid support, such as an activated carboxyl, hydroxyl, or aldehyde group. Alternatively, streptavidin coated plates can be used in conjunction with biotinylated antigen(s).

[0480] Thus, the invention provides an assay system or kit for carrying out this diagnostic method. The kit generally includes a support with surface-bound recombinant antigens, and a reporter-labeled anti-human antibody for detecting surface-bound anti-antigen antibody.

Fusion Proteins

[0481] Any polypeptide of the present invention can be used to generate fusion proteins. For example, the polypeptide of the present invention, when fused to a second protein, can be used as an antigenic tag. Antibodies raised against the polypeptide of the present invention can be used to indirectly detect the second protein by binding to the polypeptide. Moreover, because certain proteins target cellular locations based on trafficking signals, the polypeptides of the present invention can be used as targeting molecules once fused to other proteins.

[0482] Examples of domains that can be fused to polypeptides of the present inventions include not only heterologous signal sequences, but also other heterologous functional regions. The fusion does not necessarily need to be direct, but may occur through linker sequences.

[0483] Moreover, fusion proteins may also be engineered to improve characteristics of the polypeptide of the present invention. For instance, a region of additional amino acids, particularly charged amino acids, may be added to the N-terminus of the polypeptide to improve stability and persistence during purification from the host cell or subsequent handling and storage. Peptide moieties may be added to the polypeptide to facilitate purification. Such regions may be removed prior to final preparation of the polypeptide. Similarly, peptide cleavage sites can be introduced in-between such peptide moieties, which could additionally be subjected to protease activity to remove said peptide(s) from the protein of the present invention. The addition of peptide moieties, including peptide cleavage sites, to facilitate handling of polypeptides are familiar and routine techniques in the art.

[0484] Moreover, polypeptides of the present invention, including fragments, and specifically epitopes, can be combined with parts of the constant domain of immunoglobulins (IgA, IgE, IgG, IgM) or portions thereof (CH1, CH2, CH3, and any combination thereof, including both entire domains and portions thereof), resulting in chimeric polypeptides. These fusion proteins facilitate purification and show an increased half-life in vivo. One reported example describes chimeric proteins consisting of the first two domains of the human CD4-polypeptide and various domains of the constant regions of the heavy or light chains of mammalian immunoglobulins. (EP A 394,827; Traunecker et al., Nature 331:84-86 (1988).) Fusion proteins having disulfide-linked

dimeric structures (due to the IgG) can also be more efficient in binding and neutralizing other molecules, than the monomeric secreted protein or protein fragment alone. (Fountoulakis et al., J. Biochem. 270:3958-3964 (1995).)

[0485] Similarly, EP-A-O 464 533 (Canadian counterpart 2045869) discloses fusion proteins comprising various portions of the constant region of immunoglobulin-molecules together with another human protein or part thereof. In many cases, the Fc part in a fusion protein is beneficial in therapy and diagnosis, and thus can result in, for example, improved pharmacokinetic properties. (EP-A 0232 262.) Alternatively, deleting the Fc part after the fusion protein has been expressed, detected, and purified, would be desired. For example, the Fc portion may hinder therapy and diagnosis if the fusion protein is used as an antigen for immunizations. In drug discovery, for example, human proteins, such as hIL-5, have been fused with Fc portions for the purpose of high-throughput screening assays to identify antagonists of hIL-5. (See, D. Bennett et al., J. Molecular Recognition 8:52-58 (1995); K. Johanson et al., J. Biol. Chem. 270:9459-9471 (1995).)

[0486] Moreover, the polypeptides of the present invention can be fused to marker sequences (also referred to as "tags"). Due to the availability of antibodies specific to such "tags", purification of the fused polypeptide of the invention, and/or its identification is significantly facilitated since antibodies specific to the polypeptides of the invention are not required. Such purification may be in the form of an affinity purification whereby an anti-tag antibody or another type of affinity matrix (e.g., anti-tag antibody attached to the matrix of a flow-thru column) that binds to the epitope tag is present. In preferred embodiments, the marker amino acid sequence is a hexa-histidine peptide, such as the tag provided in a pQE vector (QIAGEN, Inc., 9259 Eton Avenue, Chatsworth, Calif., 91311), among others, many of which are commercially available. As described in Gentz et al., Proc. Natl. Acad. Sci. USA 86:821-824 (1989), for instance, hexa-histidine provides for convenient purification of the fusion protein. Another peptide tag useful for purification, the "HA" tag, corresponds to an epitope derived from the influenza hemagglutinin protein. (Wilson et al., Cell 37:767 (1984)).

[0487] The skilled artisan would acknowledge the existence of other "tags" which could be readily substituted for the tags referred to supra for purification and/or identification of polypeptides of the present invention (Jones C., et al., J Chromatogr A. 707(1):3-22 (1995)). For example, the c-myc tag and the 8F9, 3C7, 6E10, G4m B7 and 9E10 antibodies thereto (Evan et al., Molecular and Cellular Biology 5:3610-3616 (1985)); the Herpes Simplex virus glycoprotein D (gD) tag and its antibody (Paborsky et al., Protein Engineering, 3(6):547-553 (1990), the Flag-peptide—i.e., the octapeptide sequence DYKDDDDK (SEQ ID NO:66), (Hopp et al., Biotech. 6:1204-1210 (1988); the KT3 epitope peptide (Martin et al., Science, 255:192-194 (1992)); a-tubulin epitope peptide (Skinner et al., J. Biol. Chem., 266:15136-15166, (1991)); the T7 gene 10 protein peptide tag (Lutz-Freyermuth et al., Proc. Natl. Sci. USA, 87:6363-6397 (1990)), the FITC epitope (Zymed, Inc.), the GFP epitope (Zymed, Inc.), and the Rhodamine epitope (Zymed, Inc.).

[0488] The present invention also encompasses the attachment of up to nine codons encoding a repeating series of up to nine arginine amino acids to the coding region of a polynucleotide of the present invention. The invention also encompasses chemically derivitizing a polypeptide of the present invention with a repeating series of up to nine arginine amino acids. Such a tag, when attached to a polypeptide, has recently been shown to serve as a universal pass, allowing compounds access to the interior of cells without additional derivitization or manipulation (Wender, P., et al., unpublished data).

[0489] Protein fusions involving polypeptides of the present invention, including fragments and/or variants thereof, can be used for the following, non-limiting examples, subcellular localization of proteins, determination of protein-protein interactions via immunoprecipitation, purification of proteins via affinity chromatography, functional and/or structural characterization of protein. The present invention also encompasses the application of hapten specific antibodies for any of the uses referenced above for epitope fusion proteins. For example, the polypeptides of the present invention could be chemically derivatized to attach hapten molecules (e.g., DNP, (Zymed, Inc.)). Due to the availability of monoclonal antibodies specific to such haptens, the protein could be readily purified using immunoprecipation, for example.

[0490] Polypeptides of the present invention, including fragments and/or variants thereof, in addition to, antibodies directed against such polypeptides, fragments, and/or variants, may be fused to any of a number of known, and yet to be determined, toxins, such as ricin, saporin (Mashiba H, et al., Ann. N.Y. Acad. Sci. 1999;886:233-5), or HC toxin (Tonukari N J, et al., Plant Cell. 2000 Feb; 12(2):237-248), for example. Such fusions could be used to deliver the toxins to desired tissues for which a ligand or a protein capable of binding to the polypeptides of the invention exists.

[0491] The invention encompasses the fusion of antibodies directed against polypeptides of the present invention, including variants and fragments thereof, to said toxins for delivering the toxin to specific locations in a cell, to specific tissues, and/or to specific species. Such bifunctional antibodies are known in the art, though a review describing additional advantageous fusions, including citations for methods of production, can be found in P. J. Hudson, Curr. Opp. In. Imm. 11:548-557, (1999); this publication, in addition to the references cited therein, are hereby incorporated by reference in their entirety herein. In this context, the term "toxin" may be expanded to include any heterologous protein, a small molecule, radionucleotides, cytotoxic drugs, liposomes, adhesion molecules, glycoproteins, ligands, cell or tissue-specific ligands, enzymes, of bioactive agents, biological response modifiers, anti-fungal agents, hormones, steroids, vitamins, peptides, peptide analogs, anti-allergenic agents, anti-tubercular agents, anti-viral agents, antibiotics, anti-protozoan agents, chelates, radioactive particles, radioactive ions, X-ray contrast agents, monoclonal antibodies, polyclonal antibodies and genetic material. In view of the present disclosure, one skilled in the art could determine whether any particular "toxin" could be used in the compounds of the present invention. Examples of suitable "toxins" listed above are exemplary only and are not intended to limit the "toxins" that may be used in the present invention.

[0492] Thus, any of these above fusions can be engineered using the polynucleotides or the polypeptides of the present invention.

Vectors, Host Cells, and Protein Production

[0493] The present invention also relates to vectors containing the polynucleotide of the present invention, host cells, and the production of polypeptides by recombinant techniques. The vector may be, for example, a phage, plasmid, viral, or retroviral vector. Retroviral vectors may be replication competent or replication defective. In the latter case, viral propagation generally will occur only in complementing host cells.

[0494] The polynucleotides may be joined to a vector containing a selectable marker for propagation in a host. Generally, a plasmid vector is introduced in a precipitate, such as a calcium phosphate precipitate, or in a complex with a charged lipid. If the vector is a virus, it may be packaged in vitro using an appropriate packaging cell line and then transduced into host cells.

[0495] The polynucleotide insert should be operatively linked to an appropriate promoter, such as the phage lambda PL promoter, the *E. coli* lac, trp, phoA and tac promoters, the SV40 early and late promoters and promoters of retroviral LTRs, to name a few. Other suitable promoters will be known to the skilled artisan. The expression constructs will further contain sites for transcription initiation, termination, and, in the transcribed region, a ribosome binding site for translation. The coding portion of the transcripts expressed by the constructs will preferably include a translation initiating codon at the beginning and a termination codon (UAA, UGA or UAG) appropriately positioned at the end of the polypeptide to be translated.

[0496] As indicated, the expression vectors will preferably include at least one selectable marker. Such markers include dihydrofolate reductase, G418 or neomycin resistance for eukaryotic cell culture and tetracycline, kanamycin or ampicillin resistance genes for culturing in *E. coli* and other bacteria. Representative examples of appropriate hosts include, but are not limited to, bacterial cells, such as *E. coli*, Streptomyces and *Salmonella typhimurium* cells; fungal cells, such as yeast cells (e.g., *Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201178)); insect cells such as Drosophila S2 and Spodoptera Sf9 cells; animal cells such as CHO, COS, 293, and Bowes melanoma cells; and plant cells. Appropriate culture mediums and conditions for the above-described host cells are known in the art.

[0497] Among vectors preferred for use in bacteria include pQE70, pQE60 and pQE-9, available from QIAGEN, Inc.; pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A, available from Stratagene Cloning Systems, Inc.; and ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 available from Pharmacia Biotech, Inc. Among preferred eukaryotic vectors are pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; and pSVK3, pBPV, pMSG and pSVL available from Pharmacia. Preferred expression vectors for use in yeast systems include, but are not limited to pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, pPIC9K, and PAO815 (all available from Invitrogen, Carlsbad, Calif.). Other suitable vectors will be readily apparent to the skilled artisan.

[0498] Introduction of the construct into the host cell can be effected by calcium phosphate transfection, DEAE-dextran mediated transfection, cationic lipid-mediated transfection, electroporation, transduction, infection, or other methods. Such methods are described in many standard laboratory manuals, such as Davis et al., Basic Methods In Molecular Biology (1986). It is specifically contemplated that the polypeptides of the present invention may in fact be expressed by a host cell lacking a recombinant vector.

[0499] A polypeptide of this invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography ("HPLC") is employed for purification.

[0500] Polypeptides of the present invention, and preferably the secreted form, can also be recovered from: products purified from natural sources, including bodily fluids, tissues and cells, whether directly isolated or cultured; products of chemical synthetic procedures; and products produced by recombinant techniques from a prokaryotic or eukaryotic host, including, for example, bacterial, yeast, higher plant, insect, and mammalian cells. Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be non-glycosylated. In addition, polypeptides of the invention may also include an initial modified methionine residue, in some cases as a result of host-mediated processes. Thus, it is well known in the art that the N-terminal methionine encoded by the translation initiation codon generally is removed with high efficiency from any protein after translation in all eukaryotic cells. While the N-terminal methionine on most proteins also is efficiently removed in most prokaryotes, for some proteins, this prokaryotic removal process is inefficient, depending on the nature of the amino acid to which the N-terminal methionine is covalently linked.

[0501] In one embodiment, the yeast *Pichia pastoris* is used to express the polypeptide of the present invention in a eukaryotic system. *Pichia pastoris* is a methylotrophic yeast which can metabolize methanol as its sole carbon source. A main step in the methanol metabolization pathway is the oxidation of methanol to formaldehyde using O_2. This reaction is catalyzed by the enzyme alcohol oxidase. In order to metabolize methanol as its sole carbon source, *Pichia pastoris* must generate high levels of alcohol oxidase due, in part, to the relatively low affinity of alcohol oxidase for O_2. Consequently, in a growth medium depending on methanol as a main carbon source, the promoter region of one of the two alcohol oxidase genes (AOX1) is highly active. In the presence of methanol, alcohol oxidase produced from the AOX1 gene comprises up to approximately 30% of the total soluble protein in *Pichia pastoris*. See, Ellis, S. B., et al., Mol. Cell. Biol. 5:1111-21 (1985); Koutz, P. J, et al., Yeast 5:167-77 (1989); Tschopp, J. F., et al., Nucl. Acids Res. 15:3859-76 (1987). Thus, a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, under the transcriptional regulation of all or part of the AOX1 regulatory sequence is expressed at exceptionally high levels in Pichia yeast grown in the presence of methanol.

[0502] In one example, the plasmid vector pPIC9K is used to express DNA encoding a polypeptide of the invention, as

set forth herein, in a Pichea yeast system essentially as described in "Pichia Protocols: Methods in Molecular Biology," D. R. Higgins and J. Cregg, eds. The Humana Press, Totowa, N.J., 1998. This expression vector allows expression and secretion of a protein of the invention by virtue of the strong AOX1 promoter linked to the *Pichia pastoris* alkaline phosphatase (PHO) secretory signal peptide (i.e., leader) located upstream of a multiple cloning site.

[0503] Many other yeast vectors could be used in place of pPIC9K, such as, pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalpha, pPIC9, pPIC3.5, pHIL-D2, pHIL-S1, pPIC3.5K, and PAO815, as one skilled in the art would readily appreciate, as long as the proposed expression construct provides appropriately located signals for transcription, translation, secretion (if desired), and the like, including an in-frame AUG, as required.

[0504] In another embodiment, high-level expression of a heterologous coding sequence, such as, for example, a polynucleotide of the present invention, may be achieved by cloning the heterologous polynucleotide of the invention into an expression vector such as, for example, pGAPZ or pGAPZalpha, and growing the yeast culture in the absence of methanol.

[0505] In addition to encompassing host cells containing the vector constructs discussed herein, the invention also encompasses primary, secondary, and immortalized host cells of vertebrate origin, particularly mammalian origin, that have been engineered to delete or replace endogenous genetic material (e.g., coding sequence), and/or to include genetic material (e.g., heterologous polynucleotide sequences) that is operably associated with the polynucleotides of the invention, and which activates, alters, and/or amplifies endogenous polynucleotides. For example, techniques known in the art may be used to operably associate heterologous control regions (e.g., promoter and/or enhancer) and endogenous polynucleotide sequences via homologous recombination, resulting in the formation of a new transcription unit (see, e.g., U.S. Pat. No. 5,641,670, issued Jun. 24, 1997; U.S. Pat. No. 5,733,761, issued Mar. 31, 1998; International Publication No. WO 96/29411, published Sep. 26, 1996; International Publication No. WO 94/12650, published Aug. 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA 86:8932-8935 (1989); and Zijlstra et al., Nature 342:435-438 (1989), the disclosures of each of which are incorporated by reference in their entireties).

[0506] In addition, polypeptides of the invention can be chemically synthesized using techniques known in the art (e.g., see Creighton, 1983, Proteins: Structures and Molecular Principles, W. H. Freeman & Co., N.Y., and Hunkapiller et al., Nature, 310:105-111 (1984)). For example, a polypeptide corresponding to a fragment of a polypeptide sequence of the invention can be synthesized by use of a peptide synthesizer. Furthermore, if desired, nonclassical amino acids or chemical amino acid analogs can be introduced as a substitution or addition into the polypeptide sequence. Non-classical amino acids include, but are not limited to, to the D-isomers of the common amino acids, 2,4-diaminobutyric acid, a-amino isobutyric acid, 4-aminobutyric acid, Abu, 2-amino butyric acid, g-Abu, e-Ahx, 6-amino hexanoic acid, Aib, 2-amino isobutyric acid, 3-amino propionic acid, omithine, norleucine, norvaline, hydroxyproline, sarcosine, citrulline, homocitrulline, cysteic acid, t-butylglycine, t-butylalanine, phenylglycine, cyclohexylalanine, b-alanine, fluoro-amino acids, designer amino acids such as b-methyl amino acids, Ca-methyl amino acids, Na-methyl amino acids, and amino acid analogs in general. Furthermore, the amino acid can be D (dextrorotary) or L (levorotary).

[0507] The invention encompasses polypeptides which are differentially modified during or after translation, e.g., by glycosylation, acetylation, phosphorylation, amidation, derivatization by known protecting/blocking groups, proteolytic cleavage, linkage to an antibody molecule or other cellular ligand, etc. Any of numerous chemical modifications may be carried out by known techniques, including but not limited, to specific chemical cleavage by cyanogen bromide, trypsin, chymotrypsin, papain, V8 protease, NaBH4; acetylation, formylation, oxidation, reduction; metabolic synthesis in the presence of tunicamycin; etc.

[0508] Additional post-translational modifications encompassed by the invention include, for example, e.g., N-linked or O-linked carbohydrate chains, processing of N-terminal or C-terminal ends), attachment of chemical moieties to the amino acid backbone, chemical modifications of N-linked or O-linked carbohydrate chains, and addition or deletion of an N-terminal methionine residue as a result of prokaryotic host cell expression. The polypeptides may also be modified with a detectable label, such as an enzymatic, fluorescent, isotopic or affinity label to allow for detection and isolation of the protein, the addition of epitope tagged peptide fragments (e.g., FLAG, HA, GST, thioredoxin, maltose binding protein, etc.), attachment of affinity tags such as biotin and/or streptavidin, the covalent attachment of chemical moieties to the amino acid backbone, N- or C-terminal processing of the polypeptides ends (e.g., proteolytic processing), deletion of the N-terminal methionine residue, etc.

[0509] Also provided by the invention are chemically modified derivatives of the polypeptides of the invention which may provide additional advantages such as increased solubility, stability and circulating time of the polypeptide, or decreased immunogenicity (see U.S. Pat. No. 4,179,337). The chemical moieties for derivitization may be selected from water soluble polymers such as polyethylene glycol, ethylene glycol/propylene glycol copolymers, carboxymethylcellulose, dextran, polyvinyl alcohol and the like. The polypeptides may be modified at random positions within the molecule, or at predetermined positions within the molecule and may include one, two, three or more attached chemical moieties.

[0510] The invention further encompasses chemical derivitization of the polypeptides of the present invention, preferably where the chemical is a hydrophilic polymer residue. Exemplary hydrophilic polymers, including derivatives, may be those that include polymers in which the repeating units contain one or more hydroxy groups (polyhydroxy polymers), including, for example, poly(vinyl alcohol); polymers in which the repeating units contain one or more amino groups (polyamine polymers), including, for example, peptides, polypeptides, proteins and lipoproteins, such as albumin and natural lipoproteins; polymers in which the repeating units contain one or more carboxy groups (polycarboxy polymers), including, for example, carboxymethylcellulose, alginic acid and salts thereof, such as sodium and calcium alginate, glycosaminoglycans and salts thereof, including salts of hyaluronic acid, phosphorylated and sul-

fonated derivatives of carbohydrates, genetic material, such as interleukin-2 and interferon, and phosphorothioate oligomers; and polymers in which the repeating units contain one or more saccharide moieties (polysaccharide polymers), including, for example, carbohydrates.

[0511] The molecular weight of the hydrophilic polymers may vary, and is generally about 50 to about 5,000,000, with polymers having a molecular weight of about 100 to about 50,000 being preferred. The polymers may be branched or unbranched. More preferred polymers have a molecular weight of about 150 to about 10,000, with molecular weights of 200 to about 8,000 being even more preferred.

[0512] For polyethylene glycol, the preferred molecular weight is between about 1 kDa and about 100 kDa (the term "about" indicating that in preparations of polyethylene glycol, some molecules will weigh more, some less, than the stated molecular weight) for ease in handling and manufacturing. Other sizes may be used, depending on the desired therapeutic profile (e.g., the duration of sustained release desired, the effects, if any on biological activity, the ease in handling, the degree or lack of antigenicity and other known effects of the polyethylene glycol to a therapeutic protein or analog).

[0513] Additional preferred polymers which may be used to derivatize polypeptides of the invention, include, for example, poly(ethylene glycol) (PEG), poly(vinylpyrrolidine), polyoxomers, polysorbate and poly(vinyl alcohol), with PEG polymers being particularly preferred. Preferred among the PEG polymers are PEG polymers having a molecular weight of from about 100 to about 10,000. More preferably, the PEG polymers have a molecular weight of from about 200 to about 8,000, with PEG 2,000, PEG 5,000 and PEG 8,000, which have molecular weights of 2,000, 5,000 and 8,000, respectively, being even more preferred. Other suitable hydrophilic polymers, in addition to those exemplified above, will be readily apparent to one skilled in the art based on the present disclosure. Generally, the polymers used may include polymers that can be attached to the polypeptides of the invention via alkylation or acylation reactions.

[0514] The polyethylene glycol molecules (or other chemical moieties) should be attached to the protein with consideration of effects on functional or antigenic domains of the protein. There are a number of attachment methods available to those skilled in the art, e.g., EP 0 401 384, herein incorporated by reference (coupling PEG to G-CSF), see also Malik et al., Exp. Hematol. 20:1028-1035 (1992) (reporting pegylation of GM-CSF using tresyl chloride). For example, polyethylene glycol may be covalently bound through amino acid residues via a reactive group, such as, a free amino or carboxyl group. Reactive groups are those to which an activated polyethylene glycol molecule may be bound. The amino acid residues having a free amino group may include lysine residues and the N-terminal amino acid residues; those having a free carboxyl group may include aspartic acid residues glutamic acid residues and the C-terminal amino acid residue. Sulfhydryl groups may also be used as a reactive group for attaching the polyethylene glycol molecules. Preferred for therapeutic purposes is attachment at an amino group, such as attachment at the N-terminus or lysine group.

[0515] One may specifically desire proteins chemically modified at the N-terminus. Using polyethylene glycol as an illustration of the present composition, one may select from a variety of polyethylene glycol molecules (by molecular weight, branching, etc.), the proportion of polyethylene glycol molecules to protein (polypeptide) molecules in the reaction mix, the type of pegylation reaction to be performed, and the method of obtaining the selected N-terminally pegylated protein. The method of obtaining the N-terminally pegylated preparation (i.e., separating this moiety from other monopegylated moieties if necessary) may be by purification of the N-terminally pegylated material from a population of pegylated protein molecules. Selective proteins chemically modified at the N-terminus modification may be accomplished by reductive alkylation which exploits differential reactivity of different types of primary amino groups (lysine versus the N-terminus) available for derivatization in a particular protein. Under the appropriate reaction conditions, substantially selective derivatization of the protein at the N-terminus with a carbonyl group containing polymer is achieved.

[0516] As with the various polymers exemplified above, it is contemplated that the polymeric residues may contain functional groups in addition, for example, to those typically involved in linking the polymeric residues to the polypeptides of the present invention. Such functionalities include, for example, carboxyl, amine, hydroxy and thiol groups. These functional groups on the polymeric residues can be further reacted, if desired, with materials that are generally reactive with such functional groups and which can assist in targeting specific tissues in the body including, for example, diseased tissue. Exemplary materials which can be reacted with the additional functional groups include, for example, proteins, including antibodies, carbohydrates, peptides, glycopeptides, glycolipids, lectins, and nucleosides.

[0517] In addition to residues of hydrophilic polymers, the chemical used to derivatize the polypeptides of the present invention can be a saccharide residue. Exemplary saccharides which can be derived include, for example, monosaccharides or sugar alcohols, such as erythrose, threose, ribose, arabinose, xylose, lyxose, fructose, sorbitol, mannitol and sedoheptulose, with preferred monosaccharides being fructose, mannose, xylose, arabinose, mannitol and sorbitol; and disaccharides, such as lactose, sucrose, maltose and cellobiose. Other saccharides include, for example, inositol and ganglioside head groups. Other suitable saccharides, in addition to those exemplified above, will be readily apparent to one skilled in the art based on the present disclosure. Generally, saccharides which may be used for derivitization include saccharides that can be attached to the polypeptides of the invention via alkylation or acylation reactions.

[0518] Moreover, the invention also encompasses derivitization of the polypeptides of the present invention, for example, with lipids (including cationic, anionic, polymerized, charged, synthetic, saturated, unsaturated, and any combination of the above, etc.) stabilizing agents.

[0519] The invention encompasses derivitization of the polypeptides of the present invention, for example, with compounds that may serve a stabilizing function (e.g., to increase the polypeptides half-life in solution, to make the polypeptides more water soluble, to increase the polypeptides hydrophilic or hydrophobic character, etc.). Polymers useful as stabilizing materials may be of natural, semi-synthetic (modified natural) or synthetic origin. Exemplary

natural polymers include naturally occurring polysaccharides, such as, for example, arabinans, fructans, fucans, galactans, galacturonans, glucans, mannans, xylans (such as, for example, inulin), levan, fucoidan, carrageenan, galatocarolose, pectic acid, pectins, including amylose, pullulan, glycogen, amylopectin, cellulose, dextran, dextrin, dextrose, glucose, polyglucose, polydextrose, pustulan, chitin, agarose, keratin, chondroitin, dermatan, hyaluronic acid, alginic acid, xanthin gum, starch and various other natural homopolymer or heteropolymers, such as those containing one or more of the following aldoses, ketoses, acids or amines: erythose, threose, ribose, arabinose, xylose, lyxose, allose, altrose, glucose, dextrose, mannose, gulose, idose, galactose, talose, erythrulose, ribulose, xylulose, psicose, fructose, sorbose, tagatose, mannitol, sorbitol, lactose, sucrose, trehalose, maltose, cellobiose, glycine, serine, threonine, cysteine, tyrosine, asparagine, glutamine, aspartic acid, glutamic acid, lysine, arginine, histidine, glucuronic acid, gluconic acid, glucaric acid, galacturonic acid, mannuronic acid, glucosamine, galactosamine, and neuraminic acid, and naturally occurring derivatives thereof Accordingly, suitable polymers include, for example, proteins, such as albumin, polyalginates, and polylactide-coglycolide polymers. Exemplary semi-synthetic polymers include carboxymethylcellulose, hydroxymethylcellulose, hydroxypropylmethylcellulose, methylcellulose, and methoxycellulose. Exemplary synthetic polymers include polyphosphazenes, hydroxyapatites, fluoroapatite polymers, polyethylenes (such as, for example, polyethylene glycol (including for example, the class of compounds referred to as Pluronics-.RTM., commercially available from BASF, Parsippany, N.J.), polyoxyethylene, and polyethylene terephthlate), polypropylenes (such as, for example, polypropylene glycol), polyurethanes (such as, for example, polyvinyl alcohol (PVA), polyvinyl chloride and polyvinylpyrrolidone), polyamides including nylon, polystyrene, polylactic acids, fluorinated hydrocarbon polymers, fluorinated carbon polymers (such as, for example, polytetrafluoroethylene), acrylate, methacrylate, and polymethylmethacrylate, and derivatives thereof. Methods for the preparation of derivatized polypeptides of the invention which employ polymers as stabilizing compounds will be readily apparent to one skilled in the art, in view of the present disclosure, when coupled with information known in the art, such as that described and referred to in Unger, U.S. Pat. No. 5,205,290, the disclosure of which is hereby incorporated by reference herein in its entirety.

[0520] Moreover, the invention encompasses additional modifications of the polypeptides of the present invention. Such additional modifications are known in the art, and are specifically provided, in addition to methods of derivitization, etc., in U.S. Pat. No. 6,028,066, which is hereby incorporated in its entirety herein.

[0521] The polypeptides of the invention may be in monomers or multimers (i.e., dimers, trimers, tetramers and higher multimers). Accordingly, the present invention relates to monomers and multimers of the polypeptides of the invention, their preparation, and compositions (preferably, Therapeutics) containing them. In specific embodiments, the polypeptides of the invention are monomers, dimers, trimers or tetramers. In additional embodiments, the multimers of the invention are at least dimers, at least trimers, or at least tetramers.

[0522] Multimers encompassed by the invention may be homomers or heteromers. As used herein, the term homomer, refers to a multimer containing only polypeptides corresponding to the amino acid sequence of SEQ ID NO:2 or encoded by the cDNA contained in a deposited clone (including fragments, variants, splice variants, and fusion proteins, corresponding to these polypeptides as described herein). These homomers may contain polypeptides having identical or different amino acid sequences. In a specific embodiment, a homomer of the invention is a multimer containing only polypeptides having an identical amino acid sequence. In another specific embodiment, a homomer of the invention is a multimer containing polypeptides having different amino acid sequences. In specific embodiments, the multimer of the invention is a homodimer (e.g., containing polypeptides having identical or different amino acid sequences) or a homotrimer (e.g., containing polypeptides having identical and/or different amino acid sequences). In additional embodiments, the homomeric multimer of the invention is at least a homodimer, at least a homotrimer, or at least a homotetramer.

[0523] As used herein, the term heteromer refers to a multimer containing one or more heterologous polypeptides (i.e., polypeptides of different proteins) in addition to the polypeptides of the invention. In a specific embodiment, the multimer of the invention is a heterodimer, a heterotrimer, or a heterotetramer. In additional embodiments, the heteromeric multimer of the invention is at least a heterodimer, at least a heterotrimer, or at least a heterotetramer.

[0524] Multimers of the invention may be the result of hydrophobic, hydrophilic, ionic and/or covalent associations and/or may be indirectly linked, by for example, liposome formation. Thus, in one embodiment, multimers of the invention, such as, for example, homodimers or homotrimers, are formed when polypeptides of the invention contact one another in solution. In another embodiment, heteromultimers of the invention, such as, for example, heterotrimers or heterotetramers, are formed when polypeptides of the invention contact antibodies to the polypeptides of the invention (including antibodies to the heterologous polypeptide sequence in a fusion protein of the invention) in solution. In other embodiments, multimers of the invention are formed by covalent associations with and/or between the polypeptides of the invention. Such covalent associations may involve one or more amino acid residues contained in the polypeptide sequence (e.g., that recited in the sequence listing, or contained in the polypeptide encoded by a deposited clone). In one instance, the covalent associations are cross-linking between cysteine residues located within the polypeptide sequences which interact in the native (i.e., naturally occurring) polypeptide. In another instance, the covalent associations are the consequence of chemical or recombinant manipulation. Alternatively, such covalent associations may involve one or more amino acid residues contained in the heterologous polypeptide sequence in a fusion protein of the invention.

[0525] In one example, covalent associations are between the heterologous sequence contained in a fusion protein of the invention (see, e.g., U.S. Pat. No. 5,478,925). In a specific example, the covalent associations are between the heterologous sequence contained in an Fe fusion protein of the invention (as described herein). In another specific example, covalent associations of fusion proteins of the

invention are between heterologous polypeptide sequence from another protein that is capable of forming covalently associated multimers, such as for example, osteoprotegerin (see, e.g., International Publication No: WO 98/49305, the contents of which are herein incorporated by reference in its entirety). In another embodiment, two or more polypeptides of the invention are joined through peptide linkers. Examples include those peptide linkers described in U.S. Pat. No. 5,073,627 (hereby incorporated by reference). Proteins comprising multiple polypeptides of the invention separated by peptide linkers may be produced using conventional recombinant DNA technology.

[0526] Another method for preparing multimer polypeptides of the invention involves use of polypeptides of the invention fused to a leucine zipper or isoleucine zipper polypeptide sequence. Leucine zipper and isoleucine zipper domains are polypeptides that promote multimerization of the proteins in which they are found. Leucine zippers were originally identified in several DNA-binding proteins (Landschulz et al., Science 240:1759, (1988)), and have since been found in a variety of different proteins. Among the known leucine zippers are naturally occurring peptides and derivatives thereof that dimerize or trimerize. Examples of leucine zipper domains suitable for producing soluble multimeric proteins of the invention are those described in PCT application WO 94/10308, hereby incorporated by reference. Recombinant fusion proteins comprising a polypeptide of the invention fused to a polypeptide sequence that dimerizes or trimerizes in solution are expressed in suitable host cells, and the resulting soluble multimeric fusion protein is recovered from the culture supernatant using techniques known in the art.

[0527] Trimeric polypeptides of the invention may offer the advantage of enhanced biological activity. Preferred leucine zipper moieties and isoleucine moieties are those that preferentially form trimers. One example is a leucine zipper derived from lung surfactant protein D (SPD), as described in Hoppe et al. (FEBS Letters 344:191, (1994)) and in U.S. patent application Ser. No. 08/446,922, hereby incorporated by reference. Other peptides derived from naturally occurring trimeric proteins may be employed in preparing trimeric polypeptides of the invention.

[0528] In another example, proteins of the invention are associated by interactions between Flag® polypeptide sequence contained in fusion proteins of the invention containing Flag® polypeptide sequence. In a further embodiment, associations proteins of the invention are associated by interactions between heterologous polypeptide sequence contained in Flag® fusion proteins of the invention and anti-Flag® antibody.

[0529] The multimers of the invention may be generated using chemical techniques known in the art. For example, polypeptides desired to be contained in the multimers of the invention may be chemically cross-linked using linker molecules and linker molecule length optimization techniques known in the art (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety). Additionally, multimers of the invention may be generated using techniques known in the art to form one or more intermolecule cross-links between the cysteine residues located within the sequence of the polypeptides desired to be contained in the multimer (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety). Further, polypeptides of the invention may be routinely modified by the addition of cysteine or biotin to the C terminus or N-terminus of the polypeptide and techniques known in the art may be applied to generate multimers containing one or more of these modified polypeptides (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety). Additionally, techniques known in the art may be applied to generate liposomes containing the polypeptide components desired to be contained in the multimer of the invention (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety).

[0530] Alternatively, multimers of the invention may be generated using genetic engineering techniques known in the art. In one embodiment, polypeptides contained in multimers of the invention are produced recombinantly using fusion protein technology described herein or otherwise known in the art (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety). In a specific embodiment, polynucleotides coding for a homodimer of the invention are generated by ligating a polynucleotide sequence encoding a polypeptide of the invention to a sequence encoding a linker polypeptide and then further to a synthetic polynucleotide encoding the translated product of the polypeptide in the reverse orientation from the original C-terminus to the N-terminus (lacking the leader sequence) (see, e.g., U.S. Pat. No. 5,478,925, which is herein incorporated by reference in its entirety). In another embodiment, recombinant techniques described herein or otherwise known in the art are applied to generate recombinant polypeptides of the invention which contain a transmembrane domain (or hydrophobic or signal peptide) and which can be incorporated by membrane reconstitution techniques into liposomes (see, e.g., U.S. Pat. No. 5,478, 925, which is herein incorporated by reference in its entirety).

[0531] In addition, the polynucleotide insert of the present invention could be operatively linked to "artificial" or chimeric promoters and transcription factors. Specifically, the artificial promoter could comprise, or alternatively consist, of any combination of cis-acting DNA sequence elements that are recognized by trans-acting transcription factors. Preferably, the cis acting DNA sequence elements and trans-acting transcription factors are operable in mammals. Further, the trans-acting transcription factors of such "artificial" promoters could also be "artificial" or chimeric in design themselves and could act as activators or repressors to said "artificial" promoter.

Uses of the Polynucleotides

[0532] Each of the polynucleotides identified herein can be used in numerous ways as reagents. The following description should be considered exemplary and utilizes known techniques.

[0533] The polynucleotides of the present invention are useful for chromosome identification. There exists an ongoing need to identify new chromosome markers, since few chromosome marking reagents, based on actual sequence data (repeat polymorphisms), are presently available. Each polynucleotide of the present invention can be used as a chromosome marker.

[0534] Briefly, sequences can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp) from the

sequences shown in SEQ ID NO:1. Primers can be selected using computer analysis so that primers do not span more than one predicted exon in the genomic DNA. These primers are then used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the SEQ ID NO:1 will yield an amplified fragment.

[0535] Similarly, somatic hybrids provide a rapid method of PCR mapping the polynucleotides to particular chromosomes. Three or more clones can be assigned per day using a single thermal cycler. Moreover, sublocalization of the polynucleotides can be achieved with panels of specific chromosome fragments. Other gene mapping strategies that can be used include in situ hybridization, prescreening with labeled flow-sorted chromosomes, and preselection by hybridization to construct chromosome specific-cDNA libraries.

[0536] Precise chromosomal location of the polynucleotides can also be achieved using fluorescence in situ hybridization (FISH) of a metaphase chromosomal spread. This technique uses polynucleotides as short as 500 or 600 bases; however, polynucleotides 2,000-4,000 bp are preferred. For a review of this technique, see Verma et al., "Human Chromosomes: a Manual of Basic Techniques," Pergamon Press, New York (1988).

[0537] For chromosome mapping, the polynucleotides can be used individually (to mark a single chromosome or a single site on that chromosome) or in panels (for marking multiple sites and/or multiple chromosomes). Preferred polynucleotides correspond to the noncoding regions of the cDNAs because the coding sequences are more likely conserved within gene families, thus increasing the chance of cross hybridization during chromosomal mapping.

[0538] Once a polynucleotide has been mapped to a precise chromosomal location, the physical position of the polynucleotide can be used in linkage analysis. Linkage analysis establishes coinheritance between a chromosomal location and presentation of a particular disease. Disease mapping data are known in the art. Assuming 1 megabase mapping resolution and one gene per 20 kb, a cDNA precisely localized to a chromosomal region associated with the disease could be one of 50-500 potential causative genes.

[0539] Thus, once coinheritance is established, differences in the polynucleotide and the corresponding gene between affected and unaffected organisms can be examined. First, visible structural alterations in the chromosomes, such as deletions or translocations, are examined in chromosome spreads or by PCR. If no structural alterations exist, the presence of point mutations are ascertained. Mutations observed in some or all affected organisms, but not in normal organisms, indicates that the mutation may cause the disease. However, complete sequencing of the polypeptide and the corresponding gene from several normal organisms is required to distinguish the mutation from a polymorphism. If a new polymorphism is identified, this polymorphic polypeptide can be used for further linkage analysis.

[0540] Furthermore, increased or decreased expression of the gene in affected organisms as compared to unaffected organisms can be assessed using polynucleotides of the present invention. Any of these alterations (altered expression, chromosomal rearrangement, or mutation) can be used as a diagnostic or prognostic marker.

[0541] Thus, the invention also provides a diagnostic method useful during diagnosis of a disorder, involving measuring the expression level of polynucleotides of the present invention in cells or body fluid from an organism and comparing the measured gene expression level with a standard level of polynucleotide expression level, whereby an increase or decrease in the gene expression level compared to the standard is indicative of a disorder.

[0542] By "measuring the expression level of a polynucleotide of the present invention" is intended qualitatively or quantitatively measuring or estimating the level of the polypeptide of the present invention or the level of the mRNA encoding the polypeptide in a first biological sample either directly (e.g., by determining or estimating absolute protein level or mRNA level) or relatively (e.g., by comparing to the polypeptide level or mRNA level in a second biological sample). Preferably, the polypeptide level or mRNA level in the first biological sample is measured or estimated and compared to a standard polypeptide level or mRNA level, the standard being taken from a second biological sample obtained from an individual not having the disorder or being determined by averaging levels from a population of organisms not having a disorder. As will be appreciated in the art, once a standard polypeptide level or mRNA level is known, it can be used repeatedly as a standard for comparison.

[0543] By "biological sample" is intended any biological sample obtained from an organism, body fluids, cell line, tissue culture, or other source which contains the polypeptide of the present invention or mRNA. As indicated, biological samples include body fluids (such as the following non-limiting examples, sputum, amniotic fluid, urine, saliva, breast milk, secretions, interstitial fluid, blood, serum, spinal fluid, etc.) which contain the polypeptide of the present invention, and other tissue sources found to express the polypeptide of the present invention. Methods for obtaining tissue biopsies and body fluids from organisms are well known in the art. Where the biological sample is to include mRNA, a tissue biopsy is the preferred source.

[0544] The method(s) provided above may Preferably be applied in a diagnostic method and/or kits in which polynucleotides and/or polypeptides are attached to a solid support. In one exemplary method, the support may be a "gene chip" or a "biological chip" as described in U.S. Pat. Nos. 5,837,832, 5,874,219, and 5,856,174. Further, such a gene chip with polynucleotides of the present invention attached may be used to identify polymorphisms between the polynucleotide sequences, with polynucleotides isolated from a test subject. The knowledge of such polymorphisms (i.e. their location, as well as, their existence) would be beneficial in identifying disease loci for many disorders, including proliferative diseases and conditions. Such a method is described in U.S. Pat. Nos. 5,858,659 and 5,856,104. The US Patents referenced supra are hereby incorporated by reference in their entirety herein.

[0545] The present invention encompasses polynucleotides of the present invention that are chemically synthesized, or reproduced as peptide nucleic acids (PNA), or according to other methods known in the art. The use of PNAs would serve as the preferred form if the polynucleotides are incorporated onto a solid support, or gene chip. For the purposes of the present invention, a peptide nucleic

acid (PNA) is a polyamide type of DNA analog and the monomeric units for adenine, guanine, thymine and cytosine are available commercially (Perceptive Biosystems). Certain components of DNA, such as phosphorus, phosphorus oxides, or deoxyribose derivatives, are not present in PNAs. As disclosed by P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, Science 254, 1497 (1991); and M. Egholm, O. Buchardt, L. Christensen, C. Behrens, S. M. Freier, D. A. Driver, R. H. Berg, S. K. Kim, B. Norden, and P. E. Nielsen, Nature 365, 666 (1993), PNAs bind specifically and tightly to complementary DNA strands and are not degraded by nucleases. In fact, PNA binds more strongly to DNA than DNA itself does. This is probably because there is no electrostatic repulsion between the two strands, and also the polyamide backbone is more flexible. Because of this, PNA/DNA duplexes bind under a wider range of stringency conditions than DNA/DNA duplexes, making it easier to perform multiplex hybridization. Smaller probes can be used than with DNA due to the stronger binding characteristics of PNA:DNA hybrids. In addition, it is more likely that single base mismatches can be determined with PNA/DNA hybridization because a single mismatch in a PNA/DNA 15-mer lowers the melting point (T.sub.m) by 8°-20° C., vs. 4°-16° C. for the DNA/DNA 15-mer duplex. Also, the absence of charge groups in PNA means that hybridization can be done at low ionic strengths and reduce possible interference by salt during the analysis.

[0546] In addition to the foregoing, a polynucleotide can be used to control gene expression through triple helix formation or antisense DNA or RNA. Antisense techniques are discussed, for example, in Okano, J. Neurochem. 56: 560 (1991); "Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, Fla. (1988). Triple helix formation is discussed in, for instance Lee et al., Nucleic Acids Research 6: 3073 (1979); Cooney et al., Science 241: 456 (1988); and Dervan et al., Science 251: 1360 (1991). Both methods rely on binding of the polynucleotide to a complementary DNA or RNA. For these techniques, preferred polynucleotides are usually oligonucleotides 20 to 40 bases in length and complementary to either the region of the gene involved in transcription (triple helix—see Lee et al., Nucl. Acids Res. 6:3073 (1979); Cooney et al., Science 241:456 (1988); and Dervan et al., Science 251:1360 (1991)) or to the mRNA itself (antisense—Okano, J. Neurochem. 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, Fla. (1988).) Triple helix formation optimally results in a shut-off of RNA transcription from DNA, while antisense RNA hybridization blocks translation of an mRNA molecule into polypeptide. Both techniques are effective in model systems, and the information disclosed herein can be used to design antisense or triple helix polynucleotides in an effort to treat or prevent disease.

[0547] The present invention encompasses the addition of a nuclear localization signal, operably linked to the 5' end, 3' end, or any location therein, to any of the oligonucleotides, antisense oligonucleotides, triple helix oligonucleotides, ribozymes, PNA oligonucleotides, and/or polynucleotides, of the present invention. See, for example, G. Cutrona, et al., Nat. Biotech., 18:300-303, (2000); which is hereby incorporated herein by reference.

[0548] Polynucleotides of the present invention are also useful in gene therapy. One goal of gene therapy is to insert a normal gene into an organism having a defective gene, in an effort to correct the genetic defect. The polynucleotides disclosed in the present invention offer a means of targeting such genetic defects in a highly accurate manner. Another goal is to insert a new gene that was not present in the host genome, thereby producing a new trait in the host cell. In one example, polynucleotide sequences of the present invention may be used to construct chimeric RNA/DNA oligonucleotides corresponding to said sequences, specifically designed to induce host cell mismatch repair mechanisms in an organism upon systemic injection, for example (Bartlett, R. J., et al., Nat. Biotech, 18:615-622 (2000), which is hereby incorporated by reference herein in its entirety). Such RNA/DNA oligonucleotides could be designed to correct genetic defects in certain host strains, and/or to introduce desired phenotypes in the host (e.g., introduction of a specific polymorphism within an endogenous gene corresponding to a polynucleotide of the present invention that may ameliorate and/or prevent a disease symptom and/or disorder, etc.). Alternatively, the polynucleotide sequence of the present invention may be used to construct duplex oligonucleotides corresponding to said sequence, specifically designed to correct genetic defects in certain host strains, and/or to introduce desired phenotypes into the host (e.g., introduction of a specific polymorphism within an endogenous gene corresponding to a polynucleotide of the present invention that may ameliorate and/or prevent a disease symptom and/or disorder, etc). Such methods of using duplex oligonucleotides are known in the art and are encompassed by the present invention (see EP1007712, which is hereby incorporated by reference herein in its entirety).

[0549] The polynucleotides are also useful for identifying organisms from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or more restriction enzymes, and probed on a Southern blot to yield unique bands for identifying personnel. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The polynucleotides of the present invention can be used as additional DNA markers for RFLP.

[0550] The polynucleotides of the present invention can also be used as an alternative to RFLP, by determining the actual base-by-base DNA sequence of selected portions of an organisms genome. These sequences can be used to prepare PCR primers for amplifying and isolating such selected DNA, which can then be sequenced. Using this technique, organisms can be identified because each organism will have a unique set of DNA sequences. Once an unique ID database is established for an organism, positive identification of that organism, living or dead, can be made from extremely small tissue samples. Similarly, polynucleotides of the present invention can be used as polymorphic markers, in addition to, the identification of transformed or non-transformed cells and/or tissues.

[0551] There is also a need for reagents capable of identifying the source of a particular tissue. Such need arises, for example, when presented with tissue of unknown origin. Appropriate reagents can comprise, for example, DNA probes or primers specific to particular tissue prepared from

the sequences of the present invention. Panels of such reagents can identify tissue by species and/or by organ type. In a similar fashion, these reagents can be used to screen tissue cultures for contamination. Moreover, as mentioned above, such reagents can be used to screen and/or identify transformed and non-transformed cells and/or tissues.

[0552] In the very least, the polynucleotides of the present invention can be used as molecular weight markers on Southern gels, as diagnostic probes for the presence of a specific mRNA in a particular cell type, as a probe to "subtract-out" known sequences in the process of discovering novel polynucleotides, for selecting and making oligomers for attachment to a "gene chip" or other support, to raise anti-DNA antibodies using DNA immunization techniques, and as an antigen to elicit an immune response.

Uses of the Polypeptides

[0553] Each of the polypeptides identified herein can be used in numerous ways. The following description should be considered exemplary and utilizes known techniques.

[0554] A polypeptide of the present invention can be used to assay protein levels in a biological sample using antibody-based techniques. For example, protein expression in tissues can be studied with classical immunohistological methods. (Jalkanen, M., et al., J. Cell. Biol. 101:976-985 (1985); Jalkanen, M., et al., J. Cell. Biol. 105:3087-3096 (1987).) Other antibody-based methods useful for detecting protein gene expression include immunoassays, such as the enzyme linked immunosorbent assay (ELISA) and the radioimmunoassay (RIA). Suitable antibody assay labels are known in the art and include enzyme labels, such as, glucose oxidase, and radioisotopes, such as iodine (1251, 1211), carbon (14C), sulfur (35S), tritium (3H), indium (112In), and technetium (99 mTc), and fluorescent labels, such as fluorescein and rhodamine, and biotin.

[0555] In addition to assaying protein levels in a biological sample, proteins can also be detected in vivo by imaging. Antibody labels or markers for in vivo imaging of protein include those detectable by X-radiography, NMR or ESR. For X-radiography, suitable labels include radioisotopes such as barium or cesium, which emit detectable radiation but are not overtly harmful to the subject. Suitable markers for NMR and ESR include those with a detectable characteristic spin, such as deuterium, which may be incorporated into the antibody by labeling of nutrients for the relevant hybridoma.

[0556] A protein-specific antibody or antibody fragment which has been labeled with an appropriate detectable imaging moiety, such as a radioisotope (for example, 131I, 112In, 99 mTc), a radio-opaque substance, or a material detectable by nuclear magnetic resonance, is introduced (for example, parenterally, subcutaneously, or intraperitoneally) into the mammal. It will be understood in the art that the size of the subject and the imaging system used will determine the quantity of imaging moiety needed to produce diagnostic images. In the case of a radioisotope moiety, for a human subject, the quantity of radioactivity injected will normally range from about 5 to 20 millicuries of 99mTc. The labeled antibody or antibody fragment will then preferentially accumulate at the location of cells which contain the specific protein. In vivo tumor imaging is described in S. W. Burchiel et al., "Immunopharmacokinetics of Radiolabeled Antibodies and Their Fragments." (Chapter 13 in Tumor Imaging: The Radiochemical Detection of Cancer, S. W. Burchiel and B. A. Rhodes, eds., Masson Publishing Inc. (1982).)

[0557] Thus, the invention provides a diagnostic method of a disorder, which involves (a) assaying the expression of a polypeptide of the present invention in cells or body fluid of an individual; (b) comparing the level of gene expression with a standard gene expression level, whereby an increase or decrease in the assayed polypeptide gene expression level compared to the standard expression level is indicative of a disorder. With respect to cancer, the presence of a relatively high amount of transcript in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

[0558] Moreover, polypeptides of the present invention can be used to treat, prevent, and/or diagnose disease. For example, patients can be administered a polypeptide of the present invention in an effort to replace absent or decreased levels of the polypeptide (e.g., insulin), to supplement absent or decreased levels of a different polypeptide (e.g., hemoglobin S for hemoglobin B, SOD, catalase, DNA repair proteins), to inhibit the activity of a polypeptide (e.g., an oncogene or tumor suppressor), to activate the activity of a polypeptide (e.g., by binding to a receptor), to reduce the activity of a membrane bound receptor by competing with it for free ligand (e.g., soluble TNF receptors used in reducing inflammation), or to bring about a desired response (e.g., blood vessel growth inhibition, enhancement of the immune response to proliferative cells or tissues).

[0559] Similarly, antibodies directed to a polypeptide of the present invention can also be used to treat, prevent, and/or diagnose disease. For example, administration of an antibody directed to a polypeptide of the present invention can bind and reduce overproduction of the polypeptide. Similarly, administration of an antibody can activate the polypeptide, such as by binding to a polypeptide bound to a membrane (receptor).

[0560] At the very least, the polypeptides of the present invention can be used as molecular weight markers on SDS-PAGE gels or on molecular sieve gel filtration columns using methods well known to those of skill in the art. Polypeptides can also be used to raise antibodies, which in turn are used to measure protein expression from a recombinant cell, as a way of assessing transformation of the host cell. Moreover, the polypeptides of the present invention can be used to test the following biological activities.

Gene Therapy Methods

[0561] Another aspect of the present invention is to gene therapy methods for treating or preventing disorders, diseases and conditions. The gene therapy methods relate to the introduction of nucleic acid (DNA, RNA and antisense DNA or RNA) sequences into an animal to achieve expression of a polypeptide of the present invention. This method requires a polynucleotide which codes for a polypeptide of the invention that operatively linked to a promoter and any other genetic elements necessary for the expression of the

polypeptide by the target tissue. Such gene therapy and delivery techniques are known in the art, see, for example, WO90/11092, which is herein incorporated by reference.

[0562] Thus, for example, cells from a patient may be engineered with a polynucleotide (DNA or RNA) comprising a promoter operably linked to a polynucleotide of the invention ex vivo, with the engineered cells then being provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, see Belldegrun et al., J. Natl. Cancer Inst., 85:207-216 (1993); Ferrantini et al., Cancer Research, 53:107-1112 (1993); Ferrantini et al., J. Immunology 153: 4604-4615 (1994); Kaido, T., et al., Int. J. Cancer 60: 221-229 (1995); Ogura et al., Cancer Research 50: 5102-5106 (1990); Santodonato, et al., Human Gene Therapy 7:1-10 (1996); Santodonato, et al., Gene Therapy 4:1246-1255 (1997); and Zhang, et al., Cancer Gene Therapy 3: 31-38 (1996)), which are herein incorporated by reference. In one embodiment, the cells which are engineered are arterial cells. The arterial cells may be reintroduced into the patient through direct injection to the artery, the tissues surrounding the artery, or through catheter injection.

[0563] As discussed in more detail below, the polynucleotide constructs can be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, and the like). The polynucleotide constructs may be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

[0564] In one embodiment, the polynucleotide of the invention is delivered as a naked polynucleotide. The term "naked" polynucleotide, DNA or RNA refers to sequences that are free from any delivery vehicle that acts to assist, promote or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the invention can also be delivered in liposome formulations and lipofectin formulations and the like can be prepared by methods well known to those skilled in the art. Such methods are described, for example, in U.S. Pat. Nos. 5,593,972, 5,589,466, and 5,580,859, which are herein incorporated by reference.

[0565] The polynucleotide vector constructs of the invention used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Appropriate vectors include pWLNEO, pSV2CAT, pOG44, pXT1 and pSG available from Stratagene; pSVK3, pBPV, pMSG and pSVL available from Pharmacia; and pEF1V5, pcDNA3.1, and pRc/CMV2 available from Invitrogen. Other suitable vectors will be readily apparent to the skilled artisan.

[0566] Any strong promoter known to those skilled in the art can be used for driving the expression of polynucleotide sequence of the invention. Suitable promoters include adenoviral promoters, such as the adenoviral major late promoter; or heterologous promoters, such as the cytomegalovirus (CMV) promoter; the respiratory syncytial virus (RSV) promoter; inducible promoters, such as the MMT promoter, the metallothionein promoter; heat shock promoters; the albumin promoter; the ApoAI promoter; human globin promoters; viral thymidine kinase promoters, such as the Herpes Simplex thymidine kinase promoter; retroviral LTRs; the b-actin promoter; and human growth hormone promoters. The promoter also may be the native promoter for the polynucleotides of the invention.

[0567] Unlike other gene therapy techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

[0568] The polynucleotide construct of the invention can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular, fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

[0569] For the naked nucleic acid sequence injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 mg/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration.

[0570] The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked DNA constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

[0571] The naked polynucleotides are delivered by any method known in the art, including, but not limited to, direct needle injection at the delivery site, intravenous injection, topical administration, catheter infusion, and so-called "gene guns". These delivery methods are known in the art.

[0572] The constructs may also be delivered with delivery vehicles such as viral sequences, viral particles, liposome

formulations, lipofectin, precipitating agents, etc. Such methods of delivery are known in the art.

[0573] In certain embodiments, the polynucleotide constructs of the invention are complexed in a liposome preparation. Liposomal preparations for use in the instant invention include cationic (positively charged), anionic (negatively charged) and neutral preparations. However, cationic liposomes are particularly preferred because a tight charge complex can be formed between the cationic liposome and the polyanionic nucleic acid. Cationic liposomes have been shown to mediate intracellular delivery of plasmid DNA (Feigner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7416 (1987), which is herein incorporated by reference); mRNA (Malone et al., Proc. Natl. Acad. Sci. USA, 86:6077-6081 (1989), which is herein incorporated by reference); and purified transcription factors (Debs et al., J. Biol. Chem., 265:10189-10192 (1990), which is herein incorporated by reference), in functional form.

[0574] Cationic liposomes are readily available. For example, N[1-2,3-dioleyloxy)propyl]-N,N,N-triethylammonium (DOTMA) liposomes are particularly useful and are available under the trademark Lipofectin, from GIBCO BRL, Grand Island, N.Y. (See, also, Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7416 (1987), which is herein incorporated by reference). Other commercially available liposomes include transfectace (DDAB/DOPE) and DOTAP/DOPE (Boehringer).

[0575] Other cationic liposomes can be prepared from readily available materials using techniques well known in the art. See, e.g. PCT Publication No: WO 90/11092 (which is herein incorporated by reference) for a description of the synthesis of DOTAP (1,2-bis(oleoyloxy)-3-(trimethylammonio)propane) liposomes. Preparation of DOTMA liposomes is explained in the literature, see, e.g., Felgner et al., Proc. Natl. Acad. Sci. USA, 84:7413-7417, which is herein incorporated by reference. Similar methods can be used to prepare liposomes from other cationic lipid materials.

[0576] Similarly, anionic and neutral liposomes are readily available, such as from Avanti Polar Lipids (Birmingham, Ala.), or can be easily prepared using readily available materials. Such materials include phosphatidyl, choline, cholesterol, phosphatidyl ethanolamine, dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), dioleoylphoshatidyl ethanolamine (DOPE), among others. These materials can also be mixed with the DOTMA and DOTAP starting materials in appropriate ratios. Methods for making liposomes using these materials are well known in the art.

[0577] For example, commercially dioleoylphosphatidyl choline (DOPC), dioleoylphosphatidyl glycerol (DOPG), and dioleoylphosphatidyl ethanolamine (DOPE) can be used in various combinations to make conventional liposomes, with or without the addition of cholesterol. Thus, for example, DOPG/DOPC vesicles can be prepared by drying 50 mg each of DOPG and DOPC under a stream of nitrogen gas into a sonication vial. The sample is placed under a vacuum pump overnight and is hydrated the following day with deionized water. The sample is then sonicated for 2 hours in a capped vial, using a Heat Systems model 351 sonicator equipped with an inverted cup (bath type) probe at the maximum setting while the bath is circulated at 15EC. Alternatively, negatively charged vesicles can be prepared without sonication to produce multilamellar vesicles or by extrusion through nucleopore membranes to produce unilamellar vesicles of discrete size. Other methods are known and available to those of skill in the art.

[0578] The liposomes can comprise multilamellar vesicles (MLVs), small unilamellar vesicles (SUVs), or large unilamellar vesicles (LUVs), with SUVs being preferred. The various liposome-nucleic acid complexes are prepared using methods well known in the art. See, e.g., Straubinger et al., Methods of Immunology, 101:512-527 (1983), which is herein incorporated by reference. For example, MLVs containing nucleic acid can be prepared by depositing a thin film of phospholipid on the walls of a glass tube and subsequently hydrating with a solution of the material to be encapsulated. SUVs are prepared by extended sonication of MLVs to produce a homogeneous population of unilamellar liposomes. The material to be entrapped is added to a suspension of preformed MLVs and then sonicated. When using liposomes containing cationic lipids, the dried lipid film is resuspended in an appropriate solution such as sterile water or an isotonic buffer solution such as 10 mM Tris/NaCl, sonicated, and then the preformed liposomes are mixed directly with the DNA. The liposome and DNA form a very stable complex due to binding of the positively charged liposomes to the cationic DNA. SUVs find use with small nucleic acid fragments. LUVs are prepared by a number of methods, well known in the art. Commonly used methods include Ca2+-EDTA chelation (Papahadjopoulos et al., Biochim. Biophys. Acta, 394:483 (1975); Wilson et al., Cell, 17:77 (1979)); ether injection (Deamer et al., Biochim. Biophys. Acta, 443:629 (1976); Ostro et al., Biochem. Biophys. Res. Commun., 76:836 (1977); Fraley et al., Proc. Natl. Acad. Sci. USA, 76:3348 (1979)); detergent dialysis (Enoch et al., Proc. Natl. Acad. Sci. USA, 76:145 (1979)); and reverse-phase evaporation (REV) (Fraley et al., J. Biol. Chem., 255:10431 (1980); Szoka et al., Proc. Natl. Acad. Sci. USA, 75:145 (1978); Schaefer-Ridder et al., Science, 215:166 (1982)), which are herein incorporated by reference.

[0579] Generally, the ratio of DNA to liposomes will be from about 10:1 to about 1:10. Preferably, the ration will be from about 5:1 to about 1:5. More preferably, the ration will be about 3:1 to about 1:3. Still more preferably, the ratio will be about 1:1.

[0580] U.S. Pat. No. 5,676,954 (which is herein incorporated by reference) reports on the injection of genetic material, complexed with cationic liposomes carriers, into mice. U.S. Pat. Nos. 4,897,355, 4,946,787, 5,049,386, 5,459,127, 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication No: WO 94/9469 (which are herein incorporated by reference) provide cationic lipids for use in transfecting DNA into cells and mammals. U.S. Pat. Nos. 5,589,466, 5,693,622, 5,580,859, 5,703,055, and international publication No: WO 94/9469 (which are herein incorporated by reference) provide methods for delivering DNA-cationic lipid complexes to mammals.

[0581] In certain embodiments, cells are engineered, ex vivo or in vivo, using a retroviral particle containing RNA which comprises a sequence encoding polypeptides of the invention. Retroviruses from which the retroviral plasmid vectors may be derived include, but are not limited to, Moloney Murine Leukemia Virus, spleen necrosis virus,

Rous sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, gibbon ape leukemia virus, human immunodeficiency virus, Myeloproliferative Sarcoma Virus, and mammary tumor virus.

[0582] The retroviral plasmid vector is employed to transduce packaging cell lines to form producer cell lines. Examples of packaging cells which may be transfected include, but are not limited to, the PE501, PA317, R-2, R-AM, PA12, T19-14, VT-19-17-H2, RCRE, RCRIP, GP+E-86, GP+envAm12, and DAN cell lines as described in Miller, Human Gene Therapy, 1:5-14 (1990), which is incorporated herein by reference in its entirety. The vector may transduce the packaging cells through any means known in the art. Such means include, but are not limited to, electroporation, the use of liposomes, and CaPO4 precipitation. In one alternative, the retroviral plasmid vector may be encapsulated into a liposome, or coupled to a lipid, and then administered to a host.

[0583] The producer cell line generates infectious retroviral vector particles which include polynucleotide encoding polypeptides of the invention. Such retroviral vector particles then may be employed, to transduce eukaryotic cells, either in vitro or in vivo. The transduced eukaryotic cells will express polypeptides of the invention.

[0584] In certain other embodiments, cells are engineered, ex vivo or in vivo, with polynucleotides of the invention contained in an adenovirus vector. Adenovirus can be manipulated such that it encodes and expresses polypeptides of the invention, and at the same time is inactivated in terms of its ability to replicate in a normal lytic viral life cycle. Adenovirus expression is achieved without integration of the viral DNA into the host cell chromosome, thereby alleviating concerns about insertional mutagenesis. Furthermore, adenoviruses have been used as live enteric vaccines for many years with an excellent safety profile (Schwartzet al., Am. Rev. Respir. Dis., 109:233-238 (1974)). Finally, adenovirus mediated gene transfer has been demonstrated in a number of instances including transfer of alpha-1-antitrypsin and CFTR to the lungs of cotton rats (Rosenfeld et al., Science, 252:431-434 (1991); Rosenfeld et al., Cell, 68:143-155 (1992)). Furthermore, extensive studies to attempt to establish adenovirus as a causative agent in human cancer were uniformly negative (Green et al. Proc. Natl. Acad. Sci. USA, 76:6606 (1979)).

[0585] Suitable adenoviral vectors useful in the present invention are described, for example, in Kozarsky and Wilson, Curr. Opin. Genet. Devel., 3:499-503 (1993); Rosenfeld et al., Cell, 68:143-155 (1992); Engelhardt et al., Human Genet. Ther., 4:759-769 (1993); Yang et al., Nature Genet., 7:362-369 (1994); Wilson et al., Nature, 365:691-692 (1993); and U.S. Pat. No. 5,652,224, which are herein incorporated by reference. For example, the adenovirus vector Ad2 is useful and can be grown in human 293 cells. These cells contain the El region of adenovirus and constitutively express E1a and E1b, which complement the defective adenoviruses by providing the products of the genes deleted from the vector. In addition to Ad2, other varieties of adenovirus (e.g., Ad3, Ad5, and Ad7) are also useful in the present invention.

[0586] Preferably, the adenoviruses used in the present invention are replication deficient. Replication deficient adenoviruses require the aid of a helper virus and/or packaging cell line to form infectious particles. The resulting virus is capable of infecting cells and can express a polynucleotide of interest which is operably linked to a promoter, but cannot replicate in most cells. Replication deficient adenoviruses may be deleted in one or more of all or a portion of the following genes: E1a, E1b, E3, E4, E2a, or L1 through L5.

[0587] In certain other embodiments, the cells are engineered, ex vivo or in vivo, using an adeno-associated virus (AAV). AAVs are naturally occurring defective viruses that require helper viruses to produce infectious particles (Muzyczka, Curr. Topics in Microbiol. Immunol., 158:97 (1992)). It is also one of the few viruses that may integrate its DNA into non-dividing cells. Vectors containing as little as 300 base pairs of AAV can be packaged and can integrate, but space for exogenous DNA is limited to about 4.5 kb. Methods for producing and using such AAVs are known in the art. See, for example, U.S. Pat. Nos. 5,139,941, 5,173,414, 5,354,678, 5,436,146, 5,474,935, 5,478,745, and 5,589,377.

[0588] For example, an appropriate AAV vector for use in the present invention will include all the sequences necessary for DNA replication, encapsidation, and host-cell integration. The polynucleotide construct containing polynucleotides of the invention is inserted into the AAV vector using standard cloning methods, such as those found in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press (1989). The recombinant AAV vector is then transfected into packaging cells which are infected with a helper virus, using any standard technique, including lipofection, electroporation, calcium phosphate precipitation, etc. Appropriate helper viruses include adenoviruses, cytomegaloviruses, vaccinia viruses, or herpes viruses. Once the packaging cells are transfected and infected, they will produce infectious AAV viral particles which contain the polynucleotide construct of the invention. These viral particles are then used to transduce eukaryotic cells, either ex vivo or in vivo. The transduced cells will contain the polynucleotide construct integrated into its genome, and will express the desired gene product.

[0589] Another method of gene therapy involves operably associating heterologous control regions and endogenous polynucleotide sequences (e.g. encoding the polypeptide sequence of interest) via homologous recombination (see, e.g., U.S. Pat. No. 5,641,670, issued Jun. 24, 1997; International Publication No: WO 96/29411, published Sep. 26, 1996; International Publication No: WO 94/12650, published Aug. 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA, 86:8932-8935 (1989); and Zijlstra et al., Nature, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not normally expressed in the cells, or is expressed at a lower level than desired.

[0590] Polynucleotide constructs are made, using standard techniques known in the art, which contain the promoter with targeting sequences flanking the promoter. Suitable promoters are described herein. The targeting sequence is sufficiently complementary to an endogenous sequence to permit homologous recombination of the promoter-targeting sequence with the endogenous sequence. The targeting sequence will be sufficiently near the 5' end of the desired

endogenous polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination.

[0591] The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter. The amplified promoter and targeting sequences are digested and ligated together.

[0592] The promoter-targeting sequence construct is delivered to the cells, either as naked polynucleotide, or in conjunction with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, whole viruses, lipofection, precipitating agents, etc., described in more detail above. The P promoter-targeting sequence can be delivered by any method, included direct needle injection, intravenous injection, topical administration, catheter infusion, particle accelerators, etc. The methods are described in more detail below.

[0593] The promoter-targeting sequence construct is taken up by cells. Homologous recombination between the construct and the endogenous sequence takes place, such that an endogenous sequence is placed under the control of the promoter. The promoter then drives the expression of the endogenous sequence.

[0594] The polynucleotides encoding polypeptides of the present invention may be administered along with other polynucleotides encoding angiogenic proteins. Angiogenic proteins include, but are not limited to, acidic and basic fibroblast growth factors, VEGF-1, VEGF-2 (VEGF-C), VEGF-3 (VEGF-B), epidermal growth factor alpha and beta, platelet-derived endothelial cell growth factor, platelet-derived growth factor, tumor necrosis factor alpha, hepatocyte growth factor, insulin like growth factor, colony stimulating factor, macrophage colony stimulating factor, granulocyte/macrophage colony stimulating factor, and nitric oxide synthase.

[0595] Preferably, the polynucleotide encoding a polypeptide of the invention contains a secretory signal sequence that facilitates secretion of the protein. Typically, the signal sequence is positioned in the coding region of the polynucleotide to be expressed towards or at the 5' end of the coding region. The signal sequence may be homologous or heterologous to the polynucleotide of interest and may be homologous or heterologous to the cells to be transfected. Additionally, the signal sequence may be chemically synthesized using methods known in the art.

[0596] Any mode of administration of any of the above-described polynucleotides constructs can be used so long as the mode results in the expression of one or more molecules in an amount sufficient to provide a therapeutic effect. This includes direct needle injection, systemic injection, catheter infusion, biolistic injectors, particle accelerators (i.e., "gene guns"), gelfoam sponge depots, other commercially available depot materials, osmotic pumps (e.g., Alza minipumps), oral or suppositorial solid (tablet or pill) pharmaceutical formulations, and decanting or topical applications during surgery. For example, direct injection of naked calcium phosphate-precipitated plasmid into rat liver and rat spleen or a protein-coated plasmid into the portal vein has resulted in gene expression of the foreign gene in the rat livers. (Kaneda et al., Science, 243:375 (1989)).

[0597] A preferred method of local administration is by direct injection. Preferably, a recombinant molecule of the present invention complexed with a delivery vehicle is administered by direct injection into or locally within the area of arteries. Administration of a composition locally within the area of arteries refers to injecting the composition centimeters and preferably, millimeters within arteries.

[0598] Another method of local administration is to contact a polynucleotide construct of the present invention in or around a surgical wound. For example, a patient can undergo surgery and the polynucleotide construct can be coated on the surface of tissue inside the wound or the construct can be injected into areas of tissue inside the wound.

[0599] Therapeutic compositions useful in systemic administration, include recombinant molecules of the present invention complexed to a targeted delivery vehicle of the present invention. Suitable delivery vehicles for use with systemic administration comprise liposomes comprising ligands for targeting the vehicle to a particular site.

[0600] Preferred methods of systemic administration, include intravenous injection, aerosol, oral and percutaneous (topical) delivery. Intravenous injections can be performed using methods standard in the art. Aerosol delivery can also be performed using methods standard in the art (see, for example, Stribling et al., Proc. Natl. Acad. Sci. USA, 189:11277-11281 (1992), which is incorporated herein by reference). Oral delivery can be performed by complexing a polynucleotide construct of the present invention to a carrier capable of withstanding degradation by digestive enzymes in the gut of an animal. Examples of such carriers, include plastic capsules or tablets, such as those known in the art. Topical delivery can be performed by mixing a polynucleotide construct of the present invention with a lipophilic reagent (e.g., DMSO) that is capable of passing into the skin.

[0601] Determining an effective amount of substance to be delivered can depend upon a number of factors including, for example, the chemical structure and biological activity of the substance, the age and weight of the animal, the precise condition requiring treatment and its severity, and the route of administration. The frequency of treatments depends upon a number of factors, such as the amount of polynucleotide constructs administered per dose, as well as the health and history of the subject. The precise amount, number of doses, and timing of doses will be determined by the attending physician or veterinarian. Therapeutic compositions of the present invention can be administered to any animal, preferably to mammals and birds. Preferred mammals include humans, dogs, cats, mice, rats, rabbits sheep, cattle, horses and pigs, with humans being particularly preferred.

Biological Activities

[0602] The polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used in assays to test for one or more biological activities. If these polynucleotides and polypeptides do exhibit activity in a particular assay, it is likely that these molecules may be involved in the

diseases associated with the biological activity. Thus, the polynucleotides or polypeptides, or agonists or antagonists could be used to treat the associated disease.

Immune Activity

[0603] The polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders, and/or conditions of the immune system, by activating or inhibiting the proliferation, differentiation, or mobilization (chemotaxis) of immune cells. Immune cells develop through a process called hematopoiesis, producing myeloid (platelets, red blood cells, neutrophils, and macrophages) and lymphoid (B and T lymphocytes) cells from pluripotent stem cells. The etiology of these immune diseases, disorders, and/or conditions may be genetic, somatic, such as cancer or some autoimmune diseases, disorders, and/or conditions, acquired (e.g., by chemotherapy or toxins), or infectious. Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention can be used as a marker or detector of a particular immune system disease or disorder.

[0604] A polynucleotides or polypeptides, or agonists or antagonists of the present invention may be useful in treating, preventing, and/or diagnosing diseases, disorders, and/or conditions of hematopoietic cells. A polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to increase differentiation and proliferation of hematopoietic cells, including the pluripotent stem cells, in an effort to treat or prevent those diseases, disorders, and/or conditions associated with a decrease in certain (or many) types hematopoietic cells. Examples of immunologic deficiency syndromes include, but are not limited to: blood protein diseases, disorders, and/or conditions (e.g. agammaglobulinemia, dysgammaglobulinemia), ataxia telangiectasia, common variable immunodeficiency, Digeorge Syndrome, HIV infection, HTLV-BLV infection, leukocyte adhesion deficiency syndrome, lymphopenia, phagocyte bactericidal dysfunction, severe combined immunodeficiency (SCIDs), Wiskott-Aldrich Disorder, anemia, thrombocytopenia, or hemoglobinuria.

[0605] Moreover, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could also be used to modulate hemostatic (the stopping of bleeding) or thrombolytic activity (clot formation). For example, by increasing hemostatic or thrombolytic activity, a polynucleotides or polypeptides, or agonists or antagonists of the present invention could be used to treat or prevent blood coagulation diseases, disorders, and/or conditions (e.g., afibrinogenemia, factor deficiencies), blood platelet diseases, disorders, and/or conditions (e.g. thrombocytopenia), or wounds resulting from trauma, surgery, or other causes. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention that can decrease hemostatic or thrombolytic activity could be used to inhibit or dissolve clotting. These molecules could be important in the treatment or prevention of heart attacks (infarction), strokes, or scarring.

[0606] A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be useful in treating, preventing, and/or diagnosing autoimmune diseases, disorders, and/or conditions. Many autoimmune diseases, disorders, and/or conditions result from inappropriate recognition of self as foreign material by immune cells. This inappropriate recognition results in an immune response leading to the destruction of the host tissue. Therefore, the administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing autoimmune diseases, disorders, and/or conditions.

[0607] Examples of autoimmune diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed or detected by the present invention include, but are not limited to: Addison's Disease, hemolytic anemia, antiphospholipid syndrome, rheumatoid arthritis, dermatitis, allergic encephalomyelitis, glomerulonephritis, Goodpasture's Syndrome, Graves' Disease, Multiple Sclerosis, Myasthenia Gravis, Neuritis, Ophthalmia, Bullous Pemphigoid, Pemphigus, Polyendocrinopathies, Purpura, Reiter's Disease, Stiff-Man Syndrome, Autoimmune Thyroiditis, Systemic Lupus Erythematosus, Autoimmune Pulmonary Inflammation, Guillain-Barre Syndrome, insulin dependent diabetes mellitis, and autoimmune inflammatory eye disease.

[0608] Similarly, allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems, may also be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or agonists or antagonists of the present invention. Moreover, these molecules can be used to treat anaphylaxis, hypersensitivity to an antigenic molecule, or blood group incompatibility.

[0609] A polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to treat, prevent, and/or diagnose organ rejection or graft-versus-host disease (GVHD). Organ rejection occurs by host immune cell destruction of the transplanted tissue through an immune response. Similarly, an immune response is also involved in GVHD, but, in this case, the foreign transplanted immune cells destroy the host tissues. The administration of a polynucleotides or polypeptides, or agonists or antagonists of the present invention that inhibits an immune response, particularly the proliferation, differentiation, or chemotaxis of T-cells, may be an effective therapy in preventing organ rejection or GVHD.

[0610] Similarly, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may also be used to modulate inflammation. For example, the polypeptide or polynucleotide or agonists or antagonist may inhibit the proliferation and differentiation of cells involved in an inflammatory response. These molecules can be used to treat, prevent, and/or diagnose inflammatory conditions, both chronic and acute conditions, including chronic prostatitis, granulomatous prostatitis and malacoplakia, inflammation associated with infection (e.g., septic shock, sepsis, or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine induced lung injury, inflammatory bowel disease, Crohn's disease, or resulting from over production of cytokines (e.g., TNF or IL-1.)

Hyperproliferative Disorders

[0611] Polynucleotides or polypeptides, or agonists or antagonists of the invention can be used to treat, prevent, and/or diagnose hyperproliferative diseases, disorders, and/or conditions, including neoplasms. A polynucleotides or polypeptides, or agonists or antagonists of the present invention may inhibit the proliferation of the disorder through direct or indirect interactions. Alternatively, a polynucleotides or polypeptides, or agonists or antagonists of the present invention may proliferate other cells which can inhibit the hyperproliferative disorder.

[0612] For example, by increasing an immune response, particularly increasing antigenic qualities of the hyperproliferative disorder or by proliferating, differentiating, or mobilizing T-cells, hyperproliferative diseases, disorders, and/or conditions can be treated, prevented, and/or diagnosed. This immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, decreasing an immune response may also be a method of treating, preventing, and/or diagnosing hyperproliferative diseases, disorders, and/or conditions, such as a chemotherapeutic agent.

[0613] Examples of hyperproliferative diseases, disorders, and/or conditions that can be treated, prevented, and/or diagnosed by polynucleotides or polypeptides, or agonists or antagonists of the present invention include, but are not limited to neoplasms located in the: colon, abdomen, bone, breast, digestive system, liver, pancreas, peritoneum, endocrine glands (adrenal, parathyroid, pituitary, testicles, ovary, thymus, thyroid), eye, head and neck, nervous (central and peripheral), lymphatic system, pelvic, skin, soft tissue, spleen, thoracic, and urogenital.

[0614] Similarly, other hyperproliferative diseases, disorders, and/or conditions can also be treated, prevented, and/or diagnosed by a polynucleotides or polypeptides, or agonists or antagonists of the present invention. Examples of such hyperproliferative diseases, disorders, and/or conditions include, but are not limited to: hypergammaglobulinemia, lymphoproliferative diseases, disorders, and/or conditions, paraproteinemias, purpura, sarcoidosis, Sezary Syndrome, Waldenstron's Macroglobulinemia, Gaucher's Disease, histiocytosis, and any other hyperproliferative disease, besides neoplasia, located in an organ system listed above.

[0615] One preferred embodiment utilizes polynucleotides of the present invention to inhibit aberrant cellular division, by gene therapy using the present invention, and/or protein fusions or fragments thereof.

[0616] Thus, the present invention provides a method for treating or preventing cell proliferative diseases, disorders, and/or conditions by inserting into an abnormally proliferating cell a polynucleotide of the present invention, wherein said polynucleotide represses said expression.

[0617] Another embodiment of the present invention provides a method of treating or preventing cell-proliferative diseases, disorders, and/or conditions in individuals comprising administration of one or more active gene copies of the present invention to an abnormally proliferating cell or cells. In a preferred embodiment, polynucleotides of the present invention is a DNA construct comprising a recombinant expression vector effective in expressing a DNA sequence encoding said polynucleotides. In another preferred embodiment of the present invention, the DNA construct encoding the polynucleotides of the present invention is inserted into cells to be treated utilizing a retrovirus, or more Preferably an adenoviral vector (See G J. Nabel, et. al., PNAS 1999 96: 324-326, which is hereby incorporated by reference). In a most preferred embodiment, the viral vector is defective and will not transform non-proliferating cells, only proliferating cells. Moreover, in a preferred embodiment, the polynucleotides of the present invention inserted into proliferating cells either alone, or in combination with or fused to other polynucleotides, can then be modulated via an external stimulus (i.e. magnetic, specific small molecule, chemical, or drug administration, etc.), which acts upon the promoter upstream of said polynucleotides to induce expression of the encoded protein product. As such the beneficial therapeutic affect of the present invention may be expressly modulated (i.e. to increase, decrease, or inhibit expression of the present invention) based upon said external stimulus.

[0618] Polynucleotides of the present invention may be useful in repressing expression of oncogenic genes or antigens. By "repressing expression of the oncogenic genes" is intended the suppression of the transcription of the gene, the degradation of the gene transcript (pre-message RNA), the inhibition of splicing, the destruction of the messenger RNA, the prevention of the post-translational modifications of the protein, the destruction of the protein, or the inhibition of the normal function of the protein.

[0619] For local administration to abnormally proliferating cells, polynucleotides of the present invention may be administered by any method known to those of skill in the art including, but not limited to transfection, electroporation, microinjection of cells, or in vehicles such as liposomes, lipofectin, or as naked polynucleotides, or any other method described throughout the specification. The polynucleotide of the present invention may be delivered by known gene delivery systems such as, but not limited to, retroviral vectors (Gilboa, J. Virology 44:845 (1982); Hocke, Nature 320:275 (1986); Wilson, et al., Proc. Natl. Acad. Sci. U.S.A. 85:3014), vaccinia virus system (Chakrabarty et al., Mol. Cell Biol. 5:3403 (1985) or other efficient DNA delivery systems (Yates et al., Nature 313:812 (1985)) known to those skilled in the art. These references are exemplary only and are hereby incorporated by reference. In order to specifically deliver or transfect cells which are abnormally proliferating and spare non-dividing cells, it is preferable to utilize a retrovirus, or adenoviral (as described in the art and elsewhere herein) delivery system known to those of skill in the art. Since host DNA replication is required for retroviral DNA to integrate and the retrovirus will be unable to self replicate due to the lack of the retrovirus genes needed for its life cycle. Utilizing such a retroviral delivery system for polynucleotides of the present invention will target said gene and constructs to abnormally proliferating cells and will spare the non-dividing normal cells.

[0620] The polynucleotides of the present invention may be delivered directly to cell proliferative disorder/disease sites in internal organs, body cavities and the like by use of imaging devices used to guide an injecting needle directly to the disease site. The polynucleotides of the present invention may also be administered to disease sites at the time of surgical intervention.

[0621] By "cell proliferative disease" is meant any human or animal disease or disorder, affecting any one or any

combination of organs, cavities, or body parts, which is characterized by single or multiple local abnormal proliferations of cells, groups of cells, or tissues, whether benign or malignant.

[0622] Any amount of the polynucleotides of the present invention may be administered as long as it has a biologically inhibiting effect on the proliferation of the treated cells. Moreover, it is possible to administer more than one of the polynucleotide of the present invention simultaneously to the same site. By "biologically inhibiting" is meant partial or total growth inhibition as well as decreases in the rate of proliferation or growth of the cells. The biologically inhibitory dose may be determined by assessing the effects of the polynucleotides of the present invention on target malignant or abnormally proliferating cell growth in tissue culture, tumor growth in animals and cell cultures, or any other method known to one of ordinary skill in the art.

[0623] The present invention is further directed to antibody-based therapies which involve administering of anti-polypeptides and anti-polynucleotide antibodies to a mammalian, preferably human, patient for treating, preventing, and/or diagnosing one or more of the described diseases, disorders, and/or conditions. Methods for producing anti-polypeptides and anti-polynucleotide antibodies polyclonal and monoclonal antibodies are described in detail elsewhere herein. Such antibodies may be provided in pharmaceutically acceptable compositions as known in the art or as described herein.

[0624] A summary of the ways in which the antibodies of the present invention may be used therapeutically includes binding polynucleotides or polypeptides of the present invention locally or systemically in the body or by direct cytotoxicity of the antibody, e.g. as mediated by complement (CDC) or by effector cells (ADCC). Some of these approaches are described in more detail below. Armed with the teachings provided herein, one of ordinary skill in the art will know how to use the antibodies of the present invention for diagnostic, monitoring or therapeutic purposes without undue experimentation.

[0625] In particular, the antibodies, fragments and derivatives of the present invention are useful for treating, preventing, and/or diagnosing a subject having or developing cell proliferative and/or differentiation diseases, disorders, and/or conditions as described herein. Such treatment comprises administering a single or multiple doses of the antibody, or a fragment, derivative, or a conjugate thereof.

[0626] The antibodies of this invention may be advantageously utilized in combination with other monoclonal or chimeric antibodies, or with lymphokines or hematopoietic growth factors, for example, which serve to increase the number or activity of effector cells which interact with the antibodies.

[0627] It is preferred to use high affinity and/or potent in vivo inhibiting and/or neutralizing antibodies against polypeptides or polynucleotides of the present invention, fragments or regions thereof, for both immunoassays directed to and therapy of diseases, disorders, and/or conditions related to polynucleotides or polypeptides, including fragments thereof, of the present invention. Such antibodies, fragments, or regions, will preferably have an affinity for polynucleotides or polypeptides, including fragments thereof. Preferred binding affinities include those with a dissociation constant or Kd less than 5×10−6M, 10−6M, 5×10 −7M, 10 −7M, 5×10−8M, 10−8M, 5×10−9M, 10−9M, 5×10−10M, 10−10M, 5×1−11M, 10−11M, 5×10−12M, 10−12M, 5×10−13M, 10−13M, 5×10−14M, 10−14M, 5×10−15M, and 10−15M.

[0628] Moreover, polypeptides of the present invention may be useful in inhibiting the angiogenesis of proliferative cells or tissues, either alone, as a protein fusion, or in combination with other polypeptides directly or indirectly, as described elsewhere herein. In a most preferred embodiment, said anti-angiogenesis effect may be achieved indirectly, for example, through the inhibition of hematopoietic, tumor-specific cells, such as tumor-associated macrophages (See Joseph E B, et al. J Natl Cancer Inst, 90(21):1648-53 (1998), which is hereby incorporated by reference). Antibodies directed to polypeptides or polynucleotides of the present invention may also result in inhibition of angiogenesis directly, or indirectly (See Witte L, et al., Cancer Metastasis Rev. 17(2):155-61 (1998), which is hereby incorporated by reference)).

[0629] Polypeptides, including protein fusions, of the present invention, or fragments thereof may be useful in inhibiting proliferative cells or tissues through the induction of apoptosis. Said polypeptides may act either directly, or indirectly to induce apoptosis of proliferative cells and tissues, for example in the activation of a death-domain receptor, such as tumor necrosis factor (TNF) receptor-1, CD95 (Fas/APO-1), TNF-receptor-related apoptosis-mediated protein (TRAMP) and TNF-related apoptosis-inducing ligand (TRAIL) receptor-1 and -2 (See Schulze-Osthoff K, et al., Eur J Biochem 254(3):439-59 (1998), which is hereby incorporated by reference). Moreover, in another preferred embodiment of the present invention, said polypeptides may induce apoptosis through other mechanisms, such as in the activation of other proteins which will activate apoptosis, or through stimulating the expression of said proteins, either alone or in combination with small molecule drugs or adjuvants, such as apoptonin, galectins, thioredoxins, anti-inflammatory proteins (See for example, Mutat. Res. 400(1-2):447-55 (1998), Med Hypotheses.50(5):423-33 (1998), Chem. Biol. Interact. Apr 24;111-112:23-34 (1998), J Mol Med.76(6):402-12 (1998), Int. J. Tissue React. 20(1):3-15 (1998), which are all hereby incorporated by reference).

[0630] Polypeptides, including protein fusions to, or fragments thereof, of the present invention are useful in inhibiting the metastasis of proliferative cells or tissues. Inhibition may occur as a direct result of administering polypeptides, or antibodies directed to said polypeptides as described elsewhere herein, or indirectly, such as activating the expression of proteins known to inhibit metastasis, for example alpha 4 integrins, (See, e.g., Curr Top Microbiol Immunol 1998;231:125-41, which is hereby incorporated by reference). Such therapeutic affects of the present invention may be achieved either alone, or in combination with small molecule drugs or adjuvants.

[0631] In another embodiment, the invention provides a method of delivering compositions containing the polypeptides of the invention (e.g., compositions containing polypeptides or polypeptide antibodies associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs) to targeted cells expressing the polypeptide of

the present invention. Polypeptides or polypeptide antibodies of the invention may be associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic and/or covalent interactions.

[0632] Polypeptides, protein fusions to, or fragments thereof, of the present invention are useful in enhancing the immunogenicity and/or antigenicity of proliferating cells or tissues, either directly, such as would occur if the polypeptides of the present invention 'vaccinated' the immune response to respond to proliferative antigens and immunogens, or indirectly, such as in activating the expression of proteins known to enhance the immune response (e.g. chemokines), to said antigens and immunogens.

Anti-Angiogenesis Activity

[0633] The naturally occurring balance between endogenous stimulators and inhibitors of angiogenesis is one in which inhibitory influences predominate. Rastinejad et al., Cell 56:345-355 (1989). In those rare instances in which neovascularization occurs under normal physiological conditions, such as wound healing, organ regeneration, embryonic development, and female reproductive processes, angiogenesis is stringently regulated and spatially and temporally delimited. Under conditions of pathological angiogenesis such as that characterizing solid tumor growth, these regulatory controls fail. Unregulated angiogenesis becomes pathologic and sustains progression of many neoplastic and non-neoplastic diseases. A number of serious diseases are dominated by abnormal neovascularization including solid tumor growth and metastases, arthritis, some types of eye diseases, disorders, and/or conditions, and psoriasis. See, e.g., reviews by Moses et al., Biotech. 9:630-634 (1991); Folkman et al., N. Engl. J. Med., 333:1757-1763 (1995); Auerbach et al., J. Microvasc. Res. 29:401-411 (1985); Folkman, Advances in Cancer Research, eds. Klein and Weinhouse, Academic Press, New York, pp. 175-203 (1985); Patz, Am. J. Opthalmol. 94:715-743 (1982); and Folkman et al., Science 221:719-725 (1983). In a number of pathological conditions, the process of angiogenesis contributes to the disease state. For example, significant data have accumulated which suggest that the growth of solid tumors is dependent on angiogenesis. Folkman and Klagsbrun, Science 235:442-447 (1987).

[0634] The present invention provides for treatment of diseases, disorders, and/or conditions associated with neovascularization by administration of the polynucleotides and/or polypeptides of the invention, as well as agonists or antagonists of the present invention. Malignant and metastatic conditions which can be treated with the polynucleotides and polypeptides, or agonists or antagonists of the invention include, but are not limited to, malignancies, solid tumors, and cancers described herein and otherwise known in the art (for a review of such disorders, see Fishman et al., Medicine, 2d Ed., J. B. Lippincott Co., Philadelphia (1985)).Thus, the present invention provides a method of treating, preventing, and/or diagnosing an angiogenesis-related disease and/or disorder, comprising administering to an individual in need thereof a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist of the invention. For example, polynucleotides, polypeptides, antagonists and/or agonists may be utilized in a variety of additional methods in order to therapeutically treat or prevent a cancer or tumor. Cancers which may be treated, prevented, and/or diagnosed with polynucleotides, polypeptides, antagonists and/or agonists include, but are not limited to solid tumors, including prostate, lung, breast, ovarian, stomach, pancreas, larynx, esophagus, testes, liver, parotid, biliary tract, colon, rectum, cervix, uterus, endometrium, kidney, bladder, thyroid cancer; primary tumors and metastases; melanomas; glioblastoma; Kaposi's sarcoma; leiomyosarcoma; non-small cell lung cancer; colorectal cancer; advanced malignancies; and blood born tumors such as leukemias. For example, polynucleotides, polypeptides, antagonists and/or agonists may be delivered topically, in order to treat or prevent cancers such as skin cancer, head and neck tumors, breast tumors, and Kaposi's sarcoma.

[0635] Within yet other aspects, polynucleotides, polypeptides, antagonists and/or agonists may be utilized to treat superficial forms of bladder cancer by, for example, intravesical administration. Polynucleotides, polypeptides, antagonists and/or agonists may be delivered directly into the tumor, or near the tumor site, via injection or a catheter. Of course, as the artisan of ordinary skill will appreciate, the appropriate mode of administration will vary according to the cancer to be treated. Other modes of delivery are discussed herein.

[0636] Polynucleotides, polypeptides, antagonists and/or agonists may be useful in treating, preventing, and/or diagnosing other diseases, disorders, and/or conditions, besides cancers, which involve angiogenesis. These diseases, disorders, and/or conditions include, but are not limited to: benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas; artheroscleric plaques; ocular angiogenic diseases, for example, diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, uvietis and Pterygia (abnormal blood vessel growth) of the eye; rheumatoid arthritis; psoriasis; delayed wound healing; endometriosis; vasculogenesis; granulations; hypertrophic scars (keloids); nonunion fractures; scleroderma; trachoma; vascular adhesions; myocardial angiogenesis; coronary collaterals; cerebral collaterals; arterioyenous malformations; ischemic limb angiogenesis; Osler-Webber Syndrome; plaque neovascularization; telangiectasia; hemophiliac joints; angiofibroma; fibromuscular dysplasia; wound granulation; Crohn's disease; and atherosclerosis.

[0637] For example, within one aspect of the present invention methods are provided for treating, preventing, and/or diagnosing hypertrophic scars and keloids, comprising the step of administering a polynucleotide, polypeptide, antagonist and/or agonist of the invention to a hypertrophic scar or keloid.

[0638] Within one embodiment of the present invention polynucleotides, polypeptides, antagonists and/or agonists are directly injected into a hypertrophic scar or keloid, in order to prevent the progression of these lesions. This therapy is of particular value in the prophylactic treatment of conditions which are known to result in the development of hypertrophic scars and keloids (e.g., burns), and is preferably initiated after the proliferative phase has had time to progress (approximately 14 days after the initial injury), but before hypertrophic scar or keloid development. As noted

above, the present invention also provides methods for treating, preventing, and/or diagnosing neovascular diseases of the eye, including for example, corneal neovascularization, neovascular glaucoma, proliferative diabetic retinopathy, retrolental fibroplasia and macular degeneration.

[0639] Moreover, Ocular diseases, disorders, and/or conditions associated with neovascularization which can be treated, prevented, and/or diagnosed with the polynucleotides and polypeptides of the present invention (including agonists and/or antagonists) include, but are not limited to: neovascular glaucoma, diabetic retinopathy, retinoblastoma, retrolental fibroplasia, uveitis, retinopathy of prematurity macular degeneration, corneal graft neovascularization, as well as other eye inflammatory diseases, ocular tumors and diseases associated with choroidal or iris neovascularization. See, e.g., reviews by Waltman et al., Am. J. Ophthal. 85:704-710 (1978) and Gartner et al., Surv. Ophthal. 22:291-312 (1978).

[0640] Thus, within one aspect of the present invention methods are provided for treating or preventing neovascular diseases of the eye such as corneal neovascularization (including corneal graft neovascularization), comprising the step of administering to a patient a therapeutically effective amount of a compound (as described above) to the cornea, such that the formation of blood vessels is inhibited. Briefly, the cornea is a tissue which normally lacks blood vessels. In certain pathological conditions however, capillaries may extend into the cornea from the pericorneal vascular plexus of the limbus. When the cornea becomes vascularized, it also becomes clouded, resulting in a decline in the patient's visual acuity. Visual loss may become complete if the cornea completely opacitates. A wide variety of diseases, disorders, and/or conditions can result in corneal neovascularization, including for example, corneal infections (e.g., trachoma, herpes simplex keratitis, leishmaniasis and onchocerciasis), immunological processes (e.g., graft rejection and Stevens-Johnson's syndrome), alkali burns, trauma, inflammation (of any cause), toxic and nutritional deficiency states, and as a complication of wearing contact lenses.

[0641] Within particularly preferred embodiments of the invention, may be prepared for topical administration in saline (combined with any of the preservatives and antimicrobial agents commonly used in ocular preparations), and administered in eyedrop form. The solution or suspension may be prepared in its pure form and administered several times daily. Alternatively, anti-angiogenic compositions, prepared as described above, may also be administered directly to the cornea. Within preferred embodiments, the anti-angiogenic composition is prepared with a muco-adhesive polymer which binds to cornea. Within further embodiments, the anti-angiogenic factors or anti-angiogenic compositions may be utilized as an adjunct to conventional steroid therapy. Topical therapy may also be useful prophylactically in corneal lesions which are known to have a high probability of inducing an angiogenic response (such as chemical burns). In these instances the treatment, likely in combination with steroids, may be instituted immediately to help prevent subsequent complications.

[0642] Within other embodiments, the compounds described above may be injected directly into the corneal stroma by an ophthalmologist under microscopic guidance. The preferred site of injection may vary with the morphology of the individual lesion, but the goal of the administration would be to place the composition at the advancing front of the vasculature (i.e., interspersed between the blood vessels and the normal cornea). In most cases this would involve perilimbic corneal injection to "protect" the cornea from the advancing blood vessels. This method may also be utilized shortly after a corneal insult in order to prophylactically prevent corneal neovascularization. In this situation the material could be injected in the perilimbic cornea interspersed between the corneal lesion and its undesired potential limbic blood supply. Such methods may also be utilized in a similar fashion to prevent capillary invasion of transplanted corneas. In a sustained-release form injections might only be required 2-3 times per year. A steroid could also be added to the injection solution to reduce inflammation resulting from the injection itself.

[0643] Within another aspect of the present invention, methods are provided for treating or preventing neovascular glaucoma, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eye, such that the formation of blood vessels is inhibited. In one embodiment, the compound may be administered topically to the eye in order to treat or prevent early forms of neovascular glaucoma. Within other embodiments, the compound may be implanted by injection into the region of the anterior chamber angle. Within other embodiments, the compound may also be placed in any location such that the compound is continuously released into the aqueous humor. Within another aspect of the present invention, methods are provided for treating or preventing proliferative diabetic retinopathy, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eyes, such that the formation of blood vessels is inhibited.

[0644] Within particularly preferred embodiments of the invention, proliferative diabetic retinopathy may be treated by injection into the aqueous humor or the vitreous, in order to increase the local concentration of the polynucleotide, polypeptide, antagonist and/or agonist in the retina. Preferably, this treatment should be initiated prior to the acquisition of severe disease requiring photocoagulation.

[0645] Within another aspect of the present invention, methods are provided for treating or preventing retrolental fibroplasia, comprising the step of administering to a patient a therapeutically effective amount of a polynucleotide, polypeptide, antagonist and/or agonist to the eye, such that the formation of blood vessels is inhibited. The compound may be administered topically, via intravitreous injection and/or via intraocular implants.

[0646] Additionally, diseases, disorders, and/or conditions which can be treated, prevented, and/or diagnosed with the polynucleotides, polypeptides, agonists and/or agonists include, but are not limited to, hemangioma, arthritis, psoriasis, angiofibroma, atherosclerotic plaques, delayed wound healing, granulations, hemophilic joints, hypertrophic scars, nonunion fractures, Osler-Weber syndrome, pyogenic granuloma, scleroderma, trachoma, and vascular adhesions.

[0647] Moreover, diseases, disorders, and/or conditions and/or states, which can be treated, prevented, and/or diagnosed with the polynucleotides, polypeptides, agonists and/or agonists include, but are not limited to, solid tumors,

blood born tumors such as leukemias, tumor metastasis, Kaposi's sarcoma, benign tumors, for example hemangiomas, acoustic neuromas, neurofibromas, trachomas, and pyogenic granulomas, rheumatoid arthritis, psoriasis, ocular angiogenic diseases, for example, diabetic retinopathy, retinopathy of prematurity, macular degeneration, corneal graft rejection, neovascular glaucoma, retrolental fibroplasia, rubeosis, retinoblastoma, and uvietis, delayed wound healing, endometriosis, vascluogenesis, granulations, hypertrophic scars (keloids), nonunion fractures, scleroderma, trachoma, vascular adhesions, myocardial angiogenesis, coronary collaterals, cerebral collaterals, arteriovenous malformations, ischemic limb angiogenesis, Osler-Webber Syndrome, plaque neovascularization, telangiectasia, hemophiliac joints, angiofibroma fibromuscular dysplasia, wound granulation, Crohn's disease, atherosclerosis, birth control agent by preventing vascularization required for embryo implantation controlling menstruation, diseases that have angiogenesis as a pathologic consequence such as cat scratch disease (Rochele minalia quintosa), ulcers (Helicobacter pylori), Bartonellosis and bacillary angiomatosis.

[0648] In one aspect of the birth control method, an amount of the compound sufficient to block embryo implantation is administered before or after intercourse and fertilization have occurred, thus providing an effective method of birth control, possibly a "morning after" method. Polynucleotides, polypeptides, agonists and/or agonists may also be used in controlling menstruation or administered as either a peritoneal lavage fluid or for peritoneal implantation in the treatment of endometriosis.

[0649] Polynucleotides, polypeptides, agonists and/or agonists of the present invention may be incorporated into surgical sutures in order to prevent stitch granulomas.

[0650] Polynucleotides, polypeptides, agonists and/or agonists may be utilized in a wide variety of surgical procedures. For example, within one aspect of the present invention a compositions (in the form of, for example, a spray or film) may be utilized to coat or spray an area prior to removal of a tumor, in order to isolate normal surrounding tissues from malignant tissue, and/or to prevent the spread of disease to surrounding tissues. Within other aspects of the present invention, compositions (e.g., in the form of a spray) may be delivered via endoscopic procedures in order to coat tumors, or inhibit angiogenesis in a desired locale. Within yet other aspects of the present invention, surgical meshes which have been coated with anti-angiogenic compositions of the present invention may be utilized in any procedure wherein a surgical mesh might be utilized. For example, within one embodiment of the invention a surgical mesh laden with an anti-angiogenic composition may be utilized during abdominal cancer resection surgery (e.g., subsequent to colon resection) in order to provide support to the structure, and to release an amount of the anti-angiogenic factor.

[0651] Within further aspects of the present invention, methods are provided for treating tumor excision sites, comprising administering a polynucleotide, polypeptide, agonist and/or agonist to the resection margins of a tumor subsequent to excision, such that the local recurrence of cancer and the formation of new blood vessels at the site is inhibited. Within one embodiment of the invention, the anti-angiogenic compound is administered directly to the tumor excision site (e.g., applied by swabbing, brushing or otherwise coating the resection margins of the tumor with the anti-angiogenic compound). Alternatively, the anti-angiogenic compounds may be incorporated into known surgical pastes prior to administration. Within particularly preferred embodiments of the invention, the anti-angiogenic compounds are applied after hepatic resections for malignancy, and after neurosurgical operations.

[0652] Within one aspect of the present invention, polynucleotides, polypeptides, agonists and/or agonists may be administered to the resection margin of a wide variety of tumors, including for example, breast, colon, brain and hepatic tumors. For example, within one embodiment of the invention, anti-angiogenic compounds may be administered to the site of a neurological tumor subsequent to excision, such that the formation of new blood vessels at the site are inhibited.

[0653] The polynucleotides, polypeptides, agonists and/or agonists of the present invention may also be administered along with other anti-angiogenic factors. Representative examples of other anti-angiogenic factors include: Anti-Invasive Factor, retinoic acid and derivatives thereof, paclitaxel, Suramin, Tissue Inhibitor of Metalloproteinase-1, Tissue Inhibitor of Metalloproteinase-2, Plasminogen Activator Inhibitor-1, Plasminogen Activator Inhibitor-2, and various forms of the lighter "d group" transition metals.

[0654] Lighter "d group" transition metals include, for example, vanadium, molybdenum, tungsten, titanium, niobium, and tantalum species. Such transition metal species may form transition metal complexes. Suitable complexes of the above-mentioned transition metal species include oxo transition metal complexes.

[0655] Representative examples of vanadium complexes include oxo vanadium complexes such as vanadate and vanadyl complexes. Suitable vanadate complexes include metavanadate and orthovanadate complexes such as, for example, ammonium metavanadate, sodium metavanadate, and sodium orthovanadate. Suitable vanadyl complexes include, for example, vanadyl acetylacetonate and vanadyl sulfate including vanadyl sulfate hydrates such as vanadyl sulfate mono- and trihydrates.

[0656] Representative examples of tungsten and molybdenum complexes also include oxo complexes. Suitable oxo tungsten complexes include tungstate and tungsten oxide complexes. Suitable tungstate complexes include ammonium tungstate, calcium tungstate, sodium tungstate dihydrate, and tungstic acid. Suitable tungsten oxides include tungsten (IV) oxide and tungsten (VI) oxide. Suitable oxo molybdenum complexes include molybdate, molybdenum oxide, and molybdenyl complexes. Suitable molybdate complexes include ammonium molybdate and its hydrates, sodium molybdate and its hydrates, and potassium molybdate and its hydrates. Suitable molybdenum oxides include molybdenum (VI) oxide, molybdenum (VI) oxide, and molybdic acid. Suitable molybdenyl complexes include, for example, molybdenyl acetylacetonate. Other suitable tungsten and molybdenum complexes include hydroxo derivatives derived from, for example, glycerol, tartaric acid, and sugars.

[0657] A wide variety of other anti-angiogenic factors may also be utilized within the context of the present invention.

Representative examples include platelet factor 4; protamine sulphate; sulphated chitin derivatives (prepared from queen crab shells), (Murata et al., Cancer Res. 51:22-26, 1991); Sulphated Polysaccharide Peptidoglycan Complex (SP-PG) (the function of this compound may be enhanced by the presence of steroids such as estrogen, and tamoxifen citrate); Staurosporine; modulators of matrix metabolism, including for example, proline analogs, cishydroxyproline, d,L-3,4-dehydroproline, Thiaproline, alpha,alpha-dipyridyl, aminopropionitrile fumarate; 4-propyl-5-(4-pyridinyl)-2(3H)-oxazolone; Methotrexate; Mitoxantrone; Heparin; Interferons; 2 Macroglobulin-serum; ChIMP-3 (Pavloff et al., J. Bio. Chem. 267:17321-17326, 1992); Chymostatin (Tomkinson et al., Biochem J. 286:475-480, 1992); Cyclodextrin Tetradecasulfate; Eponemycin; Camptothecin; Fumagillin (Ingber et al., Nature 348:555-557, 1990); Gold Sodium Thiomalate ("GST"; Matsubara and Ziff, J. Clin. Invest. 79:1440-1446, 1987); anticollagenase-serum; alpha2-antiplasmin (Holmes et al., J. Biol. Chem. 262(4):1659-1664, 1987); Bisantrene (National Cancer Institute); Lobenzarit disodium (N-(2)-carboxyphenyl-4-chloroanthronilic acid disodium or "CCA"; Takeuchi et al., Agents Actions 36:312-316, 1992); Thalidomide; Angostatic steroid; AGM-1470; carboxynaminolmidazole; and metalloproteinase inhibitors such as BB94.

Diseases at the Cellular Level

[0658] Diseases associated with increased cell survival or the inhibition of apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides and/or antagonists or agonists of the invention, include cancers (such as follicular lymphomas, carcinomas with p53 mutations, and hormone-dependent tumors, including, but not limited to colon cancer, cardiac tumors, pancreatic cancer, melanoma, retinoblastoma, glioblastoma, lung cancer, intestinal cancer, testicular cancer, stomach cancer, neuroblastoma, myxoma, myoma, lymphoma, endothelioma, osteoblastoma, osteoclastoma, osteosarcoma, chondrosarcoma, adenoma, breast cancer, prostate cancer, Kaposi's sarcoma and ovarian cancer); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) and viral infections (such as herpes viruses, pox viruses and adenoviruses), inflammation, graft v. host disease, acute graft rejection, and chronic graft rejection. In preferred embodiments, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention are used to inhibit growth, progression, and/or metastasis of cancers, in particular those listed above.

[0659] Additional diseases or conditions associated with increased cell survival that could be treated, prevented or diagnosed by the polynucleotides or polypeptides, or agonists or antagonists of the invention, include, but are not limited to, progression, and/or metastases of malignancies and related disorders such as leukemia (including acute leukemias (e.g., acute lymphocytic leukemia, acute myelocytic leukemia (including myeloblastic, promyelocytic, myelomonocytic, monocytic, and erythroleukemia)) and chronic leukemias (e.g., chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia)), polycythemia vera, lymphomas (e.g., Hodgkin's disease and non-Hodgkin's disease), multiple myeloma, Waldenstrom's macroglobulinemia, heavy chain disease, and solid tumors including, but not limited to, sarcomas and carcinomas such as fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leiomyosarcoma, rhabdomyosarcoma, colon carcinoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hepatoma, bile duct carcinoma, choriocarcinoma, seminoma, embryonal carcinoma, Wilm's tumor, cervical cancer, testicular tumor, lung carcinoma, small cell lung carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependymoma, pinealoma, hemangioblastoma, acoustic neuroma, oligodendroglioma, menangioma, melanoma, neuroblastoma, and retinoblastoma.

[0660] Diseases associated with increased apoptosis that could be treated, prevented, and/or diagnosed by the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, include AIDS; neurodegenerative diseases, disorders, and/or conditions (such as Alzheimer's disease, Parkinson's disease, Amyotrophic lateral sclerosis, Retinitis pigmentosa, Cerebellar degeneration and brain tumor or prior associated disease); autoimmune diseases, disorders, and/or conditions (such as, multiple sclerosis, Sjogren's syndrome, Hashimoto's thyroiditis, biliary cirrhosis, Behcet's disease, Crohn's disease, polymyositis, systemic lupus erythematosus and immune-related glomerulonephritis and rheumatoid arthritis) myelodysplastic syndromes (such as aplastic anemia), graft v. host disease, ischemic injury (such as that caused by myocardial infarction, stroke and reperfusion injury), liver injury (e.g., hepatitis related liver injury, ischemia/reperfusion injury, cholestosis (bile duct injury) and liver cancer); toxin-induced liver disease (such as that caused by alcohol), septic shock, cachexia and anorexia.

Wound Healing and Epithelial Cell Proliferation

[0661] In accordance with yet a further aspect of the present invention, there is provided a process for utilizing the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, for therapeutic purposes, for example, to stimulate epithelial cell proliferation and basal keratinocytes for the purpose of wound healing, and to stimulate hair follicle production and healing of dermal wounds. Polynucleotides or polypeptides, as well as agonists or antagonists of the invention, may be clinically useful in stimulating wound healing including surgical wounds, excisional wounds, deep wounds involving damage of the dermis and epidermis, eye tissue wounds, dental tissue wounds, oral cavity wounds, diabetic ulcers, dermal ulcers, cubitus ulcers, arterial ulcers, venous stasis ulcers, burns resulting from heat exposure or chemicals, and other abnormal wound healing conditions such as uremia, malnutrition, vitamin deficiencies and complications associated with systemic treatment with steroids, radiation therapy and antineoplastic drugs and antimetabolites. Polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote dermal reestablishment subsequent to dermal loss

[0662] The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to increase the adherence of skin grafts to a wound bed and to stimulate re-epithelialization from the wound bed. The following are a non-exhaustive list of grafts that polynucleotides or polypeptides, agonists or antagonists of the invention, could be used to increase adherence to a wound bed: autografts, artificial skin, allografts, autodermic graft, autoepidermic grafts, avacular grafts, Blair-Brown grafts, bone graft, brephoplastic grafts, cutis graft, delayed graft, dermic graft, epidermic graft, fascia graft, full thickness graft, heterologous graft, xenograft, homologous graft, hyperplastic graft, lamellar graft, mesh graft, mucosal graft, Ollier-Thiersch graft, omenpal graft, patch graft, pedicle graft, penetrating graft, split skin graft, thick split graft. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, can be used to promote skin strength and to improve the appearance of aged skin.

[0663] It is believed that the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, will also produce changes in hepatocyte proliferation, and epithelial cell proliferation in the lung, breast, pancreas, stomach, small intestine, and large intestine. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could promote proliferation of epithelial cells such as sebocytes, hair follicles, hepatocytes, type II pneumocytes, mucin-producing goblet cells, and other epithelial cells and their progenitors contained within the skin, lung, liver, and gastrointestinal tract. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may promote proliferation of endothelial cells, keratinocytes, and basal keratinocytes.

[0664] The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to reduce the side effects of gut toxicity that result from radiation, chemotherapy treatments or viral infections. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may have a cytoprotective effect on the small intestine mucosa. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, may also stimulate healing of mucositis (mouth ulcers) that result from chemotherapy and viral infections.

[0665] The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could further be used in full regeneration of skin in full and partial thickness skin defects, including burns, (i.e., repopulation of hair follicles, sweat glands, and sebaceous glands), treatment of other skin defects such as psoriasis. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat epidermolysis bullosa, a defect in adherence of the epidermis to the underlying dermis which results in frequent, open and painful blisters by accelerating reepithelialization of these lesions. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could also be used to treat gastric and doudenal ulcers and help heal by scar formation of the mucosal lining and regeneration of glandular mucosa and duodenal mucosal lining more rapidly. Inflamamatory bowel diseases, such as Crohn's disease and ulcerative colitis, are diseases which result in destruction of the mucosal surface of the small or large intestine, respectively. Thus, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to promote the resurfacing of the mucosal surface to aid more rapid healing and to prevent progression of inflammatory bowel disease. Treatment with the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, is expected to have a significant effect on the production of mucus throughout the gastrointestinal tract and could be used to protect the intestinal mucosa from injurious substances that are ingested or following surgery. The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to treat diseases associate with the under expression of the polynucleotides of the invention.

[0666] Moreover, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to prevent and heal damage to the lungs due to various pathological states. A growth factor such as the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, which could stimulate proliferation and differentiation and promote the repair of alveoli and brochiolar epithelium to prevent or treat acute or chronic lung damage. For example, emphysema, which results in the progressive loss of aveoli, and inhalation injuries, i.e., resulting from smoke inhalation and burns, that cause necrosis of the bronchiolar epithelium and alveoli could be effectively treated, prevented, and/or diagnosed using the polynucleotides or polypeptides, and/or agonists or antagonists of the invention. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to stimulate the proliferation of and differentiation of type II pneumocytes, which may help treat or prevent disease such as hyaline membrane diseases, such as infant respiratory distress syndrome and bronchopulmonary displasia, in premature infants.

[0667] The polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could stimulate the proliferation and differentiation of hepatocytes and, thus, could be used to alleviate or treat liver diseases and pathologies such as fulminant liver failure caused by cirrhosis, liver damage caused by viral hepatitis and toxic substances (i.e., acetaminophen, carbon tetraholoride and other hepatotoxins known in the art).

[0668] In addition, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used treat or prevent the onset of diabetes mellitus. In patients with newly diagnosed Types I and II diabetes, where some islet cell function remains, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used to maintain the islet function so as to alleviate, delay or prevent permanent manifestation of the disease. Also, the polynucleotides or polypeptides, and/or agonists or antagonists of the invention, could be used as an auxiliary in islet cell transplantation to improve or promote islet cell function.

Neurological Diseases

[0669] Nervous system diseases, disorders, and/or conditions, which can be treated, prevented, and/or diagnosed with the compositions of the invention (e.g., polypeptides, polynucleotides, and/or agonists or antagonists), include, but are not limited to, nervous system injuries, and diseases, disorders, and/or conditions which result in either a disconnection of axons, a diminution or degeneration of neurons, or demyelination. Nervous system lesions which may be

treated, prevented, and/or diagnosed in a patient (including human and non-human mammalian patients) according to the invention, include but are not limited to, the following lesions of either the central (including spinal cord, brain) or peripheral nervous systems: (1) ischemic lesions, in which a lack of oxygen in a portion of the nervous system results in neuronal injury or death, including cerebral infarction or ischemia, or spinal cord infarction or ischemia; (2) traumatic lesions, including lesions caused by physical injury or associated with surgery, for example, lesions which sever a portion of the nervous system, or compression injuries; (3) malignant lesions, in which a portion of the nervous system is destroyed or injured by malignant tissue which is either a nervous system associated malignancy or a malignancy derived from non-nervous system tissue; (4) infectious lesions, in which a portion of the nervous system is destroyed or injured as a result of infection, for example, by an abscess or associated with infection by human immunodeficiency virus, herpes zoster, or herpes simplex virus or with Lyme disease, tuberculosis, syphilis; (5) degenerative lesions, in which a portion of the nervous system is destroyed or injured as a result of a degenerative process including but not limited to degeneration associated with Parkinson's disease, Alzheimer's disease, Huntington's chorea, or amyotrophic lateral sclerosis (ALS); (6) lesions associated with nutritional diseases, disorders, and/or conditions, in which a portion of the nervous system is destroyed or injured by a nutritional disorder or disorder of metabolism including but not limited to, vitamin B12 deficiency, folic acid deficiency, Wernicke disease, tobacco-alcohol amblyopia, Marchiafava-Bignami disease (primary degeneration of the corpus callosum), and alcoholic cerebellar degeneration; (7) neurological lesions associated with systemic diseases including, but not limited to, diabetes (diabetic neuropathy, Bell's palsy), systemic lupus erythematosus, carcinoma, or sarcoidosis; (8) lesions caused by toxic substances including alcohol, lead, or particular neurotoxins; and (9) demyelinated lesions in which a portion of the nervous system is destroyed or injured by a demyelinating disease including, but not limited to, multiple sclerosis, human immunodeficiency virus-associated myelopathy, transverse myelopathy or various etiologies, progressive multifocal leukoencephalopathy, and central pontine myelinolysis.

[0670] In a preferred embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to protect neural cells from the damaging effects of cerebral hypoxia. According to this embodiment, the compositions of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral hypoxia. In one aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral ischemia. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with cerebral infarction. In another aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose or prevent neural cell injury associated with a stroke. In a further aspect of this embodiment, the polypeptides, polynucleotides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose neural cell injury associated with a heart attack.

[0671] The compositions of the invention which are useful for treating or preventing a nervous system disorder may be selected by testing for biological activity in promoting the survival or differentiation of neurons. For example, and not by way of limitation, compositions of the invention which elicit any of the following effects may be useful according to the invention: (1) increased survival time of neurons in culture; (2) increased sprouting of neurons in culture or in vivo; (3) increased production of a neuron-associated molecule in culture or in vivo, e.g., choline acetyltransferase or acetylcholinesterase with respect to motor neurons; or (4) decreased symptoms of neuron dysfunction in vivo. Such effects may be measured by any method known in the art. In preferred, non-limiting embodiments, increased survival of neurons may routinely be measured using a method set forth herein or otherwise known in the art, such as, for example, the method set forth in Arakawa et al. (J. Neurosci. 10:3507-3515 (1990)); increased sprouting of neurons may be detected by methods known in the art, such as, for example, the methods set forth in Pestronk et al. (Exp. Neurol. 70:65-82 (1980)) or Brown et al. (Ann. Rev. Neurosci. 4:17-42 (1981)); increased production of neuron-associated molecules may be measured by bioassay, enzymatic assay, antibody binding, Northern blot assay, etc., using techniques known in the art and depending on the molecule to be measured; and motor neuron dysfunction may be measured by assessing the physical manifestation of motor neuron disorder, e.g., weakness, motor neuron conduction velocity, or functional disability.

[0672] In specific embodiments, motor neuron diseases, disorders, and/or conditions that may be treated, prevented, and/or diagnosed according to the invention include, but are not limited to, diseases, disorders, and/or conditions such as infarction, infection, exposure to toxin, trauma, surgical damage, degenerative disease or malignancy that may affect motor neurons as well as other components of the nervous system, as well as diseases, disorders, and/or conditions that selectively affect neurons such as amyotrophic lateral sclerosis, and including, but not limited to, progressive spinal muscular atrophy, progressive bulbar palsy, primary lateral sclerosis, infantile and juvenile muscular atrophy, progressive bulbar paralysis of childhood (Fazio-Londe syndrome), poliomyelitis and the post polio syndrome, and Hereditary Motorsensory Neuropathy (Charcot-Marie-Tooth Disease).

Infectious Disease

[0673] A polypeptide or polynucleotide and/or agonist or antagonist of the present invention can be used to treat, prevent, and/or diagnose infectious agents. For example, by increasing the immune response, particularly increasing the proliferation and differentiation of B and/or T cells, infectious diseases may be treated, prevented, and/or diagnosed. The immune response may be increased by either enhancing an existing immune response, or by initiating a new immune response. Alternatively, polypeptide or polynucleotide and/or agonist or antagonist of the present invention may also directly inhibit the infectious agent, without necessarily eliciting an immune response.

[0674] Viruses are one example of an infectious agent that can cause disease or symptoms that can be treated, pre-

vented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention. Examples of viruses, include, but are not limited to Examples of viruses, include, but are not limited to the following DNA and RNA viruses and viral families: Arbovirus, Adenoviridae, Arenaviridae, Arterivirus, Birnaviridae, Bunyaviridae, Caliciviridae, Circoviridae, Coronaviridae, Dengue, EBV, HIV, Flaviviridae, Hepadnaviridae (Hepatitis), Herpesviridae (such as, Cytomegalovirus, Herpes Simplex, Herpes Zoster), Mononegavirus (e.g., Paramyxoviridae, Morbillivirus, Rhabdoviridae), Orthomyxoviridae (e.g., Influenza A, Influenza B, and parainfluenza), Papiloma virus, Papovaviridae, Parvoviridae, Picornaviridae, Poxyiridae (such as Smallpox or Vaccinia), Reoviridae (e.g., Rotavirus), Retroviridae (HTLV-I, HTLV-II, Lentivirus), and Togaviridae (e.g., Rubivirus). Viruses falling within these families can cause a variety of diseases or symptoms, including, but not limited to: arthritis, bronchiollitis, respiratory syncytial virus, encephalitis, eye infections (e.g., conjunctivitis, keratitis), chronic fatigue syndrome, hepatitis (A, B, C, E, Chronic Active, Delta), Japanese B encephalitis, Junin, Chikungunya, Rift Valley fever, yellow fever, meningitis, opportunistic infections (e.g., AIDS), pneumonia, Burkitt's Lymphoma, chickenpox, hemorrhagic fever, Measles, Mumps, Parainfluenza, Rabies, the common cold, Polio, leukemia, Rubella, sexually transmitted diseases, skin diseases (e.g., Kaposi's, warts), and viremia. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose: meningitis, Dengue, EBV, and/or hepatitis (e.g., hepatitis B). In an additional specific embodiment polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat patients nonresponsive to one or more other commercially available hepatitis vaccines. In a further specific embodiment polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose AIDS.

[0675] Similarly, bacterial or fungal agents that can cause disease or symptoms and that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, include, but not limited to, the following Gram-Negative and Gram-positive bacteria and bacterial families and fungi: Actinomycetales (e.g., Corynebacterium, Mycobacterium, Norcardia), *Cryptococcus neoformans*, Aspergillosis, Bacillaceae (e.g., Anthrax, Clostridium), Bacteroidaceae, Blastomycosis, Bordetella, Borrelia (e.g., *Borrelia burgdorferi*), Brucellosis, Candidiasis, Campylobacter, Coccidioidomycosis, Cryptococcosis, Dermatocycoses, *E. coli* (e.g., Enterotoxigenic *E. coli* and Enterohemorrhagic *E. coli*), Enterobacteriaceae (Klebsiella, Salmonella (e.g., *Salmonella typhi*, and *Salmonella paratyphi*), Serratia, Yersinia), Erysipelothrix, Helicobacter, Legionellosis, Leptospirosis, Listeria, Mycoplasmatales, *Mycobacterium leprae, Vibrio cholerae*, Neisseriaceae (e.g., Acinetobacter, Gonorrhea, Menigococcal), *Meisseria meningitidis*, Pasteurellacea Infections (e.g., Actinobacillus, Heamophilus (e.g., *Heamophilus influenza* type B), Pasteurella), Pseudomonas, Rickettsiaceae, Chlamydiaceae, Syphilis, Shigella spp., Staphylococcal, Meningiococcal, Pneumococcal and Streptococcal (e.g., *Streptococcus pneumoniae* and Group B Streptococcus). These bacterial or fungal families can cause the following diseases or symptoms, including, but not limited to: bacteremia, endocarditis, eye infections (conjunctivitis, tuberculosis, uveitis), gingivitis, opportunistic infections (e.g., AIDS related infections), paronychia, prosthesis-related infections, Reiter's Disease, respiratory tract infections, such as Whooping Cough or Empyema, sepsis, Lyme Disease, Cat-Scratch Disease, Dysentery, Paratyphoid Fever, food poisoning, Typhoid, pneumonia, Gonorrhea, meningitis (e.g., mengitis types A and B), Chlamydia, Syphilis, Diphtheria, Leprosy, Paratuberculosis, Tuberculosis, Lupus, Botulism, gangrene, tetanus, impetigo, Rheumatic Fever, Scarlet Fever, sexually transmitted diseases, skin diseases (e.g., cellulitis, dermatocycoses), toxemia, urinary tract infections, wound infections. Polynucleotides or polypeptides, agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, agonists or antagonists of the invention are used to treat, prevent, and/or diagnose: tetanus, Diptheria, botulism, and/or meningitis type B.

[0676] Moreover, parasitic agents causing disease or symptoms that can be treated, prevented, and/or diagnosed by a polynucleotide or polypeptide and/or agonist or antagonist of the present invention include, but not limited to, the following families or class: Amebiasis, Babesiosis, Coccidiosis, Cryptosporidiosis, Dientamoebiasis, Dourine, Ectoparasitic, Giardiasis, Helminthiasis, Leishmaniasis, Theileriasis, Toxoplasmosis, Trypanosomiasis, and Trichomonas and Sporozoans (e.g., *Plasmodium virax, Plasmodium falciparium, Plasmodium malariae* and *Plasmodium ovale*). These parasites can cause a variety of diseases or symptoms, including, but not limited to: Scabies, Trombiculiasis, eye infections, intestinal disease (e.g., dysentery, giardiasis), liver disease, lung disease, opportunistic infections (e.g., AIDS related), malaria, pregnancy complications, and toxoplasmosis. polynucleotides or polypeptides, or agonists or antagonists of the invention, can be used to treat, prevent, and/or diagnose any of these symptoms or diseases. In specific embodiments, polynucleotides, polypeptides, or agonists or antagonists of the invention are used to treat, prevent, and/or diagnose malaria.

[0677] Preferably, treatment or prevention using a polypeptide or polynucleotide and/or agonist or antagonist of the present invention could either be by administering an effective amount of a polypeptide to the patient, or by removing cells from the patient, supplying the cells with a polynucleotide of the present invention, and returning the engineered cells to the patient (ex vivo therapy). Moreover, the polypeptide or polynucleotide of the present invention can be used as an antigen in a vaccine to raise an immune response against infectious disease.

Regeneration

[0678] A polynucleotide or polypeptide and/or agonist or antagonist of the present invention can be used to differentiate, proliferate, and attract cells, leading to the regeneration of tissues. (See, Science 276:59-87 (1997).) The regeneration of tissues could be used to repair, replace, or protect tissue damaged by congenital defects, trauma (wounds, burns, incisions, or ulcers), age, disease (e.g. osteoporosis, osteocarthritis, periodontal disease, liver failure), surgery, including cosmetic plastic surgery, fibrosis, reperfusion injury, or systemic cytokine damage.

[0679] Tissues that could be regenerated using the present invention include organs (e.g., pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac), vasculature (including vascular and lymphatics), nervous, hematopoietic, and skeletal (bone, cartilage, tendon, and ligament) tissue. Preferably, regeneration occurs without or decreased scarring. Regeneration also may include angiogenesis.

[0680] Moreover, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase regeneration of tissues difficult to heal. For example, increased tendon/ligament regeneration would quicken recovery time after damage. A polynucleotide or polypeptide and/or agonist or antagonist of the present invention could also be used prophylactically in an effort to avoid damage. Specific diseases that could be treated, prevented, and/or diagnosed include of tendinitis, carpal tunnel syndrome, and other tendon or ligament defects. A further example of tissue regeneration of non-healing wounds includes pressure ulcers, ulcers associated with vascular insufficiency, surgical, and traumatic wounds.

[0681] Similarly, nerve and brain tissue could also be regenerated by using a polynucleotide or polypeptide and/or agonist or antagonist of the present invention to proliferate and differentiate nerve cells. Diseases that could be treated, prevented, and/or diagnosed using this method include central and peripheral nervous system diseases, neuropathies, or mechanical and traumatic diseases, disorders, and/or conditions (e.g., spinal cord disorders, head trauma, cerebrovascular disease, and stoke). Specifically, diseases associated with peripheral nerve injuries, peripheral neuropathy (e.g., resulting from chemotherapy or other medical therapies), localized neuropathies, and central nervous system diseases (e.g., Alzheimer's disease, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome), could all be treated, prevented, and/or diagnosed using the polynucleotide or polypeptide and/or agonist or antagonist of the present invention.

Chemotaxis

[0682] A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may have chemotaxis activity. A chemotaxic molecule attracts or mobilizes cells (e.g., monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells) to a particular site in the body, such as inflammation, infection, or site of hyperproliferation. The mobilized cells can then fight off and/or heal the particular trauma or abnormality.

[0683] A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase chemotaxic activity of particular cells. These chemotactic molecules can then be used to treat, prevent, and/or diagnose inflammation, infection, hyperproliferative diseases, disorders, and/or conditions, or any immune system disorder by increasing the number of cells targeted to a particular location in the body. For example, chemotaxic molecules can be used to treat, prevent, and/or diagnose wounds and other trauma to tissues by attracting immune cells to the injured location. Chemotactic molecules of the present invention can also attract fibroblasts, which can be used to treat, prevent, and/or diagnose wounds.

[0684] It is also contemplated that a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may inhibit chemotactic activity. These molecules could also be used to treat, prevent, and/or diagnose diseases, disorders, and/or conditions. Thus, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention could be used as an inhibitor of chemotaxis.

Binding Activity

[0685] A polypeptide of the present invention may be used to screen for molecules that bind to the polypeptide or for molecules to which the polypeptide binds. The binding of the polypeptide and the molecule may activate (agonist), increase, inhibit (antagonist), or decrease activity of the polypeptide or the molecule bound. Examples of such molecules include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

[0686] Preferably, the molecule is closely related to the natural ligand of the polypeptide, e.g., a fragment of the ligand, or a natural substrate, a ligand, a structural or functional mimetic. (See, Coligan et al., Current Protocols in Immunology 1(2):Chapter 5 (1991).) Similarly, the molecule can be closely related to the natural receptor to which the polypeptide binds, or at least, a fragment of the receptor capable of being bound by the polypeptide (e.g., active site). In either case, the molecule can be rationally designed using known techniques.

[0687] Preferably, the screening for these molecules involves producing appropriate cells which express the polypeptide, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or *E. coli*. Cells expressing the polypeptide (or cell membrane containing the expressed polypeptide) are then preferably contacted with a test compound potentially containing the molecule to observe binding, stimulation, or inhibition of activity of either the polypeptide or the molecule.

[0688] The assay may simply test binding of a candidate compound to the polypeptide, wherein binding is detected by a label, or in an assay involving competition with a labeled competitor. Further, the assay may test whether the candidate compound results in a signal generated by binding to the polypeptide.

[0689] Alternatively, the assay can be carried out using cell-free preparations, polypeptide/molecule affixed to a solid support, chemical libraries, or natural product mixtures. The assay may also simply comprise the steps of mixing a candidate compound with a solution containing a polypeptide, measuring polypeptide/molecule activity or binding, and comparing the polypeptide/molecule activity or binding to a standard.

[0690] Preferably, an ELISA assay can measure polypeptide level or activity in a sample (e.g., biological sample) using a monoclonal or polyclonal antibody. The antibody can measure polypeptide level or activity by either binding, directly or indirectly, to the polypeptide or by competing with the polypeptide for a substrate.

[0691] Additionally, the receptor to which a polypeptide of the invention binds can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting (Coligan, et al., Current Protocols in Immun., 1(2), Chapter 5, (1991)). For example, expression cloning is employed wherein polyadenylated RNA is

prepared from a cell responsive to the polypeptides, for example, NIH3T3 cells which are known to contain multiple receptors for the FGF family proteins, and SC-3 cells, and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the polypeptides. Transfected cells which are grown on glass slides are exposed to the polypeptide of the present invention, after they have been labeled. The polypeptides can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase.

[0692] Following fixation and incubation, the slides are subjected to auto-radiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an iterative sub-pooling and re-screening process, eventually yielding a single clones that encodes the putative receptor.

[0693] As an alternative approach for receptor identification, the labeled polypeptides can be photoaffinity linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE analysis and exposed to X-ray film. The labeled complex containing the receptors of the polypeptides can be excised, resolved into peptide fragments, and subjected to protein microsequencing. The amino acid sequence obtained from microsequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the genes encoding the putative receptors.

[0694] Moreover, the techniques of gene-shuffling, motif-shuffling, exon-shuffling, and/or codon-shuffling (collectively referred to as "DNA shuffling") may be employed to modulate the activities of polypeptides of the invention thereby effectively generating agonists and antagonists of polypeptides of the invention. See generally, U.S. Pat. Nos. 5,605,793, 5,811,238, 5,830,721, 5,834,252, and 5,837,458, and Patten, P. A., et al., Curr. Opinion Biotechnol. 8:724-33 (1997); Harayama, S. Trends Biotechnol. 16(2):76-82 (1998); Hansson, L. O., et al., J. Mol. Biol. 287:265-76 (1999); and Lorenzo, M. M. and Blasco, R. Biotechniques 24(2):308-13 (1998) (each of these patents and publications are hereby incorporated by reference). In one embodiment, alteration of polynucleotides and corresponding polypeptides of the invention may be achieved by DNA shuffling. DNA shuffling involves the assembly of two or more DNA segments into a desired polynucleotide sequence of the invention molecule by homologous, or site-specific, recombination. In another embodiment, polynucleotides and corresponding polypeptides of the invention may be altered by being subjected to random mutagenesis by error-prone PCR, random nucleotide insertion or other methods prior to recombination. In another embodiment, one or more components, motifs, sections, parts, domains, fragments, etc., of the polypeptides of the invention may be recombined with one or more components, motifs, sections, parts, domains, fragments, etc. of one or more heterologous molecules. In preferred embodiments, the heterologous molecules are family members. In further preferred embodiments, the heterologous molecule is a growth factor such as, for example, platelet-derived growth factor (PDGF), insulin-like growth factor (IGF-I), transforming growth factor (TGF)-alpha, epidermal growth factor (EGF), fibroblast growth factor (FGF), TGF-beta, bone morphogenetic protein (BMP)-2, BMP-4, BMP-5, BMP-6, BMP-7, activins A and B, decapentaplegic(dpp), 60A, OP-2, dorsalin, growth differentiation factors (GDFs), nodal, MIS, inhibin-alpha, TGF-beta1, TGF-beta2, TGF-beta3, TGF-beta5, and glial-derived neurotrophic factor (GDNF).

[0695] Other preferred fragments are biologically active fragments of the polypeptides of the invention. Biologically active fragments are those exhibiting activity similar, but not necessarily identical, to an activity of the polypeptide. The biological activity of the fragments may include an improved desired activity, or a decreased undesirable activity.

[0696] Additionally, this invention provides a method of screening compounds to identify those which modulate the action of the polypeptide of the present invention. An example of such an assay comprises combining a mammalian fibroblast cell, a the polypeptide of the present invention, the compound to be screened and 3[H] thymidine under cell culture conditions where the fibroblast cell would normally proliferate. A control assay may be performed in the absence of the compound to be screened and compared to the amount of fibroblast proliferation in the presence of the compound to determine if the compound stimulates proliferation by determining the uptake of 3[H] thymidine in each case. The amount of fibroblast cell proliferation is measured by liquid scintillation chromatography which measures the incorporation of 3[H] thymidine. Both agonist and antagonist compounds may be identified by this procedure.

[0697] In another method, a mammalian cell or membrane preparation expressing a receptor for a polypeptide of the present invention is incubated with a labeled polypeptide of the present invention in the presence of the compound. The ability of the compound to enhance or block this interaction could then be measured. Alternatively, the response of a known second messenger system following interaction of a compound to be screened and the receptor is measured and the ability of the compound to bind to the receptor and elicit a second messenger response is measured to determine if the compound is a potential agonist or antagonist. Such second messenger systems include but are not limited to, cAMP guanylate cyclase, ion channels or phosphoinositide hydrolysis.

[0698] All of these above assays can be used as diagnostic or prognostic markers. The molecules discovered using these assays can be used to treat, prevent, and/or diagnose disease or to bring about a particular result in a patient (e.g., blood vessel growth) by activating or inhibiting the polypeptide/molecule. Moreover, the assays can discover agents which may inhibit or enhance the production of the polypeptides of the invention from suitably manipulated cells or tissues. Therefore, the invention includes a method of identifying compounds which bind to the polypeptides of the invention comprising the steps of: (a) incubating a candidate binding compound with the polypeptide; and (b) determining if binding has occurred. Moreover, the invention includes a method of identifying agonists/antagonists comprising the steps of: (a) incubating a candidate compound with the polypeptide, (b) assaying a biological activity, and (b) determining if a biological activity of the polypeptide has been altered.

[0699] Also, one could identify molecules bind a polypeptide of the invention experimentally by using the beta-pleated sheet regions contained in the polypeptide sequence

of the protein. Accordingly, specific embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, the amino acid sequence of each beta pleated sheet regions in a disclosed polypeptide sequence. Additional embodiments of the invention are directed to polynucleotides encoding polypeptides which comprise, or alternatively consist of, any combination or all of contained in the polypeptide sequences of the invention. Additional preferred embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, the amino acid sequence of each of the beta pleated sheet regions in one of the polypeptide sequences of the invention. Additional embodiments of the invention are directed to polypeptides which comprise, or alternatively consist of, any combination or all of the beta pleated sheet regions in one of the polypeptide sequences of the invention.

Targeted Delivery

[0700] In another embodiment, the invention provides a method of delivering compositions to targeted cells expressing a receptor for a polypeptide of the invention, or cells expressing a cell bound form of a polypeptide of the invention.

[0701] As discussed herein, polypeptides or antibodies of the invention may be associated with heterologous polypeptides, heterologous nucleic acids, toxins, or prodrugs via hydrophobic, hydrophilic, ionic and/or covalent interactions. In one embodiment, the invention provides a method for the specific delivery of compositions of the invention to cells by administering polypeptides of the invention (including antibodies) that are associated with heterologous polypeptides or nucleic acids. In one example, the invention provides a method for delivering a therapeutic protein into the targeted cell. In another example, the invention provides a method for delivering a single stranded nucleic acid (e.g., antisense or ribozymes) or double stranded nucleic acid (e.g., DNA that can integrate into the cell's genome or replicate episomally and that can be transcribed) into the targeted cell.

[0702] In another embodiment, the invention provides a method for the specific destruction of cells (e.g., the destruction of tumor cells) by administering polypeptides of the invention (e.g., polypeptides of the invention or antibodies of the invention) in association with toxins or cytotoxic prodrugs.

[0703] By "toxin" is meant compounds that bind and activate endogenous cytotoxic effector systems, radioisotopes, holotoxins, modified toxins, catalytic subunits of toxins, or any molecules or enzymes not normally present in or on the surface of a cell that under defined conditions cause the cell's death. Toxins that may be used according to the methods of the invention include, but are not limited to, radioisotopes known in the art, compounds such as, for example, antibodies (or complement fixing containing portions thereof) that bind an inherent or induced endogenous cytotoxic effector system, thymidine kinase, endonuclease, RNAse, alpha toxin, ricin, abrin, Pseudomonas exotoxin A, diphtheria toxin, saporin, momordin, gelonin, pokeweed antiviral protein, alpha-sarcin and cholera toxin. By "cytotoxic prodrug" is meant a non-toxic compound that is converted by an enzyme, normally present in the cell, into a cytotoxic compound. Cytotoxic prodrugs that may be used according to the methods of the invention include, but are not limited to, glutamyl derivatives of benzoic acid mustard alkylating agent, phosphate derivatives of etoposide or mitomycin C, cytosine arabinoside, daunorubisin, and phenoxyacetamide derivatives of doxorubicin.

Drug Screening

[0704] Further contemplated is the use of the polypeptides of the present invention, or the polynucleotides encoding these polypeptides, to screen for molecules which modify the activities of the polypeptides of the present invention. Such a method would include contacting the polypeptide of the present invention with a selected compound(s) suspected of having antagonist or agonist activity, and assaying the activity of these polypeptides following binding.

[0705] This invention is particularly useful for screening therapeutic compounds by using the polypeptides of the present invention, or binding fragments thereof, in any of a variety of drug screening techniques. The polypeptide or fragment employed in such a test may be affixed to a solid support, expressed on a cell surface, free in solution, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. One may measure, for example, the formulation of complexes between the agent being tested and a polypeptide of the present invention.

[0706] Thus, the present invention provides methods of screening for drugs or any other agents which affect activities mediated by the polypeptides of the present invention. These methods comprise contacting such an agent with a polypeptide of the present invention or a fragment thereof and assaying for the presence of a complex between the agent and the polypeptide or a fragment thereof, by methods well known in the art. In such a competitive binding assay, the agents to screen are typically labeled. Following incubation, free agent is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of a particular agent to bind to the polypeptides of the present invention.

[0707] Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to the polypeptides of the present invention, and is described in great detail in European Patent Application 84/03564, published on Sep. 13, 1984, which is incorporated herein by reference herein. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. The peptide test compounds are reacted with polypeptides of the present invention and washed. Bound polypeptides are then detected by methods well known in the art. Purified polypeptides are coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies may be used to capture the peptide and immobilize it on the solid support.

[0708] This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding polypeptides of the present invention specifically compete with a test compound for binding

to the polypeptides or fragments thereof. In this manner, the antibodies are used to detect the presence of any peptide which shares one or more antigenic epitopes with a polypeptide of the invention.

[0709] The human Protease-19 polypeptides and/or peptides of the present invention, or immunogenic fragments or oligopeptides thereof, can be used for screening therapeutic drugs or compounds in a variety of drug screening techniques. The fragment employed in such a screening assay may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The reduction or abolition of activity of the formation of binding complexes between the ion channel protein and the agent being tested can be measured. Thus, the present invention provides a method for screening or assessing a plurality of compounds for their specific binding affinity with a Protease-19 polypeptide, or a bindable peptide fragment, of this invention, comprising providing a plurality of compounds, combining the Protease-19 polypeptide, or a bindable peptide fragment, with each of a plurality of compounds for a time sufficient to allow binding under suitable conditions and detecting binding of the Protease-19 polypeptide or peptide to each of the plurality of test compounds, thereby identifying the compounds that specifically bind to the Protease-19 polypeptide or peptide.

[0710] Methods of identifying compounds that modulate the activity of the novel human Protease-19 polypeptides and/or peptides are provided by the present invention and comprise combining a potential or candidate compound or drug modulator of cystein protease biological activity with an Protease-19 polypeptide or peptide, for example, the Protease-19 amino acid sequence as set forth in SEQ ID NO:2, and measuring an effect of the candidate compound or drug modulator on the biological activity of the Protease-19 polypeptide or peptide. Such measurable effects include, for example, physical binding interaction; the ability to cleave a suitable cystein protease substrate; effects on native and cloned Protease-19-expressing cell line; and effects of modulators or other cystein protease-mediated physiological measures.

[0711] Another method of identifying compounds that modulate the biological activity of the novel Protease-19 polypeptides of the present invention comprises combining a potential or candidate compound or drug modulator of a cystein protease biological activity with a host cell that expresses the Protease-19 polypeptide and measuring an effect of the candidate compound or drug modulator on the biological activity of the Protease-19 polypeptide. The host cell can also be capable of being induced to express the Protease-19 polypeptide, e.g., via inducible expression. Physiological effects of a given modulator candidate on the Protease-19 polypeptide can also be measured. Thus, cellular assays for particular cystein protease modulators may be either direct measurement or quantification of the physical biological activity of the Protease-19 polypeptide, or they may be measurement or quantification of a physiological effect. Such methods preferably employ a Protease-19 polypeptide as described herein, or an overexpressed recombinant Protease-19 polypeptide in suitable host cells containing an expression vector as described herein, wherein the Protease-19 polypeptide is expressed, overexpressed, or undergoes upregulated expression.

[0712] Another aspect of the present invention embraces a method of screening for a compound that is capable of modulating the biological activity of a Protease-19 polypeptide, comprising providing a host cell containing an expression vector harboring a nucleic acid sequence encoding a Protease-19 polypeptide, or a functional peptide or portion thereof (e.g., SEQ ID NOS:2); determining the biological activity of the expressed Protease-19 polypeptide in the absence of a modulator compound; contacting the cell with the modulator compound and determining the biological activity of the expressed Protease-19 polypeptide in the presence of the modulator compound. In such a method, a difference between the activity of the Protease-19 polypeptide in the presence of the modulator compound and in the absence of the modulator compound indicates a modulating effect of the compound.

[0713] Essentially any chemical compound can be employed as a potential modulator or ligand in the assays according to the present invention. Compounds tested as cystein protease modulators can be any small chemical compound, or biological entity (e.g., protein, sugar, nucleic acid, lipid). Test compounds will typically be small chemical molecules and peptides. Generally, the compounds used as potential modulators can be dissolved in aqueous or organic (e.g., DMSO-based) solutions. The assays are designed to screen large chemical libraries by automating the assay steps and providing compounds from any convenient source. Assays are typically run in parallel, for example, in microtiter formats on microtiter plates in robotic assays. There are many suppliers of chemical compounds, including Sigma (St. Louis, Mo.), Aldrich (St. Louis, Mo.), Sigma-Aldrich (St. Louis, Mo.), Fluka Chemika-Biochemica Analytika (Buchs, Switzerland), for example. Also, compounds may be synthesized by methods known in the art.

[0714] High throughput screening methodologies are particularly envisioned for the detection of modulators of the novel Protease-19 polynucleotides and polypeptides described herein. Such high throughput screening methods typically involve providing a combinatorial chemical or peptide library containing a large number of potential therapeutic compounds (e.g., ligand or modulator compounds). Such combinatorial chemical libraries or ligand libraries are then screened in one or more assays to identify those library members (e.g., particular chemical species or subclasses) that display a desired characteristic activity. The compounds so identified can serve as conventional lead compounds, or can themselves be used as potential or actual therapeutics.

[0715] A combinatorial chemical library is a collection of diverse chemical compounds generated either by chemical synthesis or biological synthesis, by combining a number of chemical building blocks (i.e., reagents such as amino acids). As an example, a linear combinatorial library, e.g., a polypeptide or peptide library, is formed by combining a set of chemical building blocks in every possible way for a given compound length (i.e., the number of amino acids in a polypeptide or peptide compound). Millions of chemical compounds can be synthesized through such combinatorial mixing of chemical building blocks.

[0716] The preparation and screening of combinatorial chemical libraries is well known to those having skill in the pertinent art. Combinatorial libraries include, without limitation, peptide libraries (e.g. U.S. Pat. No. 5,010,175; Furka,

1991, *Int. J. Pept. Prot. Res.,* 37:487-493; and Houghton et al., 1991, *Nature,* 354:84-88). Other chemistries for generating chemical diversity libraries can also be used. Nonlimiting examples of chemical diversity library chemistries include, peptoids (PCT Publication No. WO 91/019735), encoded peptides (PCT Publication No. WO 93/20242), random bio-oligomers (PCT Publication No. WO 92/00091), benzodiazepines (U.S. Pat. No. 5,288,514), diversomers such as hydantoins, benzodiazepines and dipeptides (Hobbs et al., 1993, *Proc. Natl. Acad. Sci. USA,* 90:6909-6913), vinylogous polypeptides (Hagihara et al., 1992, *J. Amer. Chem. Soc.,* 114:6568), nonpeptidal peptidomimetics with glucose scaffolding (Hirschmann et al., 1992, *J. Amer. Chem. Soc.,* 114:9217-9218), analogous organic synthesis of small compound libraries (Chen et al., 1994, *J. Amer. Chem. Soc.,* 116:2661), oligocarbamates (Cho et al., 1993, *Science,* 261:1303), and/or peptidyl phosphonates (Campbell et al., 1994, *J. Org. Chem.,* 59:658), nucleic acid libraries (see Ausubel, Berger and Sambrook, all supra), peptide nucleic acid libraries (U.S. Pat. No. 5,539,083), antibody libraries (e.g., Vaughn et al., 1996, *Nature Biotechnology,* 14(3):309-314) and PCT/US96/10287), carbohydrate libraries (e.g., Liang et al., 1996, *Science,* 274-1520-1522) and U.S. Pat. No. 5,593,853), small organic molecule libraries (e.g., benzodiazepines, Baum C&EN, Jan. 18, 1993, page 33; and U.S. Pat. No. 5,288,514; isoprenoids, U.S. Pat. No. 5,569,588; thiazolidinones and metathiazanones, U.S. Pat. No. 5,549,974; pyrrolidines, U.S. Pat. Nos. 5,525,735 and 5,519,134; morpholino compounds, U.S. Pat. No. 5,506,337; and the like).

[**0717**] Devices for the preparation of combinatorial libraries are commercially available (e.g., 357 MPS, 390 MPS, Advanced Chem Tech, Louisville Ky.; Symphony, Rainin, Woburn, Mass.; 433A Applied Biosystems, Foster City, Calif.; 9050 Plus, Millipore, Bedford, Mass.). In addition, a large number of combinatorial libraries are commercially available (e.g., ComGenex, Princeton, N.J.; Asinex, Moscow, Russia; Tripos, Inc., St. Louis, Mo.; ChemStar, Ltd., Moscow, Russia; 3D Pharmaceuticals, Exton, Pa; Martek Biosciences, Columbia, Md, and the like).

[**0718**] In one embodiment, the invention provides solid phase based in vitro assays in a high throughput format, where the cell or tissue expressing an ion channel is attached to a solid phase substrate. In such high throughput assays, it is possible to screen up to several thousand different modulators or ligands in a single day. In particular, each well of a microtiter plate can be used to perform a separate assay against a selected potential modulator, or, if concentration or incubation time effects are to be observed, every 5-10 wells can test a single modulator. Thus, a single standard microtiter plate can assay about 96 modulators. If 1536 well plates are used, then a single plate can easily assay from about 100 to about 1500 different compounds. It is possible to assay several different plates per day; thus, for example, assay screens for up to about 6,000-20,000 different compounds are possible using the described integrated systems.

[**0719**] In another of its aspects, the present invention encompasses screening and small molecule (e.g., drug) detection assays which involve the detection or identification of small molecules that can bind to a given protein, i.e., a Protease-19 polypeptide or peptide. Particularly preferred are assays suitable for high throughput screening methodologies.

[**0720**] In such binding-based detection, identification, or screening assays, a functional assay is not typically required. All that is needed is a target protein, preferably substantially purified, and a library or panel of compounds (e.g., ligands, drugs, small molecules) or biological entities to be screened or assayed for binding to the protein target. Preferably, most small molecules that bind to the target protein will modulate activity in some manner, due to preferential, higher affinity binding to functional areas or sites on the protein.

[**0721**] An example of such an assay is the fluorescence based thermal shift assay (3-Dimensional Pharmaceuticals, Inc., 3DP, Exton, Pa.) as described in U.S. Pat. Nos. 6,020, 141 and 6,036,920 to Pantoliano et al.; see also, J. Zimmerman, 2000, *Gen. Eng. News,* 20(8)). The assay allows the detection of small molecules (e.g., drugs, ligands) that bind to expressed, and preferably purified, ion channel polypeptide based on affinity of binding determinations by analyzing thermal unfolding curves of protein-drug or ligand complexes. The drugs or binding molecules determined by this technique can be further assayed, if desired, by methods, such as those described herein, to determine if the molecules affect or modulate function or activity of the target protein.

[**0722**] To purify a Protease-19 polypeptide or peptide to measure a biological binding or ligand binding activity, the source may be a whole cell lysate that can be prepared by successive freeze-thaw cycles (e.g., one to three) in the presence of standard protease inhibitors. The Protease-19 polypeptide may be partially or completely purified by standard protein purification methods, e.g., affinity chromatography using specific antibody described infra, or by ligands specific for an epitope tag engineered into the recombinant Protease-19 polypeptide molecule, also as described herein. Binding activity can then be measured as described.

[**0723**] (Compounds which are identified according to the methods provided herein, and which modulate or regulate the biological activity or physiology of the Protease-19 polypeptides according to the present invention are a preferred embodiment of this invention. It is contemplated that such modulatory compounds may be employed in treatment and therapeutic methods for treating a condition that is mediated by the novel Protease-19 polypeptides by administering to an individual in need of such treatment a therapeutically effective amount of the compound identified by the methods described herein.

[**0724**] In addition, the present invention provides methods for treating an individual in need of such treatment for a disease, disorder, or condition that is mediated by the Protease-19 polypeptides of the invention, comprising administering to the individual a therapeutically effective amount of the Protease-19-modulating compound identified by a method provided herein.

Antisense and Ribozyme (Antagonists)

[**0725**] In specific embodiments, antagonists according to the present invention are nucleic acids corresponding to the sequences contained in SEQ ID NO:1, or the complementary strand thereof, and/or to nucleotide sequences contained a deposited clone. In one embodiment, antisense sequence is generated internally by the organism, in another embodiment, the antisense sequence is separately administered (see, for example, O'Connor, Neurochem., 56:560 (1991). Oli-

godeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, Fla. (1988). Antisense technology can be used to control gene expression through antisense DNA or RNA, or through triple-helix formation. Antisense techniques are discussed for example, in Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, Fla. (1988). Triple helix formation is discussed in, for instance, Lee et al., Nucleic Acids Research, 6:3073 (1979); Cooney et al., Science, 241:456 (1988); and Dervan et al., Science, 251:1300 (1991). The methods are based on binding of a polynucleotide to a complementary DNA or RNA.

[0726] For example, the use of c-myc and c-myb antisense RNA constructs to inhibit the growth of the non-lymphocytic leukemia cell line HL-60 and other cell lines was previously described. (Wickstrom et al. (1988); Anfossi et al. (1989)). These experiments were performed in vitro by incubating cells with the oligoribonucleotide. A similar procedure for in vivo use is described in WO 91/15580. Briefly, a pair of oligonucleotides for a given antisense RNA is produced as follows: A sequence complimentary to the first 15 bases of the open reading frame is flanked by an EcoRI site on the 5 end and a HindIII site on the 3 end. Next, the pair of oligonucleotides is heated at 90° C. for one minute and then annealed in 2×ligation buffer (20 mM TRIS HCl pH 7.5, 10 mM MgCl2, 10 MM dithiothreitol (DTT) and 0.2 mM ATP) and then ligated to the EcoRI/Hind III site of the retroviral vector PMV7 (WO 91/15580).

[0727] For example, the 5' coding portion of a polynucleotide that encodes the mature polypeptide of the present invention may be used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription thereby preventing transcription and the production of the receptor. The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into receptor polypeptide.

[0728] In one embodiment, the antisense nucleic acid of the invention is produced intracellularly by transcription from an exogenous sequence. For example, a vector or a portion thereof, is transcribed, producing an antisense nucleic acid (RNA) of the invention. Such a vector would contain a sequence encoding the antisense nucleic acid of the invention. Such a vector can remain episomal or become chromosomally integrated, as long as it can be transcribed to produce the desired antisense RNA. Such vectors can be constructed by recombinant DNA technology methods standard in the art. Vectors can be plasmid, viral, or others known in the art, used for replication and expression in vertebrate cells. Expression of the sequence encoding a polypeptide of the invention, or fragments thereof, can be by any promoter known in the art to act in vertebrate, preferably human cells. Such promoters can be inducible or constitutive. Such promoters include, but are not limited to, the SV40 early promoter region (Bernoist and Chambon, Nature, 29:304-310 (1981), the promoter contained in the 3' long terminal repeat of Rous sarcoma virus (Yamamoto et al., Cell, 22:787-797 (1980), the herpes thymidine promoter (Wagner et al., Proc. Natl. Acad. Sci. U.S.A., 78:1441-1445 (1981), the regulatory sequences of the metallothionein gene (Brinster et al., Nature, 296:39-42 (1982)), etc.

[0729] The antisense nucleic acids of the invention comprise a sequence complementary to at least a portion of an RNA transcript of a gene of interest. However, absolute complementarity, although preferred, is not required. A sequence "complementary to at least a portion of an RNA," referred to herein, means a sequence having sufficient complementarity to be able to hybridize with the RNA, forming a stable duplex; in the case of double stranded antisense nucleic acids of the invention, a single strand of the duplex DNA may thus be tested, or triplex formation may be assayed. The ability to hybridize will depend on both the degree of complementarity and the length of the antisense nucleic acid Generally, the larger the hybridizing nucleic acid, the more base mismatches with a RNA sequence of the invention it may contain and still form a stable duplex (or triplex as the case may be). One skilled in the art can ascertain a tolerable degree of mismatch by use of standard procedures to determine the melting point of the hybridized complex.

[0730] Antisense oligonucleotides may be single or double stranded. Double stranded RNA's may be designed based upon the teachings of Paddison et al., Proc. Nat. Acad. Sci., 99:1443-1448 (2002); and International Publication Nos. WO 01/29058, and WO 99/32619; which are hereby incorporated herein by reference.

[0731] Oligonucleotides that are complementary to the 5' end of the message, e.g., the 5' untranslated sequence up to and including the AUG initiation codon, should work most efficiently at inhibiting translation. However, sequences complementary to the 3' untranslated sequences of mRNAs have been shown to be effective at inhibiting translation of mRNAs as well. See generally, Wagner, R., Nature, 372:333-335 (1994). Thus, oligonucleotides complementary to either the 5'- or 3'-non-translated, non-coding regions of a polynucleotide sequence of the invention could be used in an antisense approach to inhibit translation of endogenous mRNA. Oligonucleotides complementary to the 5' untranslated region of the mRNA should include the complement of the AUG start codon. Antisense oligonucleotides complementary to mRNA coding regions are less efficient inhibitors of translation but could be used in accordance with the invention. Whether designed to hybridize to the 5'-, 3'- or coding region of mRNA, antisense nucleic acids should be at least six nucleotides in length, and are preferably oligonucleotides ranging from 6 to about 50 nucleotides in length. In specific aspects the oligonucleotide is at least 10 nucleotides, at least 17 nucleotides, at least 25 nucleotides or at least 50 nucleotides.

[0732] The polynucleotides of the invention can be DNA or RNA or chimeric mixtures or derivatives or modified versions thereof, single-stranded or double-stranded. The oligonucleotide can be modified at the base moiety, sugar moiety, or phosphate backbone, for example, to improve stability of the molecule, hybridization, etc. The oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors in vivo), or agents facilitating transport across the cell membrane (see, e.g., Letsinger et al., Proc. Natl. Acad. Sci. U.S.A. 86:6553-6556 (1989); Lemaitre et al., Proc. Natl. Acad. Sci., 84:648-652 (1987); PCT Publication No: WO88/09810, published Dec. 15, 1988) or the blood-brain barrier (see, e.g., PCT Publication No: WO89/10134, published Apr. 25, 1988), hybridization-triggered cleavage agents. (See, e.g., Krol et al.,

BioTechniques, 6:958-976 (1988)) or intercalating agents. (See, e.g., Zon, Pharm. Res., 5:539-549 (1988)). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

[0733] The antisense oligonucleotide may comprise at least one modified base moiety which is selected from the group including, but not limited to, 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine.

[0734] The antisense oligonucleotide may also comprise at least one modified sugar moiety selected from the group including, but not limited to, arabinose, 2-fluoroarabinose, xylulose, and hexose.

[0735] In yet another embodiment, the antisense oligonucleotide comprises at least one modified phosphate backbone selected from the group including, but not limited to, a phosphorothioate, a phosphorodithioate, a phosphoramidothioate, a phosphoramidate, a phosphordiamidate, a methylphosphonate, an alkyl phosphotriester, and a formacetal or analog thereof.

[0736] In yet another embodiment, the antisense oligonucleotide is an a-anomeric oligonucleotide. An a-anomeric oligonucleotide forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual b-units, the strands run parallel to each other (Gautier et al., Nucl. Acids Res., 15:6625-6641 (1987)). The oligonucleotide is a 2-O-methylribonucleotide (Inoue et al., Nucl. Acids Res., 15:6131-6148 (1987)), or a chimeric RNA-DNA analogue (Inoue et al., FEBS Lett. 215:327-330 (1987)).

[0737] Polynucleotides of the invention may be synthesized by standard methods known in the art, e.g. by use of an automated DNA synthesizer (such as are commercially available from Biosearch, Applied Biosystems, etc.). As examples, phosphorothioate oligonucleotides may be synthesized by the method of Stein et al. (Nucl. Acids Res., 16:3209 (1988)), methylphosphonate oligonucleotides can be prepared by use of controlled pore glass polymer supports (Sarin et al., Proc. Natl. Acad. Sci. U.S.A., 85:7448-7451 (1988)), etc.

[0738] While antisense nucleotides complementary to the coding region sequence of the invention could be used, those complementary to the transcribed untranslated region are most preferred.

[0739] Potential antagonists according to the invention also include catalytic RNA, or a ribozyme (See, e.g., PCT International Publication WO 90/11364, published Oct. 4, 1990; Sarver et al, Science, 247:1222-1225 (1990). While ribozymes that cleave mRNA at site specific recognition sequences can be used to destroy mRNAs corresponding to the polynucleotides of the invention, the use of hammerhead ribozymes is preferred. Hammerhead ribozymes cleave mRNAs at locations dictated by flanking regions that form complementary base pairs with the target mRNA. The sole requirement is that the target mRNA have the following sequence of two bases: 5'-UG-3'. The construction and production of hammerhead ribozymes is well known in the art and is described more fully in Haseloff and Gerlach, Nature, 334:585-591 (1988). There are numerous potential hammerhead ribozyme cleavage sites within each nucleotide sequence disclosed in the sequence listing. Preferably, the ribozyme is engineered so that the cleavage recognition site is located near the 5'end of the mRNA corresponding to the polynucleotides of the invention; i.e., to increase efficiency and minimize the intracellular accumulation of non-functional mRNA transcripts.

[0740] As in the antisense approach, the ribozymes of the invention can be composed of modified oligonucleotides (e.g. for improved stability, targeting, etc.) and should be delivered to cells which express the polynucleotides of the invention in vivo. DNA constructs encoding the ribozyme may be introduced into the cell in the same manner as described above for the introduction of antisense encoding DNA. A preferred method of delivery involves using a DNA construct "encoding" the ribozyme under the control of a strong constitutive promoter, such as, for example, pol III or pol II promoter, so that transfected cells will produce sufficient quantities of the ribozyme to destroy endogenous messages and inhibit translation. Since ribozymes unlike antisense molecules, are catalytic, a lower intracellular concentration is required for efficiency.

[0741] Antagonist/agonist compounds may be employed to inhibit the cell growth and proliferation effects of the polypeptides of the present invention on neoplastic cells and tissues, i.e. stimulation of angiogenesis of tumors, and, therefore, retard or prevent abnormal cellular growth and proliferation, for example, in tumor formation or growth.

[0742] The antagonist/agonist may also be employed to prevent hyper-vascular diseases, and prevent the proliferation of epithelial lens cells after extracapsular cataract surgery. Prevention of the mitogenic activity of the polypeptides of the present invention may also be desirous in cases such as restenosis after balloon angioplasty.

[0743] The antagonist/agonist may also be employed to prevent the growth of scar tissue during wound healing.

[0744] The antagonist/agonist may also be employed to treat, prevent, and/or diagnose the diseases described herein.

[0745] Thus, the invention provides a method of treating or preventing diseases, disorders, and/or conditions, including but not limited to the diseases, disorders, and/or conditions listed throughout this application, associated with overexpression of a polynucleotide of the present invention by administering to a patient (a) an antisense molecule directed to the polynucleotide of the present invention, and/or (b) a ribozyme directed to the polynucleotide of the present invention.

Biotic Associations

[0746] A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability, either directly or indirectly, to initiate and/or maintain biotic associations with other organisms. Such associations may be symbiotic, nonsymbiotic, endosymbiotic, macrosymbiotic, and/or microsymbiotic in nature. In general, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability to form biotic associations with any member of the fungal, bacterial, lichen, mycorrhizal, cyanobacterial, dinoflaggellate, and/or algal, kingdom, phylums, families, classes, genuses, and/or species.

[0747] The mechanism by which a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the host organisms ability, either directly or indirectly, to initiate and/or maintain biotic associations is variable, though may include, modulating osmolarity to desirable levels for the symbiont, modulating pH to desirable levels for the symbiont, modulating secretions of organic acids, modulating the secretion of specific proteins, phenolic compounds, nutrients, or the increased expression of a protein required for host-biotic organisms interactions (e.g., a receptor, ligand, etc.). Additional mechanisms are known in the art and are encompassed by the invention (see, for example, "Microbial Signalling and Communication", eds., R. England, G. Hobbs, N. Bainton, and D. McL. Roberts, Cambridge University Press, Cambridge, (1999); which is hereby incorporated herein by reference).

[0748] In an alternative embodiment, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may decrease the host organisms ability to form biotic associations with another organism, either directly or indirectly. The mechanism by which a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may decrease the host organisms ability, either directly or indirectly, to initiate and/or maintain biotic associations with another organism is variable, though may include, modulating osmolarity to undesirable levels, modulating pH to undesirable levels, modulating secretions of organic acids, modulating the secretion of specific proteins, phenolic compounds, nutrients, or the decreased expression of a protein required for host-biotic organisms interactions (e.g., a receptor, ligand, etc.). Additional mechanisms are known in the art and are encompassed by the invention (see, for example, "Microbial Signalling and Communication", eds., R. England, G. Hobbs, N. Bainton, and D. McL. Roberts, Cambridge University Press, Cambridge, (1999); which is hereby incorporated herein by reference).

[0749] The hosts ability to maintain biotic associations with a particular pathogen has significant implications for the overall health and fitness of the host. For example, human hosts have symbiosis with enteric bacteria in their gastrointestinal tracts, particularly in the small and large intestine. In fact, bacteria counts in feces of the distal colon often approach 10^{12} per milliliter of feces. Examples of bowel flora in the gastrointestinal tract are members of the Enterobacteriaceae, Bacteriodes, in addition to a-hemolytic streptococci, *E. coli*, Bifobacteria, *Anaerobic cocci,* Eubacteria, Costridia, lactobacilli, and yeasts. Such bacteria, among other things, assist the host in the assimilation of nutrients by breaking down food stuffs not typically broken down by the hosts digestive system, particularly in the hosts bowel. Therefore, increasing the hosts ability to maintain such a biotic association would help assure proper nutrition for the host.

[0750] Aberrations in the enteric bacterial population of mammals, particularly humans, has been associated with the following disorders: diarrhea, ileus, chronic inflammatory disease, bowel obstruction, duodenal diverticula, biliary calculous disease, and malnutrition. A polynucleotide or polypeptide and/or agonist or antagonist of the present invention are useful for treating, detecting, diagnosing, prognosing, and/or ameliorating, either directly or indirectly, and of the above mentioned diseases and/or disorders associated with aberrant enteric flora population.

[0751] The composition of the intestinal flora, for example, is based upon a variety of factors, which include, but are not limited to, the age, race, diet, malnutrition, gastric acidity, bile salt excretion, gut motility, and immune mechanisms. As a result, the polynucleotides and polypeptides, including agonists, antagonists, and fragments thereof, may modulate the ability of a host to form biotic associations by affecting, directly or indirectly, at least one or more of these factors.

[0752] Although the predominate intestinal flora comprises anaerobic organisms, an underlying percentage represents aerobes (e.g., *E. coli*). This is significant as such aerobes rapidly become the predominate organisms in intraabdominal infections effectively becoming opportunistic early in infection pathogenesis. As a result, there is an intrinsic need to control aerobe populations, particularly for immune compromised individuals.

[0753] In a preferred embodiment, a polynucleotides and polypeptides, including agonists, antagonists, and fragments thereof, are useful for inhibiting biotic associations with specific enteric symbiont organisms in an effort to control the population of such organisms.

[0754] Biotic associations occur not only in the gastrointestinal tract, but also on an in the integument. As opposed to the gastrointestinal flora, the cutaneous flora is comprised almost equally with aerobic and anaerobic organisms. Examples of cutaneous flora are members of the gram-positive cocci (e.g., *S. aureus*, coagulase-negative staphylococci, micrococcus, *M. sedentarius*), gram-positive bacilli (e.g., Corynebacterium species, *C. minutissimum,* Brevibacterium species, Propoionibacterium species, *P.acnes*), gram-negative bacilli (e.g., Acinebacter species), and fungi (*Pityrosporum orbiculare*). The relatively low number of flora associated with the integument is based upon the inability of many organisms to adhere to the skin. The organisms referenced above have acquired this unique ability. Therefore, the polynucleotides and polypeptides of the present invention may have uses which include modulating the population of the cutaneous flora, either directly or indirectly.

[0755] Aberrations in the cutaneous flora are associated with a number of significant diseases and/or disorders, which include, but are not limited to the following: impetigo, ecthyma, blistering distal dactulitis, pustules, folliculitis, cutaneous abscesses, pitted keratolysis, trichomycosis axcillaris, dermatophytosis complex, axillary odor, erthyrasma, cheesy foot odor, acne, tinea versicolor, seborrheic

dermititis, and *Pityrosporum folliculitis,* to name a few. A polynucleotide or polypeptide and/or agonist or antagonist of the present invention are useful for treating, detecting, diagnosing, prognosing, and/or ameliorating, either directly or indirectly, and of the above mentioned diseases and/or disorders associated with aberrant cutaneous flora population.

[0756] Additional biotic associations, including diseases and disorders associated with the aberrant growth of such associations, are known in the art and are encompassed by the invention. See, for example, "Infectious Disease", Second Edition, Eds., S. L., Gorbach, J. G., Bartlett, and N. R., Blacklow, W. B. Saunders Company, Philadelphia, (1998); which is hereby incorporated herein by reference).

Pheromones

[0757] In another embodiment, a polynucleotide or polypeptide and/or agonist or antagonist of the present invention may increase the organisms ability to synthesize, release, and/or respond to a pheromone, either directly or indirectly. Such a pheromone may, for example, alter the organisms behavior and/or metabolism.

[0758] A polynucleotide or polypeptide and/or agonist or antagonist of the present invention may modulate the biosynthesis and/or release of pheromones, the organisms ability to respond to pheromones (e.g., behaviorally, and/or metabolically), and/or the organisms ability to detect pheromones, either directly or indirectly. Preferably, any of the pheromones, and/or volatiles released from the organism, or induced, by a polynucleotide or polypeptide and/or agonist or antagonist of the invention have behavioral effects on the organism.

[0759] For example, recent studies have shown that administration of picogram quantities of androstadienone, the most prominent androstene present on male human axillary hair and on the male axillary skin, to the female vomeronasal organ resulted in a significant reduction of nervousness, tension and other negative feelings in the female recipients (Grosser-BI, et al., Psychoneuroendocrinology, 25(3): 289-99 (2000)).

Other Activities

[0760] The polypeptide of the present invention, as a result of the ability to stimulate vascular endothelial cell growth, may be employed in treatment for stimulating re-vascularization of ischemic tissues due to various disease conditions such as thrombosis, arteriosclerosis, and other cardiovascular conditions. These polypeptide may also be employed to stimulate angiogenesis and limb regeneration, as discussed above.

[0761] The polypeptide may also be employed for treating wounds due to injuries, burns, post-operative tissue repair, and ulcers since they are mitogenic to various cells of different origins, such as fibroblast cells and skeletal muscle cells, and therefore, facilitate the repair or replacement of damaged or diseased tissue.

[0762] The polypeptide of the present invention may also be employed stimulate neuronal growth and to treat, prevent, and/or diagnose neuronal damage which occurs in certain neuronal disorders or neuro-degenerative conditions such as Alzheimer's disease, Parkinson's disease, and AIDS-related complex. The polypeptide of the invention may have the ability to stimulate chondrocyte growth, therefore, they may be employed to enhance bone and periodontal regeneration and aid in tissue transplants or bone grafts.

[0763] The polypeptide of the present invention may be also be employed to prevent skin aging due to sunburn by stimulating keratinocyte growth.

[0764] The polypeptide of the invention may also be employed to maintain organs before transplantation or for supporting cell culture of primary tissues.

[0765] The polypeptide of the present invention may also be employed for inducing tissue of mesodermal origin to differentiate in early embryos.

[0766] The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also increase or decrease the differentiation or proliferation of embryonic stem cells, besides, as discussed above, hematopoietic lineage.

[0767] The polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to modulate mammalian characteristics, such as body height, weight, hair color, eye color, skin, percentage of adipose tissue, pigmentation, size, and shape (e.g., cosmetic surgery). Similarly, polypeptides or polynucleotides and/or agonist or antagonists of the present invention may be used to modulate mammalian metabolism affecting catabolism, anabolism, processing, utilization, and storage of energy.

[0768] Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may be used to change a mammal's mental state or physical state by influencing biorhythms, caricadic rhythms, depression (including depressive diseases, disorders, and/or conditions), tendency for violence, tolerance for pain, reproductive capabilities (preferably by Activin or Inhibin-like activity), hormonal or endocrine levels, appetite, libido, memory, stress, or other cognitive qualities.

[0769] Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used as a food additive or preservative, such as to increase or decrease storage capabilities, fat content, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional components.

[0770] Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to increase the efficacy of a pharmaceutical composition, either directly or indirectly. Such a use may be administered in simultaneous conjunction with said pharmaceutical, or separately through either the same or different route of administration (e.g., intravenous for the polynucleotide or polypeptide of the present invention, and orally for the pharmaceutical, among others described herein.).

[0771] Polypeptide or polynucleotides and/or agonist or antagonists of the present invention may also be used to prepare individuals for extraterrestrial travel, low gravity environments, prolonged exposure to extraterrestrial radiation levels, low oxygen levels, reduction of metabolic activity, exposure to extraterrestrial pathogens, etc. Such a use may be administered either prior to an extraterrestrial event, during an extraterrestrial event, or both. Moreover, such a use may result in a number of beneficial changes in the

recipient, such as, for example, any one of the following, non-limiting, effects: an increased level of hematopoietic cells, particularly red blood cells which would aid the recipient in coping with low oxygen levels; an increased level of B-cells, T-cells, antigen presenting cells, and/or macrophages, which would aid the recipient in coping with exposure to extraterrestrial pathogens, for example; a temporary (i.e., reversible) inhibition of hematopoietic cell production which would aid the recipient in coping with exposure to extraterrestrial radiation levels; increase and/or stability of bone mass which would aid the recipient in coping with low gravity environments; and/or decreased metabolism which would effectively facilitate the recipients ability to prolong their extraterrestrial travel by any one of the following, non-limiting means: (i) aid the recipient by decreasing their basal daily energy requirements; (ii) effectively lower the level of oxidative and/or metabolic stress in recipient (i.e., to enable recipient to cope with increased extraterrestial radiation levels by decreasing the level of internal oxidative/metabolic damage acquired during normal basal energy requirements; and/or (iii) enabling recipient to subsist at a lower metabolic temperature (i.e., cryogenic, and/or sub-cryogenic environment).

[0772] Also preferred is a method of treatment of an individual in need of an increased level of a protein activity, which method comprises administering to such an individual a pharmaceutical composition comprising an amount of an isolated polypeptide, polynucleotide, or antibody of the claimed invention effective to increase the level of said protein activity in said individual.

[0773] Having generally described the invention, the same will be more readily understood by reference to the following examples, which are provided by way of illustration and are not intended as limiting.

REFERENCES

[0774] Alejandro et al., Post-ischemic injury, delayed function and NaK-ATPase distribution in the transplanted kidney, Kidney Int, 48:1308-15, 1995.

[0775] Altschul S F; Gish W; Miller W; Myers E W; Lipman D J Basic local alignment search tool. J. Mol. Biol. 215:403-10 (1990).

[0776] Arora A S et al., Hepatocellular carcinoma cells resist necrosis during anoxia by preventing phospholipase-mediated calpain activation. J Cell Physiol 167:434-42, 1996.

[0777] Baghdiguian J et al., Calpain 3 deficiency is associated with myonuclear apoptosis and profound perturbation of the IκBα/NF-κB pathway in limb-girdle muscular dystrophy type 2A, Nature Medicine 5, 503-11, 1999.

[0778] Baier et al., A calpain-10 gene polymorphism is associated with reduced muscle mRNA levels and insulin resistance. J Clin Invest. 106:R69-73, 2000.

[0779] Bartlett et al. Molecular Recognition in Chemical and Biological Problems Special Publication, Royal Chem. Soc. 78:182-196 (1989).

[0780] Benetti R, et al., The death substrate Gas2 binds m-calpain and increases susceptibility to p53-dependent apoptosis. EMBO J. Jun. 1, 2001;20(11):2702-14.

[0781] Bohm H-J, LUDI: rule-based automatic design of new substituents for enzyme inhibitor leads. J. Comp. Aid. Molec. Design 6:61-78 (1992).

[0782] Botchkarev et al. A role for p75 neurotrophil receptor in the control of apoptosis-driven hair follicle regression. FASEB J. 14,1931-42,2000.

[0783] Brau, C et al., Expression of calpain-I mRNA in human renal cell carcinoma:correlation with lymph node metastasis and histological type. *Int.J.Cancer* 84,6-9, 1999.

[0784] Cardozo T; Totrov M; Abagyan R Homology modeling by the ICM method. Proteins 23:403-14 (1995).

[0785] Carey et al., Early hypoxia induced loss of Na+/K(+)-ATPase polarity in proximal tubules is associated with specific calpain regulated fodrin processing, J.Am. Soc. Nephrol. 6:975 (abstr)1995.

[0786] Chen, F et al., Impairment of NF-kappaB activation and modulation of gene expression by calpastatin, Am J Physiol Cell Physiol. Sep;279(3):C709-16, 2000.

[0787] Cuzzocrea S et al., Calpain inhibitor I reduces colon injury caused by dinitrobenzene sulphonic acid in the rat. Gut 48(4):478-88, 2001.

[0788] Carafoli E and Molinari M. Calpain: a protease in search of a function? Biochem Biophys Res Commun 247:193-203, 1998.

[0789] Cuzzocrea S et al., Calpain inhibitor I reduces the development of acute and chronic inflammation Am J Pathol 157:2065-79, 2000.

[0790] Dear et al., Gene structure, chromosomal localization, and expression pattern of Capn12, a new member of the calpain large subunit gene family, Genomics 68(2):152-60, 2000.

[0791] Edelstein et al., Modulation of hypoxia-induced calpain activity in rat renal proximal tubules, Kidney Int 50; 1150-57, 1996.

[0792] Fox J E, On the role of calpain and Rho proteins in regulating integrin-induced signaling. Thromb Haemost 82:385-91, 1999.

[0793] Goodford, P. J. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J. Med. Chem. 28:849-857 (1985).

[0794] Goodsell, D. S. and Olsen, A. J. Automated docking of substrates to proteins by simulated annealing. Proteins 8:195-202 (1990).

[0795] Greer J Comparative modeling of homologous proteins. Methods Enzymol 202:239-52 (1991).

[0796] Hayes et al., Drug News Perspect 11:215-222, 1998.

[0797] Hendlich M; Lackner P; Weitckus S; Floeckner H; Froschauer R; Gottsbacher K; Casari G; Sippl M J Identification of native protein folds amongst a large number of incorrect models. The calculation of low

energy conformations from potentials of mean force. J. Mol. Biol. 216:167-80 (1990).

[0798] Horikawa et al., Genetic variation in the gene encoding calpain-10 is associated with type 2 diabetes mellitus. Nat Genet. 26:163-75, 2000.

[0799] Hosfield, C., Elce, J. S., Davies, P., Jia, Z., Crystal structure of calpain reveals the structural basis for Ca^{2+}-dependent protease activity and a novel mode of enzyme activation., EMBO J., 18, 6880-6889 (1999).

[0800] Iwamoto H et al., Calpain inhibitor-1 reduces infarct size and DNA fragmentation of myocardium in ischemic/reperfused rat heart.J Cardiovasc Pharmacol 33:580-6, 1999.

[0801] Kuntz I D, Blaney J M, Oatley S J, Langridge R, Ferrin T E. A geometric approach to macromolecule-ligand interactions. J. Mol. Biol. 161:269-288 (1982).

[0802] Lee, K S, et al., Calcium-activated proteolysis as a therapeutic target in cerebrovascular disease. Annal NY Acad. Sci. 825, 95-103, 1997.

[0803] Lee et al., Proc. Natl. Acad. Sci. USA, 88:7233-7237, 1991.

[0804] Lesk, A. M., Boswell, D. R., Homology Modeling: Inferences from Tables of Aligned Sequences. Curr. Op. Struc. Biol. 2:242-247 (1992).

[0805] Liu, K et al., Antisense RNA mediated deficiency of the calpain protease nCL-4, in NIH3T3 cells is associated with neoplastic transformation and tumorigenesis, JBC 275, 31093-98, 2000.

[0806] Martin, Y. C. 3D database searching in drug design. J. Med. Chem. 35:2145-2154 (1992).

[0807] McDonald, M C et al., Calpain inhibitor I reduces the activation of nuclear factor-B and organ injury/dysfunction in hemorrhagic shock, FASEB 15,171-86, 2001.

[0808] Pearson W R Rapid and sensitive sequence comparison with FASTP and FASTA. Methods Enzymol 183:63-98 (1990).

[0809] Rami et al., Brain Research, 609:67-70, 1993.

[0810] Rawlings N. D., & Barrett A. J. Families of cysteine peptidases. Methods in Enzymol. 244 461-486 (1994).

[0811] Richard et al., Mutations in the proteolytic enzyme calpain 3 cause limb-girdle muscular dystrophy type 2A. Cell 81:27-40. 1995.

[0812] Richard I, et al., Loss of calpain 3 proteolytic activity leads to muscular dystrophy and to apoptosis-associated IkappaBalpha/nuclear factor kappaB pathway perturbation in mice, J Cell Biol 151(7):1583-90, 2000.

[0813] Sali A; Potterton L; Yuan F; van Vlijmen H; Karplus M Evaluation of comparative protein modeling by MODELLER. PROTEINS 23:318-26 (1995).

[0814] Sorimachi and Suzuki, The Structure of calpains. J. Biochem 129:653-664, 2001.

[0815] Sorimachi et al., Structure and physiological function of calpains. Biochem J. 328:721-32, 1997.

[0816] Strobl, S., Fernandez-Catalan, C., Braun, M., Huber, R., Masumoto, H., Nakagawa., K., Irie, A., Sorimachi, S., Bourenkow, G., Bartunik, H., Suzuki, K., Bode, W., The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. Proc. Natl. Acad. Sci., 97(2), 588-592, 2000.

[0817] Tompa P et al., Domain III of calpain is a Ca-regulated phospholipid-binding domain, BBRC 280,1333-9, 2001.

[0818] Wang, K K, Calpain and caspase:can you tell the difference?*Trends Neurosci.* 23;20-6, 2000.

[0819] Wang et al., An alpha-mercaptoacrylic acid derivative is a selective nonpeptide cell-permeable calpain inhibitor and is neuroprotective. Proc Natl Acad Sci USA. 93:6687-92, 1996.

[0820] Wang K K and Yuen P W. Calpain inhibition: an overview of its therapeutic potential. Trends Pharmacol. Sci. 15:412-9, 1994.

[0821] Wilson S J et al., Immunohistochemical analysis of the activation of NF-kappa B and expression of associated cytokines and adhesion molecules in human models of allergic inflammation, J. Pathol 189: 265-72, 1999.

[0822] Yoshikawa,Y et al., Isolation of two novel genes, down regulated in gastric cancer, Jpn.J.Cancer.Res 91,459-463, 2000.

EXAMPLES

Description of the Preferred Embodiments

Example 1

Bioinformatics Analysis

[0823] To search for novel cysteine protease inhibitors, a Hidden-Markov Model Peptidase_C2.hmm of cysteine proteases (obtained from the Pfam database in Sanger center) (Bateman et. al., 2000) was used to search against human genomic sequence database using the computer program GENEWISEDB. Genomic sequences that were found to have a GENEWISEDB matching score of more than 15 against Peptidase_C2.hmm HMM model were selected for further analysis. The genomic sequence contained in the sequence assembly NT_011222 was found to contain putative exon sequences that are similar to calpains. The portion of the sequence from NT_011222 that matched the Peptidase_C2.hmm HMM profile was extracted and back-searched against the non-redundant protein database using the BLASTX program (Altschul et. al., 1990). The most similar protein sequence identified was the mouse capn12 protein (Genbank Accession No. gil10303329; SEQ ID NO:16), and was used as a template to predict more exons from NT_011222 using the GENEWISEDB program (Birney and Durbin, 2000). The final predicted exons were assembled to provide the full-length clone Protease-42 (as shown in FIGS. **1**A-C and SEQ ID NO:2).

[0824] The complete protein sequence of Protease-42 was found to have significant sequence homology with a family of known protease inhibitors (see FIGS. 2A-H). The sequence of Protease-42 around the catalytic cysteine residue is well conserved and fits well with the conserved signature patterns of thiol proteases PS00139: Q-x(3)-[GE]-x-C-[YW]-x(2)-[STAGC]-[STAGCV][wherein C is the active site residue], and PS00639 [LIVMGSTAN]-x-H-[GSACE]-[LIVM]-x-[LIVMAT](2)-G-x-[GSADNH] [wherein H is the active residue].

[0825] Protein threading and molecular modeling of Protease-42 suggest that Protease-42 has a structural fold characteristic of cysteine protease inhibitors (see **FIGS. 6 and 7** herein). Based on sequence, structure, and the presence of known cysteine protease signature sequences, the novel Protease-42 has been determined to represent a novel human calpain.

Example 2

Method for Constructing a Size Fractionated Brain and Testis cDNA Library

[0826] Brain and testis poly A+RNA was purchased from Clontech and converted into double stranded cDNA using the SuperScript™ Plasmid System for cDNA Synthesis and Plasmid Cloning (Life Technologies) except that no radioisotope was incorporated in either of the cDNA synthesis steps and that the cDNA was fractionated by HPLC. This was accomplished on a TransGenomics HPLC system equipped with a size exclusion column (TosoHass) with dimensions of 7.8 mm×30 cm and a particle size of 10 nm. Tris buffered saline was used as the mobile phase and the column was run at a flow rate of 0.5 mL/min.

[0827] The resulting chromatograms were analyzed to determine which fractions should be pooled to obtain the largest cDNA's; generally fractions that eluted in the range of 12 to 15 minutes were pooled. The cDNA was precipitated prior to ligation into the Sal I/Not I sites in the pSport vector supplied with the kit. Using a combination of PCR with primers to the ends of the vector and Sal I/Not I restriction enzyme digestion of mini-prep DNA, it was determined that the average insert size of the library was greater the 3.5 Kb. The overall complexity of the library was greater that 10^7 independent clones. The library was amplified in semi-solid agar for 2 days at 30° C. An aliquot (200 microliters) of the amplified library was inoculated into a 200 ml culture for single-stranded DNA isolation by superinfection with a f1 helper phage. After overnight growth, the released phage particles with precipitated with PEG and the DNA isolated with proteinase K, SDS and phenol extractions. The single stranded circular DNA was concentrated by ethanol precipitation and used for the cDNA capture experiments.

Example 3

Cloning of the Novel Human Calpain, Protease-42

[0828] A typical RT-PCR method was used to clone the cDNA. The following is a detailed description of the procedures used.

[0829] First, PCR oligonucleotide primers were designed that flanked the predicted open reading frame based upon the sequence provided as SEQ ID NO:1 (FIGS. 1A-C). The software used for this was Primer3, written by Steve Rozen and others from the Whitehead Institute at MIT (Steve Rozen, Helen J. Skaletsky (1998) Primer3. Code available at http://www-genome.wi.mit.edu/genome_software/other/primer3.html). The predicted product size of the amplification product using these primers was 2032.

[0830] The resulting oligonucleotide sequences were as follows:

```
LEFT PRIMER    AGATGGCATCCAGCAGTG    (SEQ ID NO:59)

RIGHT PRIMER   TCCGGAGATCCTAGGAGAA   (SEQ ID NO:60)
```

[0831] Next, these oligonucleotide primers were used to amplify the target cDNA by the polymerase chain reaction (PCR). The template for the reaction was a pool of 1st strand cDNA's synthesized from human poly A+ RNA (see Example 2 and below).

[0832] The first strand cDNA was synthesized using the SuperScript™ Preamplification System for First Strand cDNA Synthesis kit from Gibco BRL®. The following was added to the reaction mix:

[0833] 2.5 μg of poly A+ human brain RNA

[0834] 50 ng random hexamers

[0835] water to 12 μl

[0836] This was incubated at 70° C. for 10 minutes. Then incubated on ice for 1 minute. Then the following reaction was set up:

[0837] all 12 μL of the RNA/primer mixture

[0838] 2 μL of 10×PCR buffer

[0839] 2 μL 25 mM $MgCl_2$

[0840] 1 μL 10 mM dNTP mix

[0841] 2 μL 0.1 M DTT

[0842] This was incubated at 25° C. for 5 minutes. Then 1 μL of SuperScript™ II reverse transcriptase was added. This was incubated at 25° C. for another 10 minutes. Then it was transferred to 42° C. for 50 minutes. The reaction was terminated by heating at 70° C. for 15 minutes, then placing on ice. Following this, 1 μL of RNase H was added to degrade the remaining RNA, and was then incubated for 20 minutes at 37° C.

[0843] Then PCR was carried out using PCR SuperMix High Fidelity reagent from GibcoBRL®. The composition of the reagent was as follows:

[0844] recombinant Taq polymerase

[0845] DNA polymerase from Pyrococus species GB-D

[0846] 66 mM Tris-SO_4 (pH 9.1)

[0847] 19.8 mM $(NH_4)_2SO_4$

[0848] **2.2** mM $MgSO_4$

[0849] **220**μM each dNTP (dGTP, dATP, dTTP, dCTP)

[0850] proprietary stabilizers

[0851] 47 μL of the reagent were added per reaction. 5 ng of DNA template were added to the reaction mixture along with each oligonucleotide primer at a final concentration of 0.2 μM each. The total volume of the reactions was 50 μL.

[0852] The thermal cycling conditions for the PCR were as follows:

[0853] 95° C. 3 minutes

[0854] Then 45 cycles of:

[0855] 95° C. 20 seconds

[0856] 55° C. 20 seconds

[0857] 72° C. 2 minutes

[0858] Then one cycle of:

[0859] 72° C. 10 minutes

[0860] 4° C. hold

[0861] The resulting PCR products were separated by electrophoresis on a 1% agarose gel. There was a band visualized of the correct size (~2 Kb). The band was excised from the gel with a razor blade.

[0862] The PCR product was then extracted from the agarose gel slice using the Qiagen QIAquick™ Gel Extraction kit. Briefly, 3 volumes of buffer QG were added to the gel slice. The mixture was incubated at 50° C. until the agarose was melted. Then one volume of isopropanol was added. The sample was applied to a QIAquick spin column and centrifuged for 1 minute at high speed. The bound DNA was washed on the column by applying 750 μL of buffer PE to and centrifuging for 1 minute. The column was then dried by spinning for an additional minute at high speed. The DNA was eluted from the column by applying 30 μL of elution buffer (buffer EB), allowing the column to stand for 1 minute, then centrifuging the column at high speed for 1 minute. The eluate was collected in a microcentrifuge tube.

[0863] Next, a 'TA' cloning procedure was used to insert the amplified fragment into a plasmid vector. In order to use the 'TA' cloning strategy, the PCR amplicon must have a 3''A' overhang which is generated by Taq polymerase. Since a high fidelity, proofreading enzyme was used for the PCR amplification, the proofreading properties of the enzyme mix caused the 'A' overhang to be removed. Therefore, before the 'TA' cloning could be done, 'A' overhangs had to be added to the PCR product. To do this, the PCR product was incubated for 15 minutes at 72° C. in a mixture containing 5 units of Taq polymerase, 1×PCR buffer and 0.2 mM dATP (all from Roche). The Taq polymerase was from *Thermus aquaticus* BM, recombinant *E. coli*. The 10×PCR buffer contained 100 mM Tris-HCl, 15 mM $MgCl_2$, 500 mM KCl, pH 8.3.

[0864] The PCR products with added 3''A' overhangs was then immediately used for 'TA' cloning. To do this, the pGEM-T-Easy® Kit for Sequencing from Promega was used.

[0865] The following reaction mixture was used:

[0866] 5 μL 2× rapid ligation buffer

[0867] 1 μL pGEM-T-Easy vector (50 ng)

[0868] 3 μL PCR product

[0869] 1 μL T4 DNA Ligase

[0870] This was incubated at room temperature for 1 hour.

[0871] The ligation reaction was diluted 5 fold in water. Then 2 μL of this reaction was transformed by electroporation into DH10B electromax electrocompetant *E. Coli* bacteria (Gibco BRL) followng the manufacturer's instructions. The cells were then incubated at 37° C. for 1 hour in S.O.C. media with shaking for aeration. 1 μL of cells were spread on selective plates containing 50 μg/μL carbenicillin and incubated at 37° C. overnight.

[0872] The next step was to screen colonies that grew on the selective plates for positive clones. This was done by growing colonies overnight in 1.2 mL of LB broth containing 50 μg/μL carbenicillin. The plasmid DNA was then isolated from the bacteria using the Qiagen QIAquick Spin Miniprep Kit. Protocols for this are available from the Qiagen company web site (http://www.qiagen.com).

[0873] Once the plasmid DNA was purified, a PCR assay using internal primers was performed to determine if the clones were correct.

Internal PCR Primer Sequences

[0874]

LEFT PRIMER	CAGAGCTATGAGGCAATTCG	(SEQ ID NO:17)
RIGHT PRIMER	TCATCCATTTCACGCCTTT	(SEQ ID NO:18)

[0875] Four clones tested gave an amplicon of the correct size (142 bp) which indicated that they contained the correct insert. The clone was sequenced using Applied Biosystems BigDye™ dideoxy terminator cycle sequencing on an Applied Biosystems 3700 capillary array DNA sequencer.

[0876] The full-length nucleotide sequence Protease-42 is shown in FIGS. **1**A-C (SEQ ID NO:1) and the conceptual translation of the full-length cDNA sequence for Protease-42 is shown in FIGS. **1**A-C (SEQ ID NO:2). The sequence was analyzed and plotted in a hydrophobicity plot using the BioPlot Hydrophobicity algorithm within Vector NTI (version 5.5). The plot showed no detectable transmembrane domain or signal sequence at the NH_2 terminus (**FIG. 3**).

Example 4

Expression Profiling of the Novel Human Calpain Protease-42

[0877] The same PCR primer pair (SEQ ID NO:17 and 18) that was used to identify the Protease-42 cDNA clones was used to measure the steady state levels of mRNA by quantitative PCR. Briefly, first strand cDNA was made from commercially available mRNA (Clontech) and subjected to real time quantitative PCR using a PE 5700 instrument (Applied Biosystems, Foster City, Calif.) which detects the amount of DNA amplified during each cycle by the fluorescent output of SYBR green, a DNA binding dye specific for double strands. The specificity of the primer pair for its target is verified by performing a thermal denaturation profile at the end of the run which gives an indication of the number of different DNA sequences present by determining melting Tm. In the case of the Protease-42 primer pair, only

one DNA fragment was detected having a homogeneous melting point. Contributions of contaminating genomic DNA to the assessment of tissue abundance is controlled for by performing the PCR with first strand made with and without reverse transcriptase. In all cases, the contribution of material amplified in the no reverse transcriptase controls was negligible.

[0878] Small variations in the amount of cDNA used in each tube was determined by performing a parallel experiment using a primer pair for a gene expressed in equal amounts in all tissues, cyclophilin. These data were used to normalize the data obtained with the Protease-42 primer pair. The PCR data was converted into a relative assessment of the difference in transcript abundance amongst the tissues tested and the data are presented in bar graph form. Transcripts corresponding to Protease-42 were expressed highly in the brain, liver, kidney, spleen, and to a lesser extent, in other tissues as shown in **FIG. 4**. The mouse ortholog of Protease-42, the mouse Can 12, is expressed at high level in skin (Dear et al., 2000).

[0879] The tissue distribution of Protease-42 and our current knowledge of calpain function suggests that the Protease-42 protein (SEQ ID NO:2) could be useful in the treatment of ischemia-reperfusion injury or tumorigenesis processes in brain, liver, spleen, lung, kidney, and in the digestive track. Also, identification of endogenous substrate (s) of Protease-42 might help to define underlying mechanisms in hair proliferation and differentiation and lead to the development of novel drug target for the treatment of of alopecia.

Example 5

Method of Assessing the Expression Profile of the Novel Protease-42 Polypeptide of the Present Invention Using Expanded mRNA Tissue and Cell Sources

[0880] Total RNA from tissues was isolated using the TriZol protocol (Invitrogen) and quantified by determining its absorbance at 260 nM. An assessment of the 18s and 28s ribosomal RNA bands was made by denaturing gel electrophoresis to determine RNA integrity.

[0881] The specific sequence to be measured was aligned with related genes found in GenBank to identity regions of significant sequence divergence to maximize primer and probe specificity. Gene-specific primers and probes were designed using the ABI primer express software to amplify small amplicons (150 base pairs or less) to maximize the likelihood that the primers function at 100% efficiency. All primer/probe sequences were searched against Public Genbank databases to ensure target specificity. Primers and probes were obtained from ABI.

[0882] For Protease-42, the primer probe sequences were as follows

```
Forward Primer
5'-GCACGTCCACACCTTCCAA-3'        (SEQ ID NO:55)

Reverse Primer
5'-TTGGTCCAGAAGGTTTCAGCAT-3'     (SEQ ID NO:56)
```

-continued

```
TaqMan Probe
5'-TCCGGCGGGAGCCAGCCT-3'         (SEQ ID NO:57)
```

DNA Contamination

[0883] To access the level of contaminating genomic DNA in the RNA, the RNA was divided into 2 aliquots and one half was treated with Rnase-free Dnase (Invitrogen). Samples from both the Dnase-treated and non-treated were then subjected to reverse transcription reactions with (RT+) and without (RT−) the presence of reverse transcriptase. TaqMan assays were carried out with gene-specific primers (see above) and the contribution of genomic DNA to the signal detected was evaluated by comparing the threshold cycles obtained with the RT+/RT− non-Dnase treated RNA to that on the RT+/RT− Dnase treated RNA. The amount of signal contributed by genomic DNA in the Dnased RT− RNA must be less that 10% of that obtained with Dnased RT+RNA. If not the RNA was not used in actual experiments.

Reverse Transcription Reaction and Sequence Detection

[0884] 100 ng of Dnase-treated total RNA was annealed to 2.5 μM of the respective gene-specific reverse primer in the presence of 5.5 mM Magnesium Chloride by heating the sample to 72° C. for 2 min and then cooling to 55° C. for 30 min. 1.25 U/l of MuLv reverse transcriptase and 500 μM of each dNTP was added to the reaction and the tube was incubated at 37° C. for 30 min. The sample was then heated to 90° C. for 5 min to denature enzyme.

[0885] Quantitative sequence detection was carried out on an ABI PRISM 7700 by adding to the reverse transcribed reaction 2.5 μM forward and reverse primers, 500 μM of each dNTP, buffer and 5U AmpliTaq Gold™. The PCR reaction was then held at 94° C. for 12 min, followed by 40 cycles of 94° C. for 15 sec and 60° C. for 30 sec.

Data Handling

[0886] The threshold cycle (Ct) of the lowest expressing tissue (the highest Ct value) was used as the baseline of expression and all other tissues were expressed as the relative abundance to that tissue by calculating the difference in Ct value between the baseline and the other tissues and using it as the exponent in $2^{(\Delta Ct)}$

[0887] The expanded expression profile of Protease-42 is provided in **FIG. 9** and is described elsewhere herein.

Example 6

Method of Measuring the Protease Activity of Protease-42 Polypeptides

[0888] Protease activity of the Protease-42 polypeptide are measured by following the inhibition of proteolytic activity in cells, tissues, and/or in in vitro assays. Cysteine proteases of the calpain family (of which the present invention is a member) catalyze the hydrolysis of peptide, amide, ester, thiol ester and thiono ester bonds. Any assay that measures cleavage of these bonds can be used to quantitate enzymatic activity. In vitro assays for measuring protease activity using

synthetic peptide fluorescent, spectrophotometric either through the use of single substrates (see below for examples), and fluorescence resonance transfer assays are well described in the art, as single substrates or as part of substrate libraries (Backes et al., 2000; Knight, C. G. Fluorimetric Assays of Proteolytic Enzymes. Meth. Enzymol. 248: 18-34 (1995)). In addition proteolytic activity is measured by following production of peptide products. Such approaches are well known to those familiar with the art (reviewed in McGeehan, G. M., Bickett, D. M., Wiseman, J. S., Green, M., Berman, Meth. Enzymol. 248: 35-46 (1995))

[0889] A complete set of protocols that have been used to evaluate calpain activity and are provided in Calpain Methods and Protocols John Elce ed. In Meth. Mol. Biol. Volume 144, 2000 (Humana Press, Totowa, N.J.).

Inhibitor Identification

[0890] Early work on calpain inhibitors produced nonselective enzyme inhibitors. Peptidyl aldehydes such as leupeptin and antipain inhibit calpain but also other proteases including serine proteases. Irreversible inhibitors such as the E64 family have also been studied, and peptidyl halomethanes and diazomethanes have long been used as protease inhibitors (Hayes et al., Drug News Perspect 11:215-222, 1998). Given the multiple therapeutic indications for the inhibition of calpain it appears that the achievement of selective modulators including specific inhibitors of this enzyme is an important goal.

[0891] The Protease-42 may be incubated with potential inhibitors (preferably small molecule inhibitors or antibodies provided elsewhere herein) for different times and with varying concentrations. Residual protease activity could then be measured according to any appropriate means known in the art. Enzyme activity in the presence of control may be expressed as fraction of control and curve fit to pre-incubation time and serpin concentration to determine inhibitory parameters including concentration that half maximally inhibits the enzyme activity.

[0892] Non-limiting examples of in vitro protease assays are well described in the art. Non-limiting examples of a spectrophotometric protease assays are the thrombin and tryptase assays measuring time-dependent optical density change followed at 405 nm using a kinetic microplate reader (Molecular Devices UVmax)(Balasubramanian, et al., Active site-directed synthetic thrombin inhibitors: synthesis, in vitro and in vivo activity profile of BMY 44621 and analogs an examination of the role of the amino group in the D-Phe-Pro-Arg-H series. J. Med. Chem. 36:300-303 (1993); and Combrink et al., Novel 1,2-Benzisothiazol-3-one-1,1-dioxide Inhibitors of Human Mast Cell Tryptase. J. Med. Chem. 41:4854-4860 (1998)).

[0893] An example of a fluorescence assay which may be used for the present invention is the Factor VIIa assay. Briefly, the Factor VIIa assay is measured in the presence of human recombinant tissue factor (INNOVIN from Dade Behring Cat.# B4212-100). Human Factor VIIa may be obtained from Enzyme Research Labs (Cat.# HFVIIA 1640). Enzymatic activity could be measured in a buffer containing 150 mM NaCl, 5 mM $CaCl_2$, 1 mM CHAPS and 1 mg/ml PEG 6000 (pH 7.4) with 1 nM FVIIa and 100 μM D-Ile-Pro-Arg-AFC (Enzyme Systems Products, Km>200 μM) 0.66% DMSO. The assay (302 μl total volume) may be incubated at room temperature for 2 hr prior to reading fluorometric signal (Ex 405/Em 535) using a Molecular Devices or Victor 2 (Wallac) fluorescent plate reader.

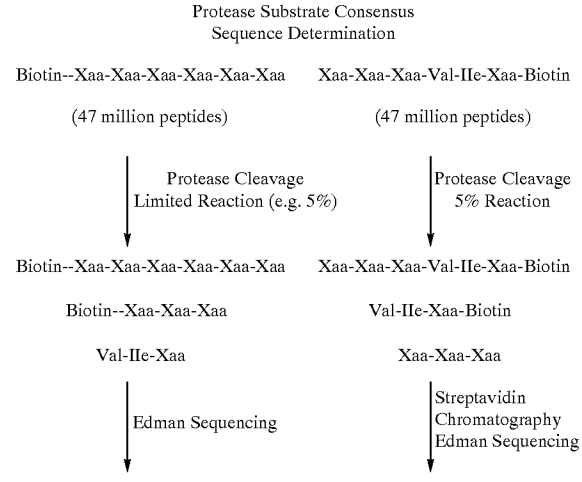
[0894] In addition to the methods described above, protease activity (and therefore serpin activity) can be measured using fluorescent resonance energy transfer (FRET with Quencher-P_n-P_3-P_2-P_1- -P_1'-P_2'-Fluorophore), fluorescent peptide bound to beads (Fluorophore-P_n-P_3-P_2-P_1- -P_1'-P_2'-Bead), dye-protein substrates and serpin-protease gel shifts. All of which are well known to those skilled in the art (see a non-limiting review in Knight, C. G. Fluorimetric Assays of Proteolytic Enzymes. Meth. Enzymol. 248: 18-34 (1995)).

[0895] Additional assay methods are known in the art and are encompassed by the present invention. See, for example, Backes B J, Harris J L, Leonetti F, Craik C S, Ellman J A. Synthesis of positional-scanning libraries of fluorogenic peptide substrates to define the extended substrate specificity of plasmin and thrombin. Nat Biotechnol. 18:187-93 (2000); Balasubramanian, et al., Active site-directed synthetic thrombin inhibitors: synthesis, in vitro and in vivo activity profile of BMY 44621 and analogs an examination of the role of the amino group in the D-Phe-Pro-Arg-H series. J. Med. Chem. 36:300-303 (1993); and Combrink et al., Novel 1,2-Benzisothiazol-3-one-1,1-dioxide Inhibitors of Human Mast Cell Tryptase. J. Med. Chem. 41:4854-4860 (1998) and those methods described in: Calpain Methods and Protocols (ed J. S. Elce) Meth. Mol. Biol. 144, 2000 and Calpain: Pharmacology and Toxicology of a calcium-dependent protease (K. Wang & P.-W. Yuen editors) Taylor & Francis Philadelphia, Pa., 1999; which are hereby incorporated herein by reference in their entirety.

Example 7

Determination of the Preferred Substrate Sequence of the Novel Calpain, Protease-42

[0896] The preferred substrate sequence specificity of the Protease-42 calpain of the present invention may be determined using using two redundant peptide libraries and Edman peptide sequencing (1-2) as shown below.



-continued

P1′-P2′-P3′ = Val-IIe-Xaa	P3-P2-P1 = Xaa-Xaa-Xaa

[0897] The first peptide library is random, can vary in length and incorporates a modification at the N-terminus to block Edman sequencing. In the example provided, biotin is used as the blocking group. Proteolytic cleavage of the library is allowed to proceed long enough to turn over approximately 5-10% of the library. Edman sequencing of the peptide mixture provides the preferred substrate residues for the P' sites on the protease. The second peptide library has fixed P' residues to restrict the proteolytic cleavage site and an affinity tag for removing the C-terminal product of the proteolysis, leaving the N-terminal peptide product pool behind for Edman sequencing to determine the amino acid residues preferred in the P1, P2, P3 etc. . . . sites of the protease.

Reagents

[0898] The endoproteases Factor Xa (New England BioLabs, Inc., Beverly, Mass.) and human kidney Renin (Calbiochem, San Diego, Calif.) were purchased for validation experiments. A hexapeptide library containing 4.7×10^7 peptide species was synthesized by the Molecular Redesign group (Natarajan & Riexinger) at Bristol-Myers Squibb Company (Princeton, N.J.). The library contained equivalent representation of 19 amino acid residues at each of the six degenerate positions and incorporated an N-terminal biotin group and a C-terminal amide. Cysteine residues were excluded from the peptide pool and Methionine residues were replaced with Norleucine.

Endoprotease Cleavage of the Peptide Library

[0899] The following method may be used to determine the preferred substrate sequence downstream of the cleavage site. A 1.88 mM peptide library solution is prepared in phosphate buffer (10 mM Sodium Phosphate (pH 7.6), 0.1 M NaCl, and 10% DMSO) and is incubated with 2-30 μg endoprotease at 37° C. Using a fluorescamine assay to estimate the extent of peptide cleavage, the reaction is stopped at 5-10% completion with incubation at 100° C. for 2.0 minutes. Peptide pools are subjected to Edman sequencing. The data obtained is normalized and corrected for differences in efficiency of cleavage and recovery in the sequencer.

Fluorescamine Assay to Monitor Peptide Cleavage

[0900] Primary amines generated during peptide cleavage is measured by reaction with fluorescamine (Aldrich, St. Louis, Mo.), as described in reference 3. The relative fluorescence is determined by measuring signals at λ^{ex}=355 nm and λ^{em}=460 nm on a PerkinElmer Wallac 1420 Spectrofluorometer. Reactions are sampled at multiple time points and assayed in triplicate. The amount of cleavage product formed is determined using the relative fluorescence produced by varying concentrations of a peptide standard of known concentration.

REFERENCES

[0901] (1) "Substrate Specificity of Cathepsins D and E Determined by N-Terminal and C-Terminal Sequencing of Peptide Pools" D. Arnold et al. (1997) *Eur. J. Biochem.* 249, 171.

[0902] (2) "Determination of Protease Cleavage Site Motifs Using Mixture-Based Oriented Peptide Libraries" B. E. Turk et al. (2001) *Nature Biotech.* 19, 661.

[0903] (3) "Fluorescamine: a Reagent for Assay of Amino Acids, Peptides, Proteins, and Primary Amines in the Picomole Range" S. Udenfirend, S. Stein, P. Bohlen, W. Dairman, W. Leimgruber, and M. Weigele (1972) *Science* 178, 87.

Example 8

Method of Screening for Compounds that Interact with the Protease-42 Polypeptide

[0904] The following assays are designed to identify compounds that bind to the Protease-42 polypeptide, bind to other cellular proteins that interact with the Protease-42 polypeptide, and to compounds that interfere with the interaction of the Protease-42 polypeptide with other cellular proteins.

[0905] Such compounds can include, but are not limited to, other cellular proteins. Specifically, such compounds can include, but are not limited to, peptides, such as, for example, soluble peptides, including, but not limited to Ig-tailed fusion peptides, comprising extracellular portions of Protease-42 polypeptide transmembrane receptors, and members of random peptide libraries (see, e.g., Lam, K. S. et al., 1991, Nature 354:82-84; Houghton, R. et al., 1991, Nature 354:84-86), made of D-and/or L-configuration amino acids, phosphopeptides (including, but not limited to, members of random or partially degenerate phosphopeptide libraries; see, e.g., Songyang, Z., et al., 1993, Cell 72:767-778), antibodies (including, but not limited to, polyclonal, monoclonal, humanized, anti-idiotypic, chimeric or single chain antibodies, and FAb, F(ab').sub.2 and FAb expression libary fragments, and epitope-binding fragments thereof), and small organic or inorganic molecules.

[0906] Compounds identified via assays such as those described herein can be useful, for example, in elaborating the biological function of the Protease-42 polypeptide, and for ameliorating symptoms of tumor progression, for example. In instances, for example, whereby a tumor progression state or disorder results from a lower overall level of Protease-42 expression, Protease-42 polypeptide, and/or Protease-42 polypeptide activity in a cell involved in the tumor progression state or disorder, compounds that interact with the Protease-42 polypeptide can include ones which accentuate or amplify the activity of the bound Protease-42 polypeptide. Such compounds would bring about an effective increase in the level of Protease-42 polypeptide activity, thus ameliorating symptoms of the tumor progression disorder or state. In instances whereby mutations within the Protease-42 polypeptide cause aberrant Protease-42 polypeptides to be made which have a deleterious effect that leads to tumor progression, compounds that bind Protease-42 polypeptide can be identified that inhibit the activity of the bound Protease-42 polypeptide. Assays for testing the effectiveness of such compounds are known in the art and discussed, elsewhere herein.

Example 9

Method of Screening, In Vitro, Compounds that Bind to the Protease-42 Polypeptide

[0907] In vitro systems can be designed to identify compounds capable of binding the Protease-42 polypeptide of

the invention. Compounds identified can be useful, for example, in modulating the activity of wild type and/or mutant Protease-42 polypeptide, preferably mutant Protease-42 polypeptide, can be useful in elaborating the biological function of the Protease-42 polypeptide, can be utilized in screens for identifying compounds that disrupt normal Protease-42 polypeptide interactions, or can in themselves disrupt such interactions.

[0908] The principle of the assays used to identify compounds that bind to the Protease-42 polypeptide involves preparing a reaction mixture of the Protease-42 polypeptide and the test compound under conditions and for a time sufficient to allow the two components to interact and bind, thus forming a complex which can be removed and/or detected in the reaction mixture. These assays can be conducted in a variety of ways. For example, one method to conduct such an assay would involve anchoring Protease-42 polypeptide or the test substance onto a solid phase and detecting Protease-42 polypeptide/test compound complexes anchored on the solid phase at the end of the reaction. In one embodiment of such a method, the Protease-42 polypeptide can be anchored onto a solid surface, and the test compound, which is not anchored, can be labeled, either directly or indirectly.

[0909] In practice, microtitre plates can conveniently be utilized as the solid phase. The anchored component can be immobilized by non-covalent or covalent attachments. Non-covalent attachment can be accomplished by simply coating the solid surface with a solution of the protein and drying. Alternatively, an immobilized antibody, preferably a monoclonal antibody, specific for the protein to be immobilized can be used to anchor the protein to the solid surface. The surfaces can be prepared in advance and stored.

[0910] In order to conduct the assay, the nonimmobilized component is added to the coated surface containing the anchored component. After the reaction is complete, unreacted components are removed (e.g., by washing) under conditions such that any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the previously immobilized component is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the previously nonimmobilized component is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the immobilized component (the antibody, in turn, can be directly labeled or indirectly labeled with a labeled anti-Ig antibody).

[0911] Alternatively, a reaction can be conducted in a liquid phase, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for Protease-42 polypeptide or the test compound to anchor any complexes formed in solution, and a labeled antibody specific for the other component of the possible complex to detect anchored complexes.

Example 10

Method of Identifying Compounds that Interfere with Protease-42 Polypeptide/Cellular Product Interaction

[0912] The Protease-42 polypeptide of the invention can, in vivo, interact with one or more cellular or extracellular macromolecules, such as proteins. Such macromolecules include, but are not limited to, nucleic acid molecules and those products identified via methods such as those described, elsewhere herein. For the purposes of this discussion, such cellular and extracellular macromolecules are referred to herein as "binding partner(s)". For the purpose of the present invention, "binding partner" may also encompass small molecule compounds, polysaccarides, lipids, and any other molecule or molecule type referenced herein. Compounds that disrupt such interactions can be useful in regulating the activity of the Protease-42 polypeptide, especially mutant Protease-42 polypeptide. Such compounds can include, but are not limited to molecules such as antibodies, peptides, and the like described in elsewhere herein.

[0913] The basic principle of the assay systems used to identify compounds that interfere with the interaction between the Protease-42 polypeptide and its cellular or extracellular binding partner or partners involves preparing a reaction mixture containing the Protease-42 polypeptide, and the binding partner under conditions and for a time sufficient to allow the two products to interact and bind, thus forming a complex. In order to test a compound for inhibitory activity, the reaction mixture is prepared in the presence and absence of the test compound. The test compound can be initially included in the reaction mixture, or can be added at a time subsequent to the addition of Protease-42 polypeptide and its cellular or extracellular binding partner. Control reaction mixtures are incubated without the test compound or with a placebo. The formation of any complexes between the Protease-42 polypeptide and the cellular or extracellular binding partner is then detected. The formation of a complex in the control reaction, but not in the reaction mixture containing the test compound, indicates that the compound interferes with the interaction of the Protease-42 polypeptide and the interactive binding partner. Additionally, complex formation within reaction mixtures containing the test compound and normal Protease-42 polypeptide can also be compared to complex formation within reaction mixtures containing the test compound and mutant Protease-42 polypeptide. This comparison can be important in those cases wherein it is desirable to identify compounds that disrupt interactions of mutant but not normal Protease-42 polypeptide.

[0914] The assay for compounds that interfere with the interaction of the Protease-42 polypeptide and binding partners can be conducted in a heterogeneous or homogeneous format. Heterogeneous assays involve anchoring either the Protease-42 polypeptide or the binding partner onto a solid phase and detecting complexes anchored on the solid phase at the end of the reaction. In homogeneous assays, the entire reaction is carried out in a liquid phase. In either approach, the order of addition of reactants can be varied to obtain different information about the compounds being tested. For example, test compounds that interfere with the interaction between the Protease-42 polypeptide and the binding partners, e.g., by competition, can be identified by conducting the reaction in the presence of the test substance; i.e., by adding the test substance to the reaction mixture prior to or simultaneously with the Protease-42 polypeptide and interactive cellular or extracellular binding partner. Alternatively, test compounds that disrupt preformed complexes, e.g. compounds with higher binding constants that displace one of the components from the complex, can be tested by

adding the test compound to the reaction mixture after complexes have been formed. The various formats are described briefly below.

[0915] In a heterogeneous assay system, either the Protease-42 polypeptide or the interactive cellular or extracellular binding partner, is anchored onto a solid surface, while the non-anchored species is labeled, either directly or indirectly. In practice, microtitre plates are conveniently utilized. The anchored species can be immobilized by non-covalent or covalent attachments. Non-covalent attachment can be accomplished simply by coating the solid surface with a solution of the Protease-42 polypeptide or binding partner and drying. Alternatively, an immobilized antibody specific for the species to be anchored can be used to anchor the species to the solid surface. The surfaces can be prepared in advance and stored.

[0916] In order to conduct the assay, the partner of the immobilized species is exposed to the coated surface with or without the test compound. After the reaction is complete, unreacted components are removed (e.g., by washing) and any complexes formed will remain immobilized on the solid surface. The detection of complexes anchored on the solid surface can be accomplished in a number of ways. Where the non-immobilized species is pre-labeled, the detection of label immobilized on the surface indicates that complexes were formed. Where the non-immobilized species is not pre-labeled, an indirect label can be used to detect complexes anchored on the surface; e.g., using a labeled antibody specific for the initially non-immobilized species (the antibody, in turn, can be directly labeled or indirectly labeled with a labeled anti-Ig antibody). Depending upon the order of addition of reaction components, test compounds which inhibit complex formation or which disrupt preformed complexes can be detected.

[0917] Alternatively, the reaction can be conducted in a liquid phase in the presence or absence of the test compound, the reaction products separated from unreacted components, and complexes detected; e.g., using an immobilized antibody specific for one of the binding components to anchor any complexes formed in solution, and a labeled antibody specific for the other partner to detect anchored complexes. Again, depending upon the order of addition of reactants to the liquid phase, test compounds which inhibit complex or which disrupt preformed complexes can be identified.

[0918] In an alternate embodiment of the invention, a homogeneous assay can be used. In this approach, a preformed complex of the Protease-42 polypeptide and the interactive cellular or extracellular binding partner product is prepared in which either the Protease-42 polypeptide or their binding partners are labeled, but the signal generated by the label is quenched due to complex formation (see, e.g., U.S. Pat. No. 4,109,496 by Rubenstein which utilizes this approach for immunoassays). The addition of a test substance that competes with and displaces one of the species from the preformed complex will result in the generation of a signal above background. In this way, test substances which disrupt Protease-42 polypeptide-cellular or extracellular binding partner interaction can be identified.

[0919] In a particular embodiment, the Protease-42 polypeptide can be prepared for immobilization using recombinant DNA techniques known in the art. For example, the Protease-42 polypeptide coding region can be fused to a glutathione-S-transferase (GST) gene using a fusion vector such as pGEX-5X-1, in such a manner that its binding activity is maintained in the resulting fusion product. The interactive cellular or extracellular product can be purified and used to raise a monoclonal antibody, using methods routinely practiced in the art and described above. This antibody can be labeled with the radioactive isotope sup. 125 I, for example, by methods routinely practiced in the art. In a heterogeneous assay, e.g., the GST-Protease-42 polypeptide fusion product can be anchored to glutathione-agarose beads. The interactive cellular or extracellular binding partner product can then be added in the presence or absence of the test compound in a manner that allows interaction and binding to occur. At the end of the reaction period, unbound material can be washed away, and the labeled monoclonal antibody can be added to the system and allowed to bind to the complexed components. The interaction between the Protease-42 polypeptide and the interactive cellular or extracellular binding partner can be detected by measuring the amount of radioactivity that remains associated with the glutathione-agarose beads. A successful inhibition of the interaction by the test compound will result in a decrease in measured radioactivity.

[0920] Alternatively, the GST-Protease-42 polypeptide fusion product and the interactive cellular or extracellular binding partner product can be mixed together in liquid in the absence of the solid glutathione-agarose beads. The test compound can be added either during or after the binding partners are allowed to interact. This mixture can then be added to the glutathione-agarose beads and unbound material is washed away. Again the extent of inhibition of the binding partner interaction can be detected by adding the labeled antibody and measuring the radioactivity associated with the beads.

[0921] In another embodiment of the invention, these same techniques can be employed using peptide fragments that correspond to the binding domains of the Protease-42 polypeptide product and the interactive cellular or extracellular binding partner (in case where the binding partner is a product), in place of one or both of the full length products.

[0922] Any number of methods routinely practiced in the art can be used to identify and isolate the protein's binding site. These methods include, but are not limited to, mutagenesis of one of the genes encoding one of the products and screening for disruption of binding in a co-immunoprecipitation assay. Compensating mutations in the gene encoding the second species in the complex can be selected. Sequence analysis of the genes encoding the respective products will reveal the mutations that correspond to the region of the product involved in interactive binding. Alternatively, one product can be anchored to a solid surface using methods described in this Section above, and allowed to interact with and bind to its labeled binding partner, which has been treated with a proteolytic enzyme, such as trypsin. After washing, a short, labeled peptide comprising the binding domain can remain associated with the solid material, which can be isolated and identified by amino acid sequencing. Also, once the gene coding for the cellular or extracellular binding partner product is obtained, short gene segments can be engineered to express peptide fragments of the product, which can then be tested for binding activity and purified or synthesized.

Example 11

Method of Identifying the Cognate Ligand of the Protease-42 Polypeptide

[0923] A number of methods are known in the art for identifying the cognate binding partner of a particular polypeptide. For example, the encoding Protease-42 polynucleotide could be engineered to comprise an epitope tag. The epitope could be any epitope known in the art or disclosed elsewhere herein. Once created, the epitope tagged Protease-42 encoding polynucleotide could be cloned into an expression vector and used to transfect a variety of cell lines representing different tissue origins (e.g., brain, testis, etc.). The transfected cell lines could then be induced to overexpress the Protease-42 polypeptide. The presence of the Protease-42 polypeptide on the cell surface could be determined by fractionating whole cell lysates into cellular and membrane protein fractions and performing immunoprecipitation using the antibody directed against the epitope engineered into the Protease-42 polypeptide. Monoclonal or polyclonal antibodies directed against the Protease-42 polypeptide could be created and used in place of the antibodies directed against the epitope.

[0924] Alternatively, the cell surface proteins could be distinguished from cellular proteins by biotinylating the surface proteins and then performing immunoprecipitations with antibody specific to the Protease-42 protein. After electrophoretic separation, the biotinylated protein could be detected with streptavidin-HRP (using standard methods known to those skilled in the art). Identification of the proteins bound to Protease-42 could be made in those cells by immunoprecipation, followed by one-dimensional electrophoresis, followed by various versions of mass spectrometry. Such mass-spectrometry methods are known in the art, such as for example the methods taught by Ciphergen Biosystems Inc. (see U.S. Pat. No. 5,792,664; which is hereby incorporated herein by reference).

Example 12

Isolation of a Specific Clone from the Deposited Sample

[0925] The deposited material in the sample assigned the ATCC Deposit Number cited in Table I for any given cDNA clone also may contain one or more additional plasmids, each comprising a cDNA clone different from that given clone. Thus, deposits sharing the same ATCC Deposit Number contain at least a plasmid for each cDNA clone identified in Table 1. Typically, each ATCC deposit sample cited in Table 1 comprises a mixture of approximately equal amounts (by weight) of about 1-10 plasmid DNAs, each containing a different cDNA clone and/or partial cDNA clone; but such a deposit sample may include plasmids for more or less than 2 cDNA clones.

[0926] Two approaches can be used to isolate a particular clone from the deposited sample of plasmid DNA(s) cited for that clone in Table 1. First, a plasmid is directly isolated by screening the clones using a polynucleotide probe corresponding to SEQ ID NO:1.

[0927] Particularly, a specific polynucleotide with 30-40 nucleotides is synthesized using an Applied Biosystems DNA synthesizer according to the sequence reported. The oligonucleotide is labeled, for instance, with 32P-(-ATP using T4 polynucleotide kinase and purified according to routine methods. (E.g., Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, N.Y. (1982).) The plasmid mixture is transformed into a suitable host, as indicated above (such as XL-1 Blue (Stratagene)) using techniques known to those of skill in the art, such as those provided by the vector supplier or in related publications or patents cited above. The transformants are plated on 1.5% agar plates (containing the appropriate selection agent, e.g., ampicillin) to a density of about 150 transformants (colonies) per plate. These plates are screened using Nylon membranes according to routine methods for bacterial colony screening (e.g., Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd Edit., (1989), Cold Spring Harbor Laboratory Press, pages 1.93 to 1.104), or other techniques known to those of skill in the art.

[0928] Alternatively, two primers of 17-20 nucleotides derived from both ends of the SEQ ID NO:1 (i.e., within the region of SEQ ID NO:1 bounded by the 5' NT and the 3' NT of the clone defined in Table 1) are synthesized and used to amplify the desired cDNA using the deposited cDNA plasmid as a template. The polymerase chain reaction is carried out under routine conditions, for instance, in 25 ul of reaction mixture with 0.5 ug of the above cDNA template. A convenient reaction mixture is 1.5-5 mM MgCl2, 0.01% (w/v) gelatin, 20 uM each of dATP, dCTP, dGTP, dTTP, 25 pmol of each primer and 0.25 Unit of Taq polymerase. Thirty five cycles of PCR (denaturation at 94 degree C. for 1 min; annealing at 55 degree C. for 1 min; elongation at 72 degree C. for 1 min) are performed with a Perkin-Elmer Cetus automated thermal cycler. The amplified product is analyzed by agarose gel electrophoresis and the DNA band with expected molecular weight is excised and purified. The PCR product is verified to be the selected sequence by subcloning and sequencing the DNA product.

[0929] The polynucleotide(s) of the present invention, the polynucleotide encoding the polypeptide of the present invention, or the polypeptide encoded by the deposited clone may represent partial, or incomplete versions of the complete coding region (i.e., full-length gene). Several methods are known in the art for the identification of the 5' or 3' non-coding and/or coding portions of a gene which may not be present in the deposited clone. The methods that follow are exemplary and should not be construed as limiting the scope of the invention. These methods include but are not limited to, filter probing, clone enrichment using specific probes, and protocols similar or identical to 5' and 3'"RACE" protocols that are well known in the art. For instance, a method similar to 5' RACE is available for generating the missing 5' end of a desired full-length transcript. (Fromont-Racine et al., Nucleic Acids Res. 21(7):1683-1684 (1993)).

[0930] Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably containing full-length gene RNA transcripts. A primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest is used to PCR amplify the 5' portion of the desired full-length gene. This amplified product may then be sequenced and used to generate the full-length gene.

[0931] This above method starts with total RNA isolated from the desired source, although poly-A+ RNA can be

used. The RNA preparation can then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA that may interfere with the later RNA ligase step. The phosphatase should then be inactivated and the RNA treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase.

[0932] This modified RNA preparation is used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction is used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the gene of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the desired gene. Moreover, it may be advantageous to optimize the RACE protocol to increase the probability of isolating additional 5' or 3' coding or non-coding sequences. Various methods of optimizing a RACE protocol are known in the art, though a detailed description summarizing these methods can be found in B. C. Schaefer, Anal. Biochem., 227:255-273, (1995).

[0933] An alternative method for carrying out 5' or 3' RACE for the identification of coding or non-coding sequences is provided by Frohman, M. A., et al., Proc.Nat'l.Acad.Sci.USA, 85:8998-9002 (1988). Briefly, a cDNA clone missing either the 5' or 3' end can be reconstructed to include the absent base pairs extending to the translational start or stop codon, respectively. In some cases, cDNAs are missing the start of translation, therefor. The following briefly describes a modification of this original 5' RACE procedure. Poly A+ or total RNAs reverse transcribed with Superscript II (Gibco/BRL) and an antisense or I complementary primer specific to the cDNA sequence. The primer is removed from the reaction with a Microcon Concentrator (Amicon). The first-strand cDNA is then tailed with dATP and terminal deoxynucleotide transferase (Gibco/BRL). Thus, an anchor sequence is produced which is needed for PCR amplification. The second strand is synthesized from the dA-tail in PCR buffer, Taq DNA polymerase (Perkin-Elmer Cetus), an oligo-dT primer containing three adjacent restriction sites (XhoIJ Sail and ClaI) at the 5' end and a primer containing just these restriction sites. This double-stranded cDNA is PCR amplified for 40 cycles with the same primers as well as a nested cDNA-specific antisense primer. The PCR products are size-separated on an ethidium bromide-agarose gel and the region of gel containing cDNA products the predicted size of missing protein-coding DNA is removed. cDNA is purified from the agarose with the Magic PCR Prep kit (Promega), restriction digested with XhoI or Sail, and ligated to a plasmid such as pBluescript SKII (Stratagene) at XhoI and EcoRV sites. This DNA is transformed into bacteria and the plasmid clones sequenced to identify the correct protein-coding inserts. Correct 5' ends are confirmed by comparing this sequence with the putatively identified homologue and overlap with the partial cDNA clone. Similar methods known in the art and/or commercial kits are used to amplify and recover 3' ends.

[0934] Several quality-controlled kits are commercially available for purchase. Similar reagents and methods to those above are supplied in kit form from Gibco/BRL for both 5' and 3' RACE for recovery of full length genes. A second kit is available from Clontech which is a modification of a related technique, SLIC (single-stranded ligation to single-stranded cDNA), developed by Dumas et al., Nucleic Acids Res., 19:5227-32(1991). The major differences in procedure are that the RNA is alkaline hydrolyzed after reverse transcription and RNA ligase is used to join a restriction site-containing anchor primer to the first-strand cDNA. This obviates the necessity for the dA-tailing reaction which results in a polyT stretch that is difficult to sequence past.

[0935] An alternative to generating 5' or 3' cDNA from RNA is to use cDNA library double-stranded DNA. An asymmetric PCR-amplified antisense cDNA strand is synthesized with an antisense cDNA-specific primer and a plasmid-anchored primer. These primers are removed and a symmetric PCR reaction is performed with a nested cDNA-specific antisense primer and the plasmid-anchored primer.

RNA Ligase Protocol for Generating the 5' or 3' End Sequences to Obtain Full Length Genes

[0936] Once a gene of interest is identified, several methods are available for the identification of the 5' or 3' portions of the gene which may not be present in the original cDNA plasmid. These methods include, but are not limited to, filter probing, clone enrichment using specific probes and protocols similar and identical to 5' and 3'RACE. While the full-length gene may be present in the library and can be identified by probing, a useful method for generating the 5' or 3' end is to use the existing sequence information from the original cDNA to generate the missing information. A method similar to 5'RACE is available for generating the missing 5' end of a desired full-length gene. (This method was published by Fromont-Racine et al., Nucleic Acids Res., 21(7): 1683-1684 (1993)). Briefly, a specific RNA oligonucleotide is ligated to the 5' ends of a population of RNA presumably 30 containing full-length gene RNA transcript and a primer set containing a primer specific to the ligated RNA oligonucleotide and a primer specific to a known sequence of the gene of interest, is used to PCR amplify the 5' portion of the desired full length gene which may then be sequenced and used to generate the full length gene. This method starts with total RNA isolated from the desired source, poly A RNA may be used but is not a prerequisite for this procedure. The RNA preparation may then be treated with phosphatase if necessary to eliminate 5' phosphate groups on degraded or damaged RNA which may interfere with the later RNA ligase step. The phosphatase if used is then inactivated and the RNA is treated with tobacco acid pyrophosphatase in order to remove the cap structure present at the 5' ends of messenger RNAs. This reaction leaves a 5' phosphate group at the 5' end of the cap cleaved RNA which can then be ligated to an RNA oligonucleotide using T4 RNA ligase. This modified RNA preparation can then be used as a template for first strand cDNA synthesis using a gene specific oligonucleotide. The first strand synthesis reaction can then be used as a template for PCR amplification of the desired 5' end using a primer specific to the ligated RNA oligonucleotide and a primer specific to the known sequence of the apoptosis related of interest. The resultant product is then sequenced and analyzed to confirm that the 5' end sequence belongs to the relevant apoptosis related.

Example 13

Bacterial Expression of a Polypeptide

[0937] A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 10, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites, such as BamHI and XbaI, at the 5' end of the primers in order to clone the amplified product into the expression vector. For example, BamHI and XbaI correspond to the restriction enzyme sites on the bacterial expression vector pQE-9. (Qiagen, Inc., Chatsworth, Calif.). This plasmid vector encodes antibiotic resistance (Ampr), a bacterial origin of replication (ori), an IPTG-regulatable promoter/operator (P/O), a ribosome binding site (RBS), a 6-histidine tag (6-His), and restriction enzyme cloning sites.

[0938] The pQE-9 vector is digested with BamHI and XbaI and the amplified fragment is ligated into the pQE-9 vector maintaining the reading frame initiated at the bacterial RBS. The ligation mixture is then used to transform the *E. coli* strain M15/rep4 (Qiagen, Inc.) which contains multiple copies of the plasmid pREP4, that expresses the lacI repressor and also confers kanamycin resistance (Kanr). Transformants are identified by their ability to grow on LB plates and ampicillin/kanamycin resistant colonies are selected. Plasmid DNA is isolated and confirmed by restriction analysis.

[0939] Clones containing the desired constructs are grown overnight (O/N) in liquid culture in LB media supplemented with both Amp (100 ug/ml) and Kan (25 ug/ml). The O/N culture is used to inoculate a large culture at a ratio of 1:100 to 1:250. The cells are grown to an optical density 600 (O.D.600) of between 0.4 and 0.6. IPTG (Isopropyl-B-D-thiogalacto pyranoside) is then added to a final concentration of 1 mM. IPTG induces by inactivating the lacI repressor, clearing the P/O leading to increased gene expression.

[0940] Cells are grown for an extra 3 to 4 hours. Cells are then harvested by centrifugation (20 mins at 6000×g). The cell pellet is solubilized in the chaotropic agent 6 Molar Guanidine HCl by stirring for 3-4 hours at 4 degree C. The cell debris is removed by centrifugation, and the supernatant containing the polypeptide is loaded onto a nickel-nitrilo-tri-acetic acid ("Ni-NTA") affinity resin column (available from QIAGEN, Inc., supra). Proteins with a 6×His tag bind to the Ni-NTA resin with high affinity and can be purified in a simple one-step procedure (for details see: The QIAexpressionist (1995) QIAGEN, Inc., supra).

[0941] Briefly, the supernatant is loaded onto the column in 6 M guanidine-HCl, pH 8, the column is first washed with 10 volumes of 6 M guanidine-HCl, pH 8, then washed with 10 volumes of 6 M guanidine-HCl pH 6, and finally the polypeptide is eluted with 6 M guanidine-HCl, pH 5.

[0942] The purified protein is then renatured by dialyzing it against phosphate-buffered saline (PBS) or 50 mM Na-acetate, pH 6 buffer plus 200 mM NaCl. Alternatively, the protein can be successfully refolded while immobilized on the Ni-NTA column. The recommended conditions are as follows: renature using a linear 6M-1M urea gradient in 500 mM NaCl, 20% glycerol, 20 mM Tris/HCl pH 7.4, containing protease inhibitors. The renaturation should be performed over a period of 1.5 hours or more. After renaturation the proteins are eluted by the addition of 250 mM imidazole. Imidazole is removed by a final dialyzing step against PBS or 50 mM sodium acetate pH 6 buffer plus 200 mM NaCl. The purified protein is stored at 4 degree C. or frozen at −80 degree C.

Example 14

Purification of a Polypeptide from an Inclusion Body

[0943] The following alternative method can be used to purify a polypeptide expressed in *E coli* when it is present in the form of inclusion bodies. Unless otherwise specified, all of the following steps are conducted at 4-10 degree C.

[0944] Upon completion of the production phase of the *E. coli* fermentation, the cell culture is cooled to 4-10 degree C. and the cells harvested by continuous centrifugation at 15,000 rpm (Heraeus Sepatech). On the basis of the expected yield of protein per unit weight of cell paste and the amount of purified protein required, an appropriate amount of cell paste, by weight, is suspended in a buffer solution containing 100 mM Tris, 50 mM EDTA, pH 7.4. The cells are dispersed to a homogeneous suspension using a high shear mixer.

[0945] The cells are then lysed by passing the solution through a microfluidizer (Microfluidics, Corp. or APV Gaulin, Inc.) twice at 4000-6000 psi. The homogenate is then mixed with NaCl solution to a final concentration of 0.5 M NaCl, followed by centrifugation at 7000×g for 15 min. The resultant pellet is washed again using 0.5M NaCl, 100 mM Tris, 50 mM EDTA, pH 7.4.

[0946] The resulting washed inclusion bodies are solubilized with 1.5 M guanidine hydrochloride (GuHCl) for 2-4 hours. After 7000×g centrifugation for 15 min., the pellet is discarded and the polypeptide containing supernatant is incubated at 4 degree C. overnight to allow further GuHCl extraction.

[0947] Following high speed centrifugation (30,000×g) to remove insoluble particles, the GuHCl solubilized protein is refolded by quickly mixing the GuHCl extract with 20 volumes of buffer containing 50 mM sodium, pH 4.5, 150 mM NaCl, 2 mM EDTA by vigorous stirring. The refolded diluted protein solution is kept at 4 degree C. without mixing for 12 hours prior to further purification steps.

[0948] To clarify the refolded polypeptide solution, a previously prepared tangential filtration unit equipped with 0.16 um membrane filter with appropriate surface area (e.g., Filtron), equilibrated with 40 mM sodium acetate, pH 6.0 is employed. The filtered sample is loaded onto a cation exchange resin (e.g., Poros HS-50, Perceptive Biosystems). The column is washed with 40 mM sodium acetate, pH 6.0 and eluted with 250 mM, 500 mM, 1000 mM, and 1500 mM NaCl in the same buffer, in a stepwise manner. The absorbance at 280 nm of the effluent is continuously monitored. Fractions are collected and further analyzed by SDS-PAGE.

[0949] Fractions containing the polypeptide are then pooled and mixed with 4 volumes of water. The diluted sample is then loaded onto a previously prepared set of tandem columns of strong anion (Poros HQ-50, Perceptive

Biosystems) and weak anion (Poros CM-20, Perceptive Biosystems) exchange resins. The columns are equilibrated with 40 mM sodium acetate, pH 6.0. Both columns are washed with 40 mM sodium acetate, pH 6.0, 200 mM NaCl. The CM-20 column is then eluted using a 10 column volume linear gradient ranging from 0.2 M NaCl, 50 mM sodium acetate, pH 6.0 to 1.0 M NaCl, 50 mM sodium acetate, pH 6.5. Fractions are collected under constant A280 monitoring of the effluent. Fractions containing the polypeptide (determined, for instance, by 16% SDS-PAGE) are then pooled.

[0950] The resultant polypeptide should exhibit greater than 95% purity after the above refolding and purification steps. No major contaminant bands should be observed from Coomassie blue stained 16% SDS-PAGE gel when 5 ug of purified protein is loaded. The purified protein can also be tested for endotoxin/LPS contamination, and typically the LPS content is less than 0.1 ng/ml according to LAL assays.

Example 15

Cloning and Expression of a Polypeptide in a Baculovirus Expression System

[0951] In this example, the plasmid shuttle vector pAc373 is used to insert a polynucleotide into a baculovirus to express a polypeptide. A typical baculovirus expression vector contains the strong polyhedrin promoter of the *Autographa californica* nuclear polyhedrosis virus (AcMNPV) followed by convenient restriction sites, which may include, for example BamHI, Xba I and Asp718. The polyadenylation site of the simian virus 40 ("SV40") is often used for efficient polyadenylation. For easy selection of recombinant virus, the plasmid contains the beta-galactosidase gene from *E. coli* under control of a weak Drosophila promoter in the same orientation, followed by the polyadenylation signal of the polyhedrin gene. The inserted genes are flanked on both sides by viral sequences for cell-mediated homologous recombination with wild-type viral DNA to generate a viable virus that express the cloned polynucleotide.

[0952] Many other baculovirus vectors can be used in place of the vector above, such as pVL941 and pAcIM1, as one skilled in the art would readily appreciate, as long as the construct provides appropriately located signals for transcription, translation, secretion and the like, including a signal peptide and an in-frame AUG as required. Such vectors are described, for instance, in Luckow et al., Virology 170:31-39 (1989).

[0953] A polynucleotide encoding a polypeptide of the present invention is amplified using PCR oligonucleotide primers corresponding to the 5' and 3' ends of the DNA sequence, as outlined in Example 10, to synthesize insertion fragments. The primers used to amplify the cDNA insert should preferably contain restriction sites at the 5' end of the primers in order to clone the amplified product into the expression vector. Specifically, the cDNA sequence contained in the deposited clone, including the AUG initiation codon and the naturally associated leader sequence identified elsewhere herein (if applicable), is amplified using the PCR protocol described in Example 10. If the naturally occurring signal sequence is used to produce the protein, the vector used does not need a second signal peptide. Alternatively, the vector can be modified to include a baculovirus leader sequence, using the standard methods described in Summers et al., "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures," Texas Agricultural Experimental Station Bulletin No. 1555 (1987).

[0954] The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Calif.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

[0955] The plasmid is digested with the corresponding restriction enzymes and optionally, can be dephosphorylated using calf intestinal phosphatase, using routine procedures known in the art. The DNA is then isolated from a 1% agarose gel using a commercially available kit ("Geneclean" BIO 101 Inc., La Jolla, Calif.).

[0956] The fragment and the dephosphorylated plasmid are ligated together with T4 DNA ligase. *E. coli* HB101 or other suitable *E. coli* hosts such as XL-1 Blue (Stratagene Cloning Systems, La Jolla, Calif.) cells are transformed with the ligation mixture and spread on culture plates. Bacteria containing the plasmid are identified by digesting DNA from individual colonies and analyzing the digestion product by gel electrophoresis. The sequence of the cloned fragment is confirmed by DNA sequencing.

[0957] Five ug of a plasmid containing the polynucleotide is co-transformed with 1.0 ug of a commercially available linearized baculovirus DNA ("BaculoGoldtm baculovirus DNA", Pharmingen, San Diego, Calif.), using the lipofection method described by Felgner et al., Proc. Natl. Acad. Sci. USA 84:7413-7417 (1987). One ug of BaculoGoldtm virus DNA and 5 ug of the plasmid are mixed in a sterile well of a microtiter plate containing 50 ul of serum-free Grace's medium (Life Technologies Inc., Gaithersburg, Md.). Afterwards, 10 ul Lipofectin plus 90 ul Grace's medium are added, mixed and incubated for 15 minutes at room temperature. Then the transfection mixture is added drop-wise to Sf9 insect cells (ATCC CRL 1711) seeded in a 35 mm tissue culture plate with 1 ml Grace's medium without serum. The plate is then incubated for 5 hours at 27 degrees C. The transfection solution is then removed from the plate and 1 ml of Grace's insect medium supplemented with 10% fetal calf serum is added. Cultivation is then continued at 27 degrees C. for four days.

[0958] After four days the supernatant is collected and a plaque assay is performed, as described by Summers and Smith, supra. An agarose gel with "Blue Gal" (Life Technologies Inc., Gaithersburg) is used to allow easy identification and isolation of gal-expressing clones, which produce blue-stained plaques. (A detailed description of a "plaque assay" of this type can also be found in the user's guide for insect cell culture and baculovirology distributed by Life Technologies Inc., Gaithersburg, page 9-10.) After appro-

priate incubation, blue stained plaques are picked with the tip of a micropipettor (e.g., Eppendorf). The agar containing the recombinant viruses is then resuspended in a microcentrifuge tube containing 200 ul of Grace's medium and the suspension containing the recombinant baculovirus is used to infect Sf9 cells seeded in 35 mm dishes. Four days later the supernatants of these culture dishes are harvested and then they are stored at 4 degree C.

[0959] To verify the expression of the polypeptide, Sf9 cells are grown in Grace's medium supplemented with 10% heat-inactivated FBS. The cells are infected with the recombinant baculovirus containing the polynucleotide at a multiplicity of infection ("MOI") of about 2. If radiolabeled proteins are desired, 6 hours later the medium is removed and is replaced with SF900 II medium minus methionine and cysteine (available from Life Technologies Inc., Rockville, Md.). After 42 hours, 5 uCi of 35S-methionine and 5 uCi 35S-cysteine (available from Amersham) are added. The cells are further incubated for 16 hours and then are harvested by centrifugation. The proteins in the supernatant as well as the intracellular proteins are analyzed by SDS-PAGE followed by autoradiography (if radiolabeled).

[0960] Microsequencing of the amino acid sequence of the amino terminus of purified protein may be used to determine the amino terminal sequence of the produced protein.

Example 16

Expression of the Protease-42 Polypeptide in Mammalian Cells

[0961] The polypeptide of the present invention can be expressed in a mammalian cell. A typical mammalian expression vector contains a promoter element, which mediates the initiation of transcription of mRNA, a protein coding sequence, and signals required for the termination of transcription and polyadenylation of the transcript. Additional elements include enhancers, Kozak sequences and intervening sequences flanked by donor and acceptor sites for RNA splicing. Highly efficient transcription is achieved with the early and late promoters from SV40, the long terminal repeats (LTRs) from Retroviruses, e.g., RSV, HTLVI, HIVI and the early promoter of the cytomegalovirus (CMV). However, cellular elements can also be used (e.g., the human actin promoter).

[0962] Suitable expression vectors for use in practicing the present invention include, for example, vectors such as pSVL and pMSG (Pharmacia, Uppsala, Sweden), pRSVcat (ATCC 37152), pSV2dhfr (ATCC 37146), pBC12MI (ATCC 67109), pCMVSport 2.0, and pCMVSport 3.0. Mammalian host cells that could be used include, human Hela, 293, H9 and Jurkat cells, mouse NIH3T3 and C127 cells, Cos 1, Cos 7 and CV1, quail QCI-3 cells, mouse L cells and Chinese hamster ovary (CHO) cells.

[0963] Alternatively, the polypeptide can be expressed in stable cell lines containing the polynucleotide integrated into a chromosome. The co-transformation with a selectable marker such as dhfr, gpt, neomycin, hygromycin allows the identification and isolation of the transformed cells.

[0964] The transformed gene can also be amplified to express large amounts of the encoded protein. The DHFR (dihydrofolate reductase) marker is useful in developing cell lines that carry several hundred or even several thousand copies of the gene of interest. (See, e.g., Alt, F. W., et al., J. Biol. Chem. 253:1357-1370 (1978); Hamlin, J. L. and Ma, C., Biochem. et Biophys. Acta, 1097:107-143 (1990); Page, M. J. and Sydenham, M. A., Biotechnology 9:64-68 (1991).) Another useful selection marker is the enzyme glutamine synthase (GS) (Murphy et al., Biochem J. 227:277-279 (1991); Bebbington et al., Bio/Technology 10:169-175 (1992). Using these markers, the mammalian cells are grown in selective medium and the cells with the highest resistance are selected. These cell lines contain the amplified gene(s) integrated into a chromosome. Chinese hamster ovary (CHO) and NSO cells are often used for the production of proteins.

[0965] A polynucleotide of the present invention is amplified according to the protocol outlined in herein. If the naturally occurring signal sequence is used to produce the protein, the vector does not need a second signal peptide. Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891.) The amplified fragment is isolated from a 1% agarose gel using a commercially available kit ("Geneclean," BIO 101 Inc., La Jolla, Calif.). The fragment then is digested with appropriate restriction enzymes and again purified on a 1% agarose gel.

[0966] The amplified fragment is then digested with the same restriction enzyme and purified on a 1% agarose gel. The isolated fragment and the dephosphorylated vector are then ligated with T4 DNA ligase. *E. coli* HB101 or XL-1 Blue cells are then transformed and bacteria are identified that contain the fragment inserted into plasmid pC6 using, for instance, restriction enzyme analysis.

[0967] Chinese hamster ovary cells lacking an active DHFR gene is used for transformation. Five μg of an expression plasmid is cotransformed with 0.5 ug of the plasmid pSVneo using lipofectin (Felgner et al., supra). The plasmid pSV2-neo contains a dominant selectable marker, the neo gene from Tn5 encoding an enzyme that confers resistance to a group of antibiotics including G418. The cells are seeded in alpha minus MEM supplemented with 1 mg/ml G418. After 2 days, the cells are trypsinized and seeded in hybridoma cloning plates (Greiner, Germany) in alpha minus MEM supplemented with 10, 25, or 50 ng/ml of methotrexate plus 1 mg/ml G418. After about 10-14 days single clones are trypsinized and then seeded in 6-well petri dishes or 10 ml flasks using different concentrations of methotrexate (50 nM, 100 nM, 200 nM, 400 nM, 800 nM). Clones growing at the highest concentrations of methotrexate are then transferred to new 6-well plates containing even higher concentrations of methotrexate (1 uM, 2 uM, 5 uM, 10 mM, 20 mM). The same procedure is repeated until clones are obtained which grow at a concentration of 100-

200 uM. Expression of the desired gene product is analyzed, for instance, by SDS-PAGE and Western blot or by reversed phase HPLC analysis.

Example 17

Protein Fusions Between the Protease-42 Polypeptide and Another Polypeptide

[0968] The polypeptides of the present invention are preferably fused to other proteins. These fusion proteins can be used for a variety of applications. For example, fusion of the present polypeptides to His-tag, HA-tag, protein A, IgG domains, albumin, and maltose binding protein facilitates purification. (See Example described herein; see also EP A 394,827; Traunecker, et al., Nature 331:84-86 (1988).) Similarly, fusion to IgG-1, IgG-3, and albumin increases the half-life time in vivo. Nuclear localization signals fused to the polypeptides of the present invention can target the protein to a specific subcellular localization, while covalent heterodimer or homodimers can increase or decrease the activity of a fusion protein. Fusion proteins can also create chimeric molecules having more than one function. Finally, fusion proteins can increase solubility and/or stability of the fused protein compared to the non-fused protein. All of the types of fusion proteins described above can be made by modifying the following protocol, which outlines the fusion of a polypeptide to an IgG molecule.

[0969] Briefly, the human Fc portion of the IgG molecule can be PCR amplified, using primers that span the 5' and 3' ends of the sequence described below. These primers also should have convenient restriction enzyme sites that will facilitate cloning into an expression vector, preferably a mammalian expression vector. Note that the polynucleotide is cloned without a stop codon, otherwise a fusion protein will not be produced.

[0970] The naturally occurring signal sequence may be used to produce the protein (if applicable). Alternatively, if the naturally occurring signal sequence is not used, the vector can be modified to include a heterologous signal sequence. (See, e.g., WO 96/34891 and/or U.S. Pat. No. 6,066,781, supra.)

[0971] Human IgG Fc region:

(SEQ ID NO:67)

```
GGGATCCGGAGCCCAAATCTTCTGACAAAACTCACACATGCCCACCGTGC

CCAGCACCTGAATTCGAGGGTGCACCGTCAGTCTTCCTCTTCCCCCCAAA

ACCCAAGGACACCCTCATGATCTCCCGGACTCCTGAGGTCACATGCGTGG

TGGTGGACGTAAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTG

GACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCGCGGGAGGAGCAGTA

CAACAGCACGTACCGTGTGGTCAGCGTCCTCACCGTCCTGCACCAGGACT
```

-continued

```
GGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACAAAGCCCTCCCA

ACCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAACC

ACAGGTGTACACCCTGCCCCCATCCCGGGATGAGCTGACCAAGAACCAGG

TCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCAAGCGACATCGCCGTG

GAGTGGGAGAGCAATGGGCAGCCGGAGAACAACTACAAGACCACGCCTCC

CGTGCTGGACTCCGACGGCTCCTTCTTCCTCTACAGCAAGCTCACCGTGG

ACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCTCCGTGATGCAT

GAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGGG

TAAATGAGTGCGACGGCCGCGACTCTAGAGGAT
```

Example 18

Method of Creating N- and C-Terminal Mutants Corresponding to the Protease-42 Polypeptide of the Present Invention

[0972] As described elsewhere herein, the present invention encompasses the creation of N- and C-terminal deletion mutants, in addition to any combination of N- and C-terminal deletions thereof, corresponding to the Protease-42 polypeptide of the present invention. A number of methods are available to one skilled in the art for creating such mutants. Such methods may include a combination of PCR amplification and gene cloning methodology. Although one of skill in the art of molecular biology, through the use of the teachings provided or referenced herein, and/or otherwise known in the art as standard methods, could readily create each deletion mutant of the present invention, exemplary methods are described below.

[0973] Briefly, using the isolated cDNA clone encoding the full-length Protease-42 polypeptide sequence (as described in Example 10, for example), appropriate primers of about 15-25 nucleotides derived from the desired 5' and 3' positions of SEQ ID NO:1 may be designed to PCR amplify, and subsequently clone, the intended N- and/or C-terminal deletion mutant. Such primers could comprise, for example, an inititation and stop codon for the 5' and 3' primer, respectively. Such primers may also comprise restriction sites to facilitate cloning of the deletion mutant post amplification. Moreover, the primers may comprise additional sequences, such as, for example, flag-tag sequences, kozac sequences, or other sequences discussed and/or referenced herein.

[0974] For example, in the case of the M92 to S735 N-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

```
5' Primer 5'-GCAGCA GCGGCCGC ATGAGCCGCACAGACGTGTGTCAGG-3'  (SEQ ID NO:102)
                  NotI

3' Primer 5'-GCAGCA GTCGAC GGAGAAGGTGGCCACCTCCATCCAC-3'   (SEQ ID NO:103)
                  SalI
```

[0975] For example, in the case of the M1 to A643 C-terminal deletion mutant, the following primers could be used to amplify a cDNA fragment corresponding to this deletion mutant:

```
5' Primer 5'- GCAGCA GCGGCCGC ATGGCATCCAGCAGTGGGAGGGTC -3'  (SEQ ID NO: 104)
                   NotI

3' Primer 5'- GCAGCA GTCGAC GGCCTGCCACTCCAGGAGGTAGCCC -3'   (SEQ ID NO:105)
                   SalI
```

[0976] Representative PCR amplification conditions are provided below, although the skilled artisan would appreciate that other conditions may be required for efficient amplification. A 100 ul PCR reaction mixture may be prepared using 10 ng of the template DNA (cDNA clone of Protease-42), 200 uM 4dNTPs, 1 uM primers, 0.25U Taq DNA polymerase (PE), and standard Taq DNA polymerase buffer. Typical PCR cycling condition are as follows:

20–25 cycles:	45 sec, 93 degrees
	2 min, 50 degrees
	2 min, 72 degrees
1 cycle:	10 min, 72 degrees

[0977] After the final extension step of PCR, 5U Klenow Fragment may be added and incubated for 15 min at 30 degrees.

[0978] Upon digestion of the fragment with the NotI and SalI restriction enzymes, the fragment could be cloned into an appropriate expression and/or cloning vector which has been similarly digested (e.g., pSport1, among others). The skilled artisan would appreciate that other plasmids could be equally substituted, and may be desirable in certain circumstances. The digested fragment and vector are then ligated using a DNA ligase, and then used to transform competent *E. coli* cells using methods provided herein and/or otherwise known in the art.

[0979] The 5' primer sequence for amplifying any additional N-terminal deletion mutants may be determined by reference to the following formula: (S+(X*3)) to ((S+(X*3))+25), wherein 'S' is equal to the nucleotide position of the initiating start codon of the Protease-42 gene (SEQ ID NO:1), and 'X' is equal to the most N-terminal amino acid of the intended N-terminal deletion mutant. The first term will provide the start 5' nucleotide position of the 5' primer, while the second term will provide the end 3' nucleotide position of the 5' primer corresponding to sense strand of SEQ ID NO:1. Once the corresponding nucleotide positions of the primer are determined, the final nucleotide sequence may be created by the addition of applicable restriction site sequences to the 5' end of the sequence, for example. As referenced herein, the addition of other sequences to the 5' primer may be desired in certain circumstances (e.g., kozac sequences, etc.).

[0980] The 3' primer sequence for amplifying any additional N-terminal deletion mutants may be determined by reference to the following formula: (S+(X*3)) to ((S+(X*3))−25), wherein 'S' is equal to the nucleotide position of the initiating start codon of the Protease-42 gene (SEQ ID NO:1), and 'X' is equal to the most C-terminal amino acid of the intended N-terminal deletion mutant. The first term will provide the start 5' nucleotide position of the 3' primer, while the second term will provide the end 3' nucleotide position of the 3' primer corresponding to the anti-sense strand of SEQ ID NO:1. Once the corresponding nucleotide positions of the primer are determined, the final nucleotide sequence may be created by the addition of applicable restriction site sequences to the 5' end of the sequence, for example. As referenced herein, the addition of other sequences to the 3' primer may be desired in certain circumstances (e.g., stop codon sequences, etc.). The skilled artisan would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

[0981] The same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any C-terminal deletion mutant of the present invention. Moreover, the same general formulas provided above may be used in identifying the 5' and 3' primer sequences for amplifying any combination of N-terminal and C-terminal deletion mutant of the present invention. The skilled artisan would appreciate that modifications of the above nucleotide positions may be necessary for optimizing PCR amplification.

Example 19

Regulation of Protein Expression via Controlled Aggregation in the Endoplasmic Reticulum

[0982] As described more particularly herein, proteins regulate diverse cellular processes in higher organisms, ranging from rapid metabolic changes to growth and differentiation. Increased production of specific proteins could be used to prevent certain diseases and/or disease states. Thus, the ability to modulate the expression of specific proteins in an organism would provide significant benefits.

[0983] Numerous methods have been developed to date for introducing foreign genes, either under the control of an inducible, constitutively active, or endogenous promoter, into organisms. Of particular interest are the inducible promoters (see, M. Gossen, et al., Proc. Natl. Acad. Sci. USA., 89:5547 (1992); Y. Wang, et al., Proc. Natl. Acad. Sci. USA, 91:8180 (1994), D. No., et al., Proc. Natl. Acad. Sci. USA, 93:3346 (1996); and V. M. Rivera, et al., Nature Med, 2:1028 (1996); in addition to additional examples disclosed elsewhere herein). In one example, the gene for erthropoietin (Epo) was transferred into mice and primates under the control of a small molecule inducer for expression (e.g., tetracycline or rapamycin) (see, D. Bohl, et al., Blood, 92:1512, (1998); K. G. Rendahl, et al., Nat. Biotech, 16:757, (1998); V. M. Rivera, et al., Proc. Natl. Acad. Sci. USA, 96:8657 (1999); and X. Ye et al., Science, 283:88 (1999). Although such systems enable efficient induction of the gene of interest in the organism upon addition of the inducing

agent (i.e., tetracycline, rapamycin, etc.), the levels of expression tend to peak at 24 hours and trail off to background levels after 4 to 14 days. Thus, controlled transient expression is virtually impossible using these systems, though such control would be desirable.

[0984] A new alternative method of controlling gene expression levels of a protein from a transgene (i.e., includes stable and transient transformants) has recently been elucidated (V. M. Rivera., et al., Science, 287:826-830, (2000)). This method does not control gene expression at the level of the mRNA like the aforementioned systems. Rather, the system controls the level of protein in an active secreted form. In the absence of the inducing agent, the protein aggregates in the ER and is not secreted. However, addition of the inducing agent results in dis-aggregation of the protein and the subsequent secretion from the ER. Such a system affords low basal secretion, rapid, high level secretion in the presence of the inducing agent, and rapid cessation of secretion upon removal of the inducing agent. In fact, protein secretion reached a maximum level within 30 minutes of induction, and a rapid cessation of secretion within 1 hour of removing the inducing agent. The method is also applicable for controlling the level of production for membrane proteins.

[0985] Detailed methods are presented in V. M. Rivera., et al., Science, 287:826-830, (2000)), briefly:

[0986] Fusion protein constructs are created using polynucleotide sequences of the present invention with one or more copies (preferably at least 2, 3, 4, or more) of a conditional aggregation domain (CAD) a domain that interacts with itself in a ligand-reversible manner (i.e., in the presence of an inducing agent) using molecular biology methods known in the art and discussed elsewhere herein. The CAD domain may be the mutant domain isolated from the human FKBPI2 (Phe^{36} to Met) protein (as disclosed in V. M. Rivera., et al., Science, 287:826-830, (2000), or alternatively other proteins having domains with similar ligand-reversible, self-aggregation properties. As a principle of design the fusion protein vector would contain a furin cleavage sequence operably linked between the polynucleotides of the present invention and the CAD domains. Such a cleavage site would enable the proteolytic cleavage of the CAD domains from the polypeptide of the present invention subsequent to secretion from the ER and upon entry into the trans-Golgi (J. B. Denault, et al., FEBS Lett., 379:113, (1996)). Alternatively, the skilled artisan would recognize that any proteolytic cleavage sequence could be substituted for the furin sequence provided the substituted sequence is cleavable either endogenously (e.g., the furin sequence) or exogenously (e.g., post secretion, post purification, post production, etc.). The preferred sequence of each feature of the fusion protein construct, from the 5' to 3' direction with each feature being operably linked to the other, would be a promoter, signal sequence, "X" number of (CAD) x domains, the furin sequence (or other proteolytic sequence), and the coding sequence of the polypeptide of the present invention. The artisan would appreciate that the promotor and signal sequence, independent from the other, could be either the endogenous promotor or signal sequence of a polypeptide of the present invention, or alternatively, could be a heterologous signal sequence and promotor.

[0987] The specific methods described herein for controlling protein secretion levels through controlled ER aggregation are not meant to be limiting are would be generally applicable to any of the polynucleotides and polypeptides of the present invention, including variants, homologues, orthologs, and fragments therein.

Example 20

Alteration of Protein Glycosylation Sites to Enhance Characteristics of Polypeptides of the Invention

[0988] Many eukaryotic cell surface and proteins are post-translationally processed to incorporate N-linked and O-linked carbohydrates (Kornfeld and Kornfeld (1985) Annu. Rev. Biochem. 54:631-64; Rademacher et al., (1988) Annu. Rev. Biochem. 57:785-838). Protein glycosylation is thought to serve a variety of functions including: augmentation of protein folding, inhibition of protein aggregation, regulation of intracellular trafficking to organelles, increasing resistance to proteolysis, modulation of protein antigenicity, and mediation of intercellular adhesion (Fieldler and Simons (1995) Cell, 81:309-312; Helenius (1994) Mol. Biol. Of the Cell 5:253-265; Olden et al., (1978) Cell, 13:461-473; Caton et al., (1982) Cell, 37:417-427; Alexamnder and Elder (1984), Science, 226:1328-1330; and Flack et al., (1994), J. Biol. Chem., 269:14015-14020). In higher organisms, the nature and extent of glycosylation can markedly affect the circulating half-life and bio-availability of proteins by mechanisms involving receptor mediated uptake and clearance (Ashwell and Morrell, (1974), Adv. Enzymol., 41:99-128; Ashwell and Harford (1982), Ann. Rev. Biochem., 51:531-54). Receptor systems have been identified that are thought to play a major role in the clearance of serum proteins through recognition of various carbohydrate structures on the glycoproteins (Stockert (1995), Physiol. Rev., 75:591-609; Kery et al., (1992), Arch. Biochem. Biophys., 298:49-55). Thus, production strategies resulting in incomplete attachment of terminal sialic acid residues might provide a means of shortening the bioavailability and half-life of glycoproteins. Conversely, expression strategies resulting in saturation of terminal sialic acid attachment sites might lengthen protein bioavailability and half-life.

[0989] In the development of recombinant glycoproteins for use as pharmaceutical products, for example, it has been speculated that the pharmacodynamics of recombinant proteins can be modulated by the addition or deletion of glycosylation sites from a glycoproteins primary structure (Berman and Lasky (1985a) Trends in Biotechnol., 3:51-53). However, studies have reported that the deletion of N-linked glycosylation sites often impairs intracellular transport and results in the intracellular accumulation of glycosylation site variants (Machamer and Rose (1988), J. Biol. Chem., 263:5955-5960; Gallagher et al., (1992), J. Virology., 66:7136-7145; Collier et al., (1993), Biochem., 32:7818-7823; Claffey et al., (1995) Biochemica et Biophysica Acta, 1246:1-9; Dube et al., (1988), J. Biol. Chem. 263:17516-17521). While glycosylation site variants of proteins can be expressed intracellularly, it has proved difficult to recover useful quantities from growth conditioned cell culture medium.

[0990] Moreover, it is unclear to what extent a glycosylation site in one species will be recognized by another species glycosylation machinery. Due to the importance of glycosylation in protein metabolism, particularly the secre-

tion and/or expression of the protein, whether a glycosylation signal is recognized may profoundly determine a proteins ability to be expressed, either endogenously or recombinately, in another organism (i.e., expressing a human protein in *E. coli*, yeast, or viral organisms; or an *E. coli*, yeast, or viral protein in human, etc.). Thus, it may be desirable to add, delete, or modify a glycosylation site, and possibly add a glycosylation site of one species to a protein of another species to improve the proteins functional, bioprocess purification, and/or structural characteristics (e.g., a polypeptide of the present invention).

[0991] A number of methods may be employed to identify the location of glycosylation sites within a protein. One preferred method is to run the translated protein sequence through the PROSITE computer program (Swiss Institute of Bioinformatics). Once identified, the sites could be systematically deleted, or impaired, at the level of the DNA using mutagenesis methodology known in the art and available to the skilled artisan, Preferably using PCR-directed mutagenesis (See Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, N.Y. (1982)). Similarly, glycosylation sites could be added, or modified at the level of the DNA using similar methods, preferably PCR methods (See, Maniatis, supra). The results of modifying the glycosylation sites for a particular protein (e.g., solubility, secretion potential, activity, aggregation, proteolytic resistance, etc.) could then be analyzed using methods know in the art.

Example 21

Method of Enhancing the Biological Activity/Functional Characteristics of Invention Through Molecular Evolution

[0992] Although many of the most biologically active proteins known are highly effective for their specified function in an organism, they often possess characteristics that make them undesirable for transgenic, therapeutic, and/or industrial applications. Among these traits, a short physiological half-life is the most prominent problem, and is present either at the level of the protein, or the level of the proteins mRNA. The ability to extend the half-life, for example, would be particularly important for a proteins use in gene therapy, transgenic animal production, the bioprocess production and purification of the protein, and use of the protein as a chemical modulator among others. Therefore, there is a need to identify novel variants of isolated proteins possessing characteristics which enhance their application as a therapeutic for treating diseases of animal origin, in addition to the proteins applicability to common industrial and pharmaceutical applications.

[0993] Thus, one aspect of the present invention relates to the ability to enhance specific characteristics of invention through directed molecular evolution. Such an enhancement may, in a non-limiting example, benefit the inventions utility as an essential component in a kit, the inventions physical attributes such as its solubility, structure, or codon optimization, the inventions specific biological activity, including any associated enzymatic activity, the proteins enzyme kinetics, the proteins Ki, Kcat, Km, Vmax, Kd, protein-protein activity, protein-DNA binding activity, antagonist/inhibitory activity (including direct or indirect interaction), agonist activity (including direct or indirect interaction), the proteins antigenicity (e.g., where it would be desirable to either increase or decrease the antigenic potential of the protein), the immunogenicity of the protein, the ability of the protein to form dimers, trimers, or multimers with either itself or other proteins, the antigenic efficacy of the invention, including its subsequent use a preventative treatment for disease or disease states, or as an effector for targeting diseased genes. Moreover, the ability to enhance specific characteristics of a protein may also be applicable to changing the characterized activity of an enzyme to an activity completely unrelated to its initially characterized activity. Other desirable enhancements of the invention would be specific to each individual protein, and would thus be well known in the art and contemplated by the present invention.

[0994] For example, an engineered calpain may be constitutively active upon binding of its cognate substrate. Alternatively, an engineered calpain may be constitutively active in the absence of substrate binding, and/or may exhibit increased efficacy in inhibiting cysteine proteases. In yet another example, an engineered calpain may be capable of being activated with less than all of the regulatory factors and/or conditions typically required for calpain activation (e.g., substrate binding, phosphorylation, cofactor binding, Ca^{+} binding, Ca+ activation, conformational changes, etc.). Such calpain would be useful in screens to identify calpain modulators, among other uses described herein.

[0995] Directed evolution is comprised of several steps. The first step is to establish a library of variants for the gene or protein of interest. The most important step is to then select for those variants that entail the activity you wish to identify. The design of the screen is essential since your screen should be selective enough to eliminate non-useful variants, but not so stringent as to eliminate all variants. The last step is then to repeat the above steps using the best variant from the previous screen. Each successive cycle, can then be tailored as necessary, such as increasing the stringency of the screen, for example.

[0996] Over the years, there have been a number of methods developed to introduce mutations into macromolecules. Some of these methods include, random mutagenesis, "error-prone" PCR, chemical mutagenesis, site-directed mutagenesis, and other methods well known in the art (for a comprehensive listing of current mutagenesis methods, see Maniatis, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring, N.Y. (1982)). Typically, such methods have been used, for example, as tools for identifying the core functional region(s) of a protein or the function of specific domains of a protein (if a multi-domain protein). However, such methods have more recently been applied to the identification of macromolecule variants with specific or enhanced characteristics.

[0997] Random mutagenesis has been the most widely recognized method to date. Typically, this has been carried out either through the use of "error-prone" PCR (as described in Moore, J., et al, Nature Biotechnology 14:458, (1996), or through the application of randomized synthetic oligonucleotides corresponding to specific regions of interest (as described by Derbyshire, K. M. et al, Gene, 46:145-152, (1986), and Hill, DE, et al, Methods Enzymol., 55:559-568, (1987). Both approaches have limits to the level of mutagenesis that can be obtained. However, either approach

enables the investigator to effectively control the rate of mutagenesis. This is particularly important considering the fact that mutations beneficial to the activity of the enzyme are fairly rare. In fact, using too high a level of mutagenesis may counter or inhibit the desired benefit of a useful mutation.

[0998] While both of the aforementioned methods are effective for creating randomized pools of macromolecule variants, a third method, termed "DNA Shuffling", or "sexual PCR" (WPC, Stemmer, PNAS, 91:10747, (1994)) has recently been elucidated. DNA shuffling has also been referred to as "directed molecular evolution", "exon-shuffling", "directed enzyme evolution", "in vitro evolution", and "artificial evolution". Such reference terms are known in the art and are encompassed by the invention. This new, preferred, method apparently overcomes the limitations of the previous methods in that it not only propagates positive traits, but simultaneously eliminates negative traits in the resulting progeny.

[0999] DNA shuffling accomplishes this task by combining the principal of in vitro recombination, along with the method of "error-prone" PCR. In effect, you begin with a randomly digested pool of small fragments of your gene, created by Dnase I digestion, and then introduce said random fragments into an "error-prone" PCR assembly reaction. During the PCR reaction, the randomly sized DNA fragments not only hybridize to their cognate strand, but also may hybridize to other DNA fragments corresponding to different regions of the polynucleotide of interest—regions not typically accessible via hybridization of the entire polynucleotide. Moreover, since the PCR assembly reaction utilizes "error-prone" PCR reaction conditions, random mutations are introduced during the DNA synthesis step of the PCR reaction for all of the fragments—further diversifying the potential hybridization sites during the annealing step of the reaction.

[1000] A variety of reaction conditions could be utilized to carry-out the DNA shuffling reaction. However, specific reaction conditions for DNA shuffling are provided, for example, in PNAS, 91:10747, (1994). Briefly:

[1001] Prepare the DNA substrate to be subjected to the DNA shuffling reaction. Preparation may be in the form of simply purifying the DNA from contaminating cellular material, chemicals, buffers, oligonucleotide primers, deoxynucleotides, RNAs, etc., and may entail the use of DNA purification kits as those provided by Qiagen, Inc., or by the Promega, Corp., for example.

[1002] Once the DNA substrate has been purified, it would be subjected to Dnase I digestion. About 2-4 ug of the DNA substrate(s) would be digested with 0.0015 units of Dnase I (Sigma) per ul in 100 ul of 50 mM Tris-HCL, pH 7.4/1 mM MgCl2 for 10-20 min. at room temperature. The resulting fragments of 10-50 bp could then be purified by running them through a 2% low-melting point agarose gel by electrophoresis onto DE81 ion-exchange paper (Whatmann) or could be purified using Microcon concentrators (Amicon) of the appropriate molecular weight cutoff, or could use oligonucleotide purification columns (Qiagen), in addition to other methods known in the art. If using DE81 ion-exchange paper, the 10-50 bp fragments could be eluted from said paper using 1M NaCl, followed by ethanol precipitation.

[1003] The resulting purified fragments would then be subjected to a PCR assembly reaction by re-suspension in a PCR mixture containing: 2 mM of each dNTP, 2.2 mM MgCl2, 50 mM KCl, 10 mM Tris•HCL, pH 9.0, and 0.1% Triton X-100, at a final fragment concentration of 10-30 ng/ul. No primers are added at this point. Taq DNA polymerase (Promega) would be used at 2.5 units per 100 ul of reaction mixture. A PCR program of 94 C for 60 s; 94 C for 30 s, 50-55 C for 30 s, and 72 C for 30 s using 30-45 cycles, followed by 72 C for 5 min using an MJ Research (Cambridge, Mass.) PTC-150 thermocycler. After the assembly reaction is completed, a 1:40 dilution of the resulting primeness product would then be introduced into a PCR mixture (using the same buffer mixture used for the assembly reaction) containing 0.8 um of each primer and subjecting this mixture to 15 cycles of PCR (using 94 C for 30 s, 50 C for 30 s, and 72 C for 30 s). The referred primers would be primers corresponding to the nucleic acid sequences of the polynucleotide(s) utilized in the shuffling reaction. Said primers could consist of modified nucleic acid base pairs using methods known in the art and referred to else where herein, or could contain additional sequences (i.e., for adding restriction sites, mutating specific base-pairs, etc.).

[1004] The resulting shuffled, assembled, and amplified product can be purified using methods well known in the art (e.g., Qiagen PCR purification kits) and then subsequently cloned using appropriate restriction enzymes.

[1005] Although a number of variations of DNA shuffling have been published to date, such variations would be obvious to the skilled artisan and are encompassed by the invention. The DNA shuffling method can also be tailored to the desired level of mutagenesis using the methods described by Zhao, et al. (Nucl Acid Res., 25(6): 1307-1308, (1997).

[1006] As described above, once the randomized pool has been created, it can then be subjected to a specific screen to identify the variant possessing the desired characteristic(s). Once the variant has been identified, DNA corresponding to the variant could then be used as the DNA substrate for initiating another round of DNA shuffling. This cycle of shuffling, selecting the optimized variant of interest, and then re-shuffling, can be repeated until the ultimate variant is obtained. Examples of model screens applied to identify variants created using DNA shuffling technology may be found in the following publications: J. C., Moore, et al., J. Mol. Biol., 272:336-347, (1997), F. R., Cross, et al., Mol. Cell. Biol., 18:2923-2931, (1998), and A. Crameri., et al., Nat. Biotech., 15:436-438, (1997).

[1007] DNA shuffling has several advantages. First, it makes use of beneficial mutations. When combined with screening, DNA shuffling allows the discovery of the best mutational combinations and does not assume that the best combination contains all the mutations in a population. Secondly, recombination occurs simultaneously with point mutagenesis. An effect of forcing DNA polymerase to synthesize full-length genes from the small fragment DNA pool is a background mutagenesis rate. In combination with a stringent selection method, enzymatic activity has been evolved up to 16000 fold increase over the wild-type form of the enzyme. In essence, the background mutagenesis yielded the genetic variability on which recombination acted to enhance the activity.

[1008] A third feature of recombination is that it can be used to remove deleterious mutations. As discussed above,

during the process of the randomization, for every one beneficial mutation, there may be at least one or more neutral or inhibitory mutations. Such mutations can be removed by including in the assembly reaction an excess of the wild-type random-size fragments, in addition to the random-size fragments of the selected mutant from the previous selection. During the next selection, some of the most active variants of the polynucleotide/polypeptide/enzyme, should have lost the inhibitory mutations.

[1009] Finally, recombination enables parallel processing. This represents a significant advantage since there are likely multiple characteristics that would make a protein more desirable (e.g. solubility, activity, etc.). Since it is increasingly difficult to screen for more than one desirable trait at a time, other methods of molecular evolution tend to be inhibitory. However, using recombination, it would be possible to combine the randomized fragments of the best representative variants for the various traits, and then select for multiple properties at once.

[1010] DNA shuffling can also be applied to the polynucleotides and polypeptides of the present invention to decrease their immunogenicity in a specified host. For example, a particular variant of the present invention may be created and isolated using DNA shuffling technology. Such a variant may have all of the desired characteristics, though may be highly immunogenic in a host due to its novel intrinsic structure. Specifically, the desired characteristic may cause the polypeptide to have a non-native structure which could no longer be recognized as a "self" molecule, but rather as a "foreign", and thus activate a host immune response directed against the novel variant. Such a limitation can be overcome, for example, by including a copy of the gene sequence for a xenobiotic ortholog of the native protein in with the gene sequence of the novel variant gene in one or more cycles of DNA shuffling. The molar ratio of the ortholog and novel variant DNAs could be varied accordingly. Ideally, the resulting hybrid variant identified would contain at least some of the coding sequence which enabled the xenobiotic protein to evade the host immune system, and additionally, the coding sequence of the original novel variant that provided the desired characteristics.

[1011] Likewise, the invention encompasses the application of DNA shuffling technology to the evolution of polynucleotides and polypeptides of the invention, wherein one or more cycles of DNA shuffling include, in addition to the gene template DNA, oligonucleotides coding for known allelic sequences, optimized codon sequences, known variant sequences, known polynucleotide polymorphism sequences, known ortholog sequences, known homologue sequences, additional homologous sequences, additional non-homologous sequences, sequences from another species, and any number and combination of the above.

[1012] In addition to the described methods above, there are a number of related methods that may also be applicable, or desirable in certain cases. Representative among these are the methods discussed in PCT applications WO 98/31700, and WO 98/32845, which are hereby incorporated by reference. Furthermore, related methods can also be applied to the polynucleotide sequences of the present invention in order to evolve invention for creating ideal variants for use in gene therapy, protein engineering, evolution of whole cells containing the variant, or in the evolution of entire enzyme pathways containing polynucleotides of the invention as described in PCT applications WO 98/13485, WO 98/13487, WO 98/27230, WO 98/31837, and Crameri, A., et al., Nat. Biotech., 15:436-438, (1997), respectively.

[1013] Additional methods of applying "DNA Shuffling" technology to the polynucleotides and polypeptides of the present invention, including their proposed applications, may be found in U.S. Pat. No. 5,605,793; PCT Application No. WO 95/22625; PCT Application No. WO 97/20078; PCT Application No. WO 97/35966; and PCT Application No. WO 98/42832; PCT Application No. WO 00/09727 specifically provides methods for applying DNA shuffling to the identification of herbicide selective crops which could be applied to the polynucleotides and polypeptides of the present invention; additionally, PCT Application No. WO 00/12680 provides methods and compositions for generating, modifying, adapting, and optimizing polynucleotide sequences that confer detectable phenotypic properties on plant species; each of the above are hereby incorporated in their entirety herein for all purposes.

Example 22

Method of Determining Alterations in a Gene Corresponding to a Polynucleotide

[1014] RNA isolated from entire families or individual patients presenting with a phenotype of interest (such as a disease) is be isolated. cDNA is then generated from these RNA samples using protocols known in the art. (See, Sambrook.) The cDNA is then used as a template for PCR, employing primers surrounding regions of interest in SEQ ID NO:1. Suggested PCR conditions consist of 35 cycles at 95 degrees C. for 30 seconds; 60-120 seconds at 52-58 degrees C.; and 60-120 seconds at 70 degrees C., using buffer solutions described in Sidransky et al., Science 252:706 (1991).

[1015] PCR products are then sequenced using primers labeled at their 5' end with T4 polynucleotide kinase, employing SequiTherm Polymerase. (Epicentre Technologies). The intron-exon borders of selected exons is also determined and genomic PCR products analyzed to confirm the results. PCR products harboring suspected mutations is then cloned and sequenced to validate the results of the direct sequencing.

[1016] PCR products are cloned into T-tailed vectors as described in Holton et al., Nucleic Acids Research, 19:1156 (1991) and sequenced with T7 polymerase (United States Biochemical). Affected individuals are identified by mutations not present in unaffected individuals.

[1017] Genomic rearrangements are also observed as a method of determining alterations in a gene corresponding to a polynucleotide. Genomic clones isolated according to Example 10 are nick-translated with digoxigenindeoxyuridine 5'-triphosphate (Boehringer Manheim), and FISH performed as described in Johnson et al., Methods Cell Biol. 35:73-99 (1991). Hybridization with the labeled probe is carried out using a vast excess of human cot-1 DNA for specific hybridization to the corresponding genomic locus.

[1018] Chromosomes are counterstained with 4,6-diamino-2-phenylidole and propidium iodide, producing a combination of C- and R-bands. Aligned images for precise

mapping are obtained using a triple-band filter set (Chroma Technology, Brattleboro, Vt.) in combination with a cooled charge-coupled device camera (Photometrics, Tucson, Ariz.) and variable excitation wavelength filters. (Johnson et al., Genet. Anal. Tech. Appl., 8:75 (1991).) Image collection, analysis and chromosomal fractional length measurements are performed using the ISee Graphical Program System. (Inovision Corporation, Durham, N.C.) Chromosome alterations of the genomic region hybridized by the probe are identified as insertions, deletions, and translocations. These alterations are used as a diagnostic marker for an associated disease.

Example 23

Method of Detecting Abnormal Levels of a Polypeptide in a Biological Sample

[1019] A polypeptide of the present invention can be detected in a biological sample, and if an increased or decreased level of the polypeptide is detected, this polypeptide is a marker for a particular phenotype. Methods of detection are numerous, and thus, it is understood that one skilled in the art can modify the following assay to fit their particular needs.

[1020] For example, antibody-sandwich ELISAs are used to detect polypeptides in a sample, preferably a biological sample. Wells of a microtiter plate are coated with specific antibodies, at a final concentration of 0.2 to 10 ug/ml. The antibodies are either monoclonal or polyclonal and are produced by the method described elsewhere herein. The wells are blocked so that non-specific binding of the polypeptide to the well is reduced.

[1021] The coated wells are then incubated for >2 hours at RT with a sample containing the polypeptide. Preferably, serial dilutions of the sample should be used to validate results. The plates are then washed three times with deionized or distilled water to remove unbounded polypeptide.

[1022] Next, 50 ul of specific antibody-alkaline phosphatase conjugate, at a concentration of 25-400 ng, is added and incubated for 2 hours at room temperature. The plates are again washed three times with deionized or distilled water to remove unbounded conjugate.

[1023] Add 75 ul of 4-methylumbelliferyl phosphate (MUP) or p-nitrophenyl phosphate (NPP) substrate solution to each well and incubate 1 hour at room temperature. Measure the reaction by a microtiter plate reader. Prepare a standard curve, using serial dilutions of a control sample, and plot polypeptide concentration on the X-axis (log scale) and fluorescence or absorbance of the Y-axis (linear scale). Interpolate the concentration of the polypeptide in the sample using the standard curve.

Example 24

Formulation

[1024] The invention also provides methods of treatment and/or prevention diseases, disorders, and/or conditions (such as, for example, any one or more of the diseases or disorders disclosed herein) by administration to a subject of an effective amount of a Therapeutic. By therapeutic is meant a polynucleotides or polypeptides of the invention (including fragments and variants), agonists or antagonists thereof, and/or antibodies thereto, in combination with a pharmaceutically acceptable carrier type (e.g., a sterile carrier).

[1025] The Therapeutic will be formulated and dosed in a fashion consistent with good medical practice, taking into account the clinical condition of the individual patient (especially the side effects of treatment with the Therapeutic alone), the site of delivery, the method of administration, the scheduling of administration, and other factors known to practitioners. The "effective amount" for purposes herein is thus determined by such considerations.

[1026] As a general proposition, the total pharmaceutically effective amount of the Therapeutic administered parenterally per dose will be in the range of about 1 ug/kg/day to 10 mg/kg/day of patient body weight, although, as noted above, this will be subject to therapeutic discretion. More preferably, this dose is at least 0.01 mg/kg/day, and most preferably for humans between about 0.01 and 1 mg/kg/day for the hormone. If given continuously, the Therapeutic is typically administered at a dose rate of about 1 ug/kg/hour to about 50 ug/kg/hour, either by 1-4 injections per day or by continuous subcutaneous infusions, for example, using a mini-pump. An intravenous bag solution may also be employed. The length of treatment needed to observe changes and the interval following treatment for responses to occur appears to vary depending on the desired effect.

[1027] Therapeutics can be administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

[1028] Therapeutics of the invention are also suitably administered by sustained-release systems. Suitable examples of sustained-release Therapeutics are administered orally, rectally, parenterally, intracisternally, intravaginally, intraperitoneally, topically (as by powders, ointments, gels, drops or transdermal patch), bucally, or as an oral or nasal spray. "Pharmaceutically acceptable carrier" refers to a non-toxic solid, semisolid or liquid filler, diluent, encapsulating material or formulation auxiliary of any type. The term "parenteral" as used herein refers to modes of administration which include intravenous, intramuscular, intraperitoneal, intrasternal, subcutaneous and intraarticular injection and infusion.

[1029] Therapeutics of the invention may also be suitably administered by sustained-release systems. Suitable examples of sustained-release Therapeutics include suitable polymeric materials (such as, for example, semi-permeable polymer matrices in the form of shaped articles, e.g., films, or microcapsules), suitable hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, and sparingly soluble derivatives (such as, for example, a sparingly soluble salt).

[1030] Sustained-release matrices include polylactides (U.S. Pat. No. 3,773,919, EP 58,481), copolymers of

L-glutamic acid and gamma-ethyl-L-glutamate (Sidman et al., Biopolymers 22:547-556 (1983)), poly (2-hydroxyethyl methacrylate) (Langer et al., J. Biomed. Mater. Res. 15:167-277 (1981), and Langer, Chem. Tech. 12:98-105 (1982)), ethylene vinyl acetate (Langer et al., Id.) or poly-D-(−)-3-hydroxybutyric acid (EP 133,988).

[1031] Sustained-release Therapeutics also include liposomally entrapped Therapeutics of the invention (see, generally, Langer, Science 249:1527-1533 (1990); Treat et al., in Liposomes in the Therapy of Infectious Disease and Cancer, Lopez-Berestein and Fidler (eds.), Liss, New York, pp. 317-327 and 353-365 (1989)). Liposomes containing the Therapeutic are prepared by methods known per se: DE 3,218,121; Epstein et al., Proc. Natl. Acad. Sci. (USA) 82:3688-3692 (1985); Hwang et al., Proc. Natl. Acad. Sci.(USA) 77:4030-4034 (1980); EP 52,322; EP 36,676; EP 88,046; EP 143,949; EP 142,641; Japanese Pat. Appl. 83-118008; U.S. Pat. Nos. 4,485,045 and 4,544,545; and EP 102,324. Ordinarily, the liposomes are of the small (about 200-800 Angstroms) unilamellar type in which the lipid content is greater than about 30 mol. percent cholesterol, the selected proportion being adjusted for the optimal Therapeutic.

[1032] In yet an additional embodiment, the Therapeutics of the invention are delivered by way of a pump (see Langer, supra; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989)).

[1033] Other controlled release systems are discussed in the review by Langer (Science 249:1527-1533 (1990)).

[1034] For parenteral administration, in one embodiment, the Therapeutic is formulated generally by mixing it at the desired degree of purity, in a unit dosage injectable form (solution, suspension, or emulsion), with a pharmaceutically acceptable carrier, i.e., one that is non-toxic to recipients at the dosages and concentrations employed and is compatible with other ingredients of the formulation. For example, the formulation preferably does not include oxidizing agents and other compounds that are known to be deleterious to the Therapeutic.

[1035] Generally, the formulations are prepared by contacting the Therapeutic uniformly and intimately with liquid carriers or finely divided solid carriers or both. Then, if necessary, the product is shaped into the desired formulation. Preferably the carrier is a parenteral carrier, more preferably a solution that is isotonic with the blood of the recipient. Examples of such carrier vehicles include water, saline, Ringer's solution, and dextrose solution. Non-aqueous vehicles such as fixed oils and ethyl oleate are also useful herein, as well as liposomes.

[1036] The carrier suitably contains minor amounts of additives such as substances that enhance isotonicity and chemical stability. Such materials are non-toxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate, succinate, acetic acid, and other organic acids or their salts; antioxidants such as ascorbic acid; low molecular weight (less than about ten residues) polypeptides, e.g., polyarginine or tripeptides; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids, such as glycine, glutamic acid, aspartic acid, or arginine; monosaccharides, disaccharides, and other carbohydrates including cellulose or its derivatives, glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; counterions such as sodium; and/or nonionic surfactants such as polysorbates, poloxamers, or PEG.

[1037] The Therapeutic will typically be formulated in such vehicles at a concentration of about 0.1 mg/ml to 100 mg/ml, preferably 1-10 mg/ml, at a pH of about 3 to 8. It will be understood that the use of certain of the foregoing excipients, carriers, or stabilizers will result in the formation of polypeptide salts.

[1038] Any pharmaceutical used for therapeutic administration can be sterile. Sterility is readily accomplished by filtration through sterile filtration membranes (e.g., 0.2 micron membranes). Therapeutics generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

[1039] Therapeutics ordinarily will be stored in unit or multi-dose containers, for example, sealed ampoules or vials, as an aqueous solution or as a lyophilized formulation for reconstitution. As an example of a lyophilized formulation, 10-ml vials are filled with 5 ml of sterile-filtered 1% (w/v) aqueous Therapeutic solution, and the resulting mixture is lyophilized. The infusion solution is prepared by reconstituting the lyophilized Therapeutic using bacteriostatic Water-for-Injection.

[1040] The invention also provides a pharmaceutical pack or kit comprising one or more containers filled with one or more of the ingredients of the Therapeutics of the invention. Associated with such container(s) can be a notice in the form prescribed by a governmental agency regulating the manufacture, use or sale of pharmaceuticals or biological products, which notice reflects approval by the agency of manufacture, use or sale for human administration. In addition, the Therapeutics may be employed in conjunction with other therapeutic compounds.

[1041] The Therapeutics of the invention may be administered alone or in combination with adjuvants. Adjuvants that may be administered with the Therapeutics of the invention include, but are not limited to, alum, alum plus deoxycholate (ImmunoAg), MTP-PE (Biocine Corp.), QS21 (Genentech, Inc.), BCG, and MPL. In a specific embodiment, Therapeutics of the invention are administered in combination with alum. In another specific embodiment, Therapeutics of the invention are administered in combination with QS-21. Further adjuvants that may be administered with the Therapeutics of the invention include, but are not limited to, Monophosphoryl lipid immunomodulator, AdjuVax 100a, QS-21, QS-18, CRL1005, Aluminum salts, MF-59, and Virosomal adjuvant technology. Vaccines that may be administered with the Therapeutics of the invention include, but are not limited to, vaccines directed toward protection against MMR (measles, mumps, rubella), polio, varicella, tetanus/diptheria, hepatitis A, hepatitis B, haemophilus influenzae B, whooping cough, pneumonia, influenza, Lyme's Disease, rotavirus, cholera, yellow fever, Japanese encephalitis, poliomyelitis, rabies, typhoid fever, and pertussis. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations

in which the combined agents are administered together as a therapeutic mixture, and also procedures in which the combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given first, followed by the second.

[1042] The Therapeutics of the invention may be administered alone or in combination with other therapeutic agents. Therapeutic agents that may be administered in combination with the Therapeutics of the invention, include but not limited to, other members of the TNF family, chemotherapeutic agents, antibiotics, steroidal and non-steroidal anti-inflammatories, conventional immunotherapeutic agents, cytokines and/or growth factors. Combinations may be administered either concomitantly, e.g., as an admixture, separately but simultaneously or concurrently; or sequentially. This includes presentations in which the combined agents are administered together as a therapeutic mixture, and also procedures in which the combined agents are administered separately but simultaneously, e.g., as through separate intravenous lines into the same individual. Administration "in combination" further includes the separate administration of one of the compounds or agents given first, followed by the second.

[1043] In one embodiment, the Therapeutics of the invention are administered in combination with members of the TNF family. TNF, TNF-related or TNF-like molecules that may be administered with the Therapeutics of the invention include, but are not limited to, soluble forms of TNF-alpha, lymphotoxin-alpha (LT-alpha, also known as TNF-beta), LT-beta (found in complex heterotrimer LT-alpha2-beta), OPGL, FasL, CD27L, CD30L, CD40L, 4-1BBL, DcR3, OX40L, TNF-gamma (International Publication No. WO 96/14328), AIM-I (International Publication No. WO 97/33899), endokine-alpha (International Publication No. WO 98/07880), TR6 (International Publication No. WO 98/30694), OPG, and neutrokine-alpha (International Publication No. WO 98/18921, OX40, and nerve growth factor (NGF), and soluble forms of Fas, CD30, CD27, CD40 and 4-IBB, TR2 (International Publication No. WO 96/34095), DR3 (International Publication No. WO 97/33904), DR4 (International Publication No. WO 98/32856), TR5 (International Publication No. WO 98/30693), TR6 (International Publication No. WO 98/30694), TR7 (International Publication No. WO 98/41629), TRANK, TR9 (International Publication No. WO 98/56892),TR10 (International Publication No. WO 98/54202), 312C2 (International Publication No. WO 98/06842), and TR12, and soluble forms CD154, CD70, and CD153.

[1044] In certain embodiments, Therapeutics of the invention are administered in combination with antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or protease inhibitors. Nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, RETROVIR (zidovudine/AZT), VIDEX (didanosine/ddI), HIVID (zalcitabine/ddC), ZERIT (stavudine/d4T), EPIVIR (lamivudine/3TC), and COMBIVIR (zidovudine/lamivudine). Non-nucleoside reverse transcriptase inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, VIRAMUNE (nevirapine), RESCRIPTOR (delavirdine), and SUSTIVA (efavirenz). Protease inhibitors that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, CRIXIVAN (indinavir), NORVIR (ritonavir), INVIRASE (saquinavir), and VIRACEPT (nelfinavir). In a specific embodiment, antiretroviral agents, nucleoside reverse transcriptase inhibitors, non-nucleoside reverse transcriptase inhibitors, and/or protease inhibitors may be used in any combination with Therapeutics of the invention to treat AIDS and/or to prevent or treat HIV infection.

[1045] In other embodiments, Therapeutics of the invention may be administered in combination with anti-opportunistic infection agents. Anti-opportunistic agents that may be administered in combination with the Therapeutics of the invention, include, but are not limited to, TRIMETHOPRIM-SULFAMETHOXAZOLE, DAPSONE, PENTAMIDINE, ATOVAQUONE, ISONIAZID, RIFAMPIN, PYRAZINAMIDE, ETHAMBUTOL, RIFABUTIN, CLARITHROMYCIN, AZITHROMYCIN, GANCICLOVIR, FOSCARNET, CIDOFOVIR, FLUCONAZOLE, ITRACONAZOLE, KETOCONAZOLE, ACYCLOVIR, FAMCICOLVIR, PYRIMETHAMINE, LEUCOVORIN, NEUPOGEN (filgrastim/G-CSF), and LEUKINE (sargramostim/GM-CSF). In a specific embodiment, Therapeutics of the invention are used in any combination with TRIMETHOPRIM-SULFAMETHOXAZOLE, DAPSONE, PENTAMIDINE, and/or ATOVAQUONE to prophylactically treat or prevent an opportunistic *Pneumocystis carinii* pneumonia infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ISONIAZID, RIFAMPIN, PYRAZINAMIDE, and/or ETHAMBUTOL to prophylactically treat or prevent an opportunistic *Mycobacterium avium* complex infection. In another specific embodiment, Therapeutics of the invention are used in any combination with RIFABUTIN, CLARITHROMYCIN, and/or AZITHROMYCIN to prophylactically treat or prevent an opportunistic *Mycobacterium tuberculosis* infection. In another specific embodiment, Therapeutics of the invention are used in any combination with GANCICLOVIR, FOSCARNET, and/or CIDOFOVIR to prophylactically treat or prevent an opportunistic cytomegalovirus infection. In another specific embodiment, Therapeutics of the invention are used in any combination with FLUCONAZOLE, ITRACONAZOLE, and/or KETOCONAZOLE to prophylactically treat or prevent an opportunistic fungal infection. In another specific embodiment, Therapeutics of the invention are used in any combination with ACYCLOVIR and/or FAMCICOLVIR to prophylactically treat or prevent an opportunistic herpes simplex virus type I and/or type II infection. In another specific embodiment, Therapeutics of the invention are used in any combination with PYRIMETHAMINE and/or LEUCOVORIN to prophylactically treat or prevent an opportunistic *Toxoplasma gondii* infection. In another specific embodiment, Therapeutics of the invention are used in any combination with LEUCOVORIN and/or NEUPOGEN to prophylactically treat or prevent an opportunistic bacterial infection.

[1046] In a further embodiment, the Therapeutics of the invention are administered in combination with an antiviral agent. Antiviral agents that may be administered with the Therapeutics of the invention include, but are not limited to, acyclovir, ribavirin, amantadine, and remantidine.

[1047] In a further embodiment, the Therapeutics of the invention are administered in combination with an antibiotic agent. Antibiotic agents that may be administered with the Therapeutics of the invention include, but are not limited to, amoxicillin, beta-lactamases, aminoglycosides, beta-lactam (glycopeptide), beta-lactamases, Clindamycin, chloramphenicol, cephalosporins, ciprofloxacin, ciprofloxacin, erythromycin, fluoroquinolones, macrolides, metronidazole, penicillins, quinolones, rifampin, streptomycin, sulfonamide, tetracyclines, trimethoprim, trimethoprim-sulfamthoxazole, and vancomycin.

[1048] Conventional nonspecific immunosuppressive agents, that may be administered in combination with the Therapeutics of the invention include, but are not limited to, steroids, cyclosporine, cyclosporine analogs, cyclophosphamide methylprednisone, prednisone, azathioprine, FK-506, 15-deoxyspergualin, and other immunosuppressive agents that act by suppressing the function of responding T cells.

[1049] In specific embodiments, Therapeutics of the invention are administered in combination with immunosuppressants. Immunosuppressants preparations that may be administered with the Therapeutics of the invention include, but are not limited to, ORTHOCLONE (OKT3), SANDIMMUNE/NEORAL/SANGDYA (cyclosporin), PROGRAF (tacrolimus), CELLCEPT (mycophenolate), Azathioprine, glucorticosteroids, and RAPAMUNE (sirolimus). In a specific embodiment, immunosuppressants may be used to prevent rejection of organ or bone marrow transplantation.

[1050] In an additional embodiment, Therapeutics of the invention are administered alone or in combination with one or more intravenous immune globulin preparations. Intravenous immune globulin preparations that may be administered with the Therapeutics of the invention include, but not limited to, GAMMAR, IVEEGAM, SANDOGLOBULIN, GAMMAGARD S/D, and GAMIMUNE. In a specific embodiment, Therapeutics of the invention are administered in combination with intravenous immune globulin preparations in transplantation therapy (e.g., bone marrow transplant).

[1051] In an additional embodiment, the Therapeutics of the invention are administered alone or in combination with an anti-inflammatory agent. Anti-inflammatory agents that may be administered with the Therapeutics of the invention include, but are not limited to, glucocorticoids and the nonsteroidal anti-inflammatories, aminoarylcarboxylic acid derivatives, arylacetic acid derivatives, arylbutyric acid derivatives, arylcarboxylic acids, arylpropionic acid derivatives, pyrazoles, pyrazolones, salicylic acid derivatives, thiazinecarboxamides, e-acetamidocaproic acid, S-adenosylmethionine, 3-amino-4-hydroxybutyric acid, amixetrine, bendazac, benzydamine, bucolome, difenpiramide, ditazol, emorfazone, guaiazulene, nabumetone, nimesulide, orgotein, oxaceprol, paranyline, perisoxal, pifoxime, proquazone, proxazole, and tenidap.

[1052] In another embodiment, compositions of the invention are administered in combination with a chemotherapeutic agent. Chemotherapeutic agents that may be administered with the Therapeutics of the invention include, but are not limited to, antibiotic derivatives (e.g., doxorubicin, bleomycin, daunorubicin, and dactinomycin); antiestrogens (e.g., tamoxifen); antimetabolites (e.g., fluorouracil, 5-FU, methotrexate, floxuridine, interferon alpha-2b, glutamic acid, plicamycin, mercaptopurine, and 6-thioguanine); cytotoxic agents (e.g., carmustine, BCNU, lomustine, CCNU, cytosine arabinoside, cyclophosphamide, estramustine, hydroxyurea, procarbazine, mitomycin, busulfan, cis-platin, and vincristine sulfate); hormones (e.g., medroxyprogesterone, estramustine phosphate sodium, ethinyl estradiol, estradiol, megestrol acetate, methyltestosterone, diethylstilbestrol diphosphate, chlorotrianisene, and testolactone); nitrogen mustard derivatives (e.g., mephalen, chorambucil, mechlorethamine (nitrogen mustard) and thiotepa); steroids and combinations (e.g., bethamethasone sodium phosphate); and others (e.g., dicarbazine, asparaginase, mitotane, vincristine sulfate, vinblastine sulfate, and etoposide).

[1053] In a specific embodiment, Therapeutics of the invention are administered in combination with CHOP (cyclophosphamide, doxorubicin, vincristine, and prednisone) or any combination of the components of CHOP. In another embodiment, Therapeutics of the invention are administered in combination with Rituximab. In a further embodiment, Therapeutics of the invention are administered with Rituxmab and CHOP, or Rituxmab and any combination of the components of CHOP.

[1054] In an additional embodiment, the Therapeutics of the invention are administered in combination with cytokines. Cytokines that may be administered with the Therapeutics of the invention include, but are not limited to, IL2, IL3, IL4, IL5, IL6, IL7, IL10, IL12, IL13, IL15, anti-CD40, CD40L, IFN-gamma and TNF-alpha. In another embodiment, Therapeutics of the invention may be administered with any interleukin, including, but not limited to, IL-1alpha, IL-1beta, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, and IL-21.

[1055] In an additional embodiment, the Therapeutics of the invention are administered in combination with angiogenic proteins. Angiogenic proteins that may be administered with the Therapeutics of the invention include, but are not limited to, Glioma Derived Growth Factor (GDGF), as disclosed in European Patent Number EP-399816; Platelet Derived Growth Factor-A (PDGF-A), as disclosed in European Patent Number EP-682110; Platelet Derived Growth Factor-B (PDGF-B), as disclosed in European Patent Number EP-282317; Placental Growth Factor (PlGF), as disclosed in International Publication Number WO 92/06194; Placental Growth Factor-2 (PlGF-2), as disclosed in Hauser et al., Gorwth Factors, 4:259-268 (1993); Vascular Endothelial Growth Factor (VEGF), as disclosed in International Publication Number WO 90/13649; Vascular Endothelial Growth Factor-A (VEGF-A), as disclosed in European Patent Number EP-506477; Vascular Endothelial Growth Factor-2 (VEGF-2), as disclosed in International Publication Number WO 96/39515; Vascular Endothelial Growth Factor B (VEGF-3); Vascular Endothelial Growth Factor B-186 (VEGF-B186), as disclosed in International Publication Number WO 96/26736; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/02543; Vascular Endothelial Growth Factor-D (VEGF-D), as disclosed in International Publication Number WO 98/07832; and Vascular Endothelial Growth Factor-E (VEGF-E), as disclosed in German Patent Number DE19639601. The above mentioned references are incorporated herein by reference herein.

[1056] In an additional embodiment, the Therapeutics of the invention are administered in combination with hematopoietic growth factors. Hematopoietic growth factors that may be administered with the Therapeutics of the invention include, but are not limited to, LEUKINE (SARGRAMOSTIM) and NEUPOGEN (FILGRASTIM).

[1057] In an additional embodiment, the Therapeutics of the invention are administered in combination with Fibroblast Growth Factors. Fibroblast Growth Factors that may be administered with the Therapeutics of the invention include, but are not limited to, FGF-1, FGF-2, FGF-3, FGF-4, FGF-5, FGF-6, FGF-7, FGF-8, FGF-9, FGF-10, FGF-11, FGF-12, FGF-13, FGF-14, and FGF-15.

[1058] In a specific embodiment, formulations of the present invention may further comprise antagonists of P-glycoprotein (also referred to as the multiresistance protein, or PGP), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). P-glycoprotein is well known for decreasing the efficacy of various drug administrations due to its ability to export intracellular levels of absorbed drug to the cell exterior. While this activity has been particularly pronounced in cancer cells in response to the administration of chemotherapy regimens, a variety of other cell types and the administration of other drug classes have been noted (e.g., T-cells and anti-HIV drugs). In fact, certain mutations in the PGP gene significantly reduces PGP function, making it less able to force drugs out of cells. People who have two versions of the mutated gene—one inherited from each parent—have more than four times less PGP than those with two normal versions of the gene. People may also have one normal gene and one mutated one. Certain ethnic populations have increased incidence of such PGP mutations. Among individuals from Ghana, Kenya, the Sudan, as well as African Americans, frequency of the normal gene ranged from 73% to 84%. In contrast, the frequency was 34% to 59% among British whites, Portuguese, Southwest Asian, Chinese, Filipino and Saudi populations. As a result, certain ethnic populations may require increased administration of PGP antagonist in the formulation of the present invention to arrive at the an efficacious dose of the therapeutic (e.g., those from African descent). Conversely, certain ethnic populations, particularly those having increased frequency of the mutated PGP (e.g., of Caucasian descent, or non-African descent) may require less pharmaceutical compositions in the formulation due to an effective increase in efficacy of such compositions as a result of the increased effective absorption (e.g., less PGP activity) of said composition.

[1059] Moreover, in another specific embodiment, formulations of the present invention may further comprise antagonists of OATP2 (also referred to as the multiresistance protein, or MRP2), including antagonists of its encoding polynucleotides (e.g., antisense oligonucleotides, ribozymes, zinc-finger proteins, etc.). The invention also further comprises any additional antagonists known to inhibit proteins thought to be attributable to a multidrug resistant phenotype in proliferating cells.

[1060] In additional embodiments, the Therapeutics of the invention are administered in combination with other therapeutic or prophylactic regimens, such as, for example, radiation therapy.

Example 25

Method of Treating Decreased Levels of the Polypeptide

[1061] The present invention relates to a method for treating an individual in need of an increased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an agonist of the invention (including polypeptides of the invention). Moreover, it will be appreciated that conditions caused by a decrease in the standard or normal expression level of a secreted protein in an individual can be treated by administering the polypeptide of the present invention, preferably in the secreted form. Thus, the invention also provides a method of treatment of an individual in need of an increased level of the polypeptide comprising administering to such an individual a Therapeutic comprising an amount of the polypeptide to increase the activity level of the polypeptide in such an individual.

[1062] For example, a patient with decreased levels of a polypeptide receives a daily dose 0.1-100 ug/kg of the polypeptide for six consecutive days. Preferably, the polypeptide is in the secreted form. The exact details of the dosing scheme, based on administration and formulation, are provided herein.

Example 26

Method of Treating Increased Levels of the Polypeptide

[1063] The present invention also relates to a method of treating an individual in need of a decreased level of a polypeptide of the invention in the body comprising administering to such an individual a composition comprising a therapeutically effective amount of an antagonist of the invention (including polypeptides and antibodies of the invention).

[1064] In one example, antisense technology is used to inhibit production of a polypeptide of the present invention. This technology is one example of a method of decreasing levels of a polypeptide, preferably a secreted form, due to a variety of etiologies, such as cancer. For example, a patient diagnosed with abnormally increased levels of a polypeptide is administered intravenously antisense polynucleotides at 0.5, 1.0, 1.5, 2.0 and 3.0 mg/kg day for 21 days. This treatment is repeated after a 7-day rest period if the treatment was well tolerated. The formulation of the antisense polynucleotide is provided herein.

Example 27

Method of Treatment Using Gene Therapy-Ex Vivo

[1065] One method of gene therapy transplants fibroblasts, which are capable of expressing a polypeptide, onto a patient. Generally, fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in tissue-culture medium and separated into small pieces. Small chunks of the tissue are placed on a wet surface of a tissue culture flask, approximately ten pieces are placed in each flask. The flask is turned upside down, closed tight and left at room temperature over night. After 24 hours at room temperature, the flask is inverted and the chunks of tissue remain fixed to the

bottom of the flask and fresh media (e.g., Ham's F12 media, with 10% FBS, penicillin and streptomycin) is added. The flasks are then incubated at 37 degree C. for approximately one week.

[1066] At this time, fresh media is added and subsequently changed every several days. After an additional two weeks in culture, a monolayer of fibroblasts emerge. The monolayer is trypsinized and scaled into larger flasks.

[1067] pMV-7 (Kirschmeier, P. T. et al., DNA, 7:219-25 (1988)), flanked by the long terminal repeats of the Moloney murine sarcoma virus, is digested with EcoRI and HindIII and subsequently treated with calf intestinal phosphatase. The linear vector is fractionated on agarose gel and purified, using glass beads.

[1068] The cDNA encoding a polypeptide of the present invention can be amplified using PCR primers which correspond to the 5' and 3' end sequences respectively as set forth in Example 10 using primers and having appropriate restriction sites and initiation/stop codons, if necessary. Preferably, the 5' primer contains an EcoRI site and the 3' primer includes a HindIII site. Equal quantities of the Moloney murine sarcoma virus linear backbone and the amplified EcoRI and HindIII fragment are added together, in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The ligation mixture is then used to transform bacteria HB 101, which are then plated onto agar containing kanamycin for the purpose of confirming that the vector has the gene of interest properly inserted.

[1069] The amphotropic pA317 or GP+am12 packaging cells are grown in tissue culture to confluent density in Dulbecco's Modified Eagles Medium (DMEM) with 10% calf serum (CS), penicillin and streptomycin. The MSV vector containing the gene is then added to the media and the packaging cells transduced with the vector. The packaging cells now produce infectious viral particles containing the gene (the packaging cells are now referred to as producer cells).

[1070] Fresh media is added to the transduced producer cells, and subsequently, the media is harvested from a 10 cm plate of confluent producer cells. The spent media, containing the infectious viral particles, is filtered through a millipore filter to remove detached producer cells and this media is then used to infect fibroblast cells. Media is removed from a sub-confluent plate of fibroblasts and quickly replaced with the media from the producer cells. This media is removed and replaced with fresh media. If the titer of virus is high, then virtually all fibroblasts will be infected and no selection is required. If the titer is very low, then it is necessary to use a retroviral vector that has a selectable marker, such as neo or his. Once the fibroblasts have been efficiently infected, the fibroblasts are analyzed to determine whether protein is produced.

[1071] The engineered fibroblasts are then transplanted onto the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads.

Example 28

Gene Therapy Using Endogenous Genes Corresponding to Polynucleotides of the Invention

[1072] Another method of gene therapy according to the present invention involves operably associating the endogenous polynucleotide sequence of the invention with a promoter via homologous recombination as described, for example, in U.S. Pat. No. 5,641,670, issued Jun. 24, 1997; International Publication No: WO 96/29411, published Sep. 26, 1996; International Publication No: WO 94/12650, published Aug. 4, 1994; Koller et al., Proc. Natl. Acad. Sci. USA, 86:8932-8935 (1989); and Zijlstra et al., Nature, 342:435-438 (1989). This method involves the activation of a gene which is present in the target cells, but which is not expressed in the cells, or is expressed at a lower level than desired.

[1073] Polynucleotide constructs are made which contain a promoter and targeting sequences, which are homologous to the 5' non-coding sequence of endogenous polynucleotide sequence, flanking the promoter. The targeting sequence will be sufficiently near the 5' end of the polynucleotide sequence so the promoter will be operably linked to the endogenous sequence upon homologous recombination. The promoter and the targeting sequences can be amplified using PCR. Preferably, the amplified promoter contains distinct restriction enzyme sites on the 5' and 3' ends. Preferably, the 3' end of the first targeting sequence contains the same restriction enzyme site as the 5' end of the amplified promoter and the 5' end of the second targeting sequence contains the same restriction site as the 3' end of the amplified promoter.

[1074] The amplified promoter and the amplified targeting sequences are digested with the appropriate restriction enzymes and subsequently treated with calf intestinal phosphatase. The digested promoter and digested targeting sequences are added together in the presence of T4 DNA ligase. The resulting mixture is maintained under conditions appropriate for ligation of the two fragments. The construct is size fractionated on an agarose gel then purified by phenol extraction and ethanol precipitation.

[1075] In this Example, the polynucleotide constructs are administered as naked polynucleotides via electroporation. However, the polynucleotide constructs may also be administered with transfection-facilitating agents, such as liposomes, viral sequences, viral particles, precipitating agents, etc. Such methods of delivery are known in the art.

[1076] Once the cells are transfected, homologous recombination will take place which results in the promoter being operably linked to the endogenous polynucleotide sequence. This results in the expression of polynucleotide corresponding to the polynucleotide in the cell. Expression may be detected by immunological staining, or any other method known in the art.

[1077] Fibroblasts are obtained from a subject by skin biopsy. The resulting tissue is placed in DMEM+10% fetal calf serum. Exponentially growing or early stationary phase fibroblasts are trypsinized and rinsed from the plastic surface with nutrient medium. An aliquot of the cell suspension is removed for counting, and the remaining cells are subjected to centrifugation. The supernatant is aspirated and the pellet is resuspended in 5 ml of electroporation buffer (20 mM HEPES pH 7.3, 137 mM NaCl, 5 mM KCl, 0.7 mM Na2 HPO4, 6 mM dextrose). The cells are recentrifuged, the supernatant aspirated, and the cells resuspended in electroporation buffer containing 1 mg/ml acetylated bovine serum albumin. The final cell suspension contains approximately 3×10^6 cells/ml. Electroporation should be performed immediately following resuspension.

[1078] Plasmid DNA is prepared according to standard techniques. For example, to construct a plasmid for targeting to the locus corresponding to the polynucleotide of the invention, plasmid pUC18 (MBI Fermentas, Amherst, N.Y.) is digested with HindIII. The CMV promoter is amplified by PCR with an XbaI site on the 5' end and a BamHI site on the 3'end. Two non-coding sequences are amplified via PCR: one non-coding sequence (fragment 1) is amplified with a HindIII site at the 5' end and an Xba site at the 3'end; the other non-coding sequence (fragment 2) is amplified with a BamHI site at the 5'end and a HindIII site at the 3'end. The CMV promoter and the fragments (1 and 2) are digested with the appropriate enzymes (CMV promoter—XbaI and BamHI; fragment 1—XbaI; fragment 2—BamHI) and ligated together. The resulting ligation product is digested with HindIII, and ligated with the HindIII-digested pUC18 plasmid.

[1079] Plasmid DNA is added to a sterile cuvette with a 0.4 cm electrode gap (Bio-Rad). The final DNA concentration is generally at least 120 μg/ml. 0.5 ml of the cell suspension (containing approximately 1.5.×106 cells) is then added to the cuvette, and the cell suspension and DNA solutions are gently mixed. Electroporation is performed with a Gene-Pulser apparatus (Bio-Rad). Capacitance and voltage are set at 960 μF and 250-300 V, respectively. As voltage increases, cell survival decreases, but the percentage of surviving cells that stably incorporate the introduced DNA into their genome increases dramatically. Given these parameters, a pulse time of approximately 14-20 mSec should be observed.

[1080] Electroporated cells are maintained at room temperature for approximately 5 min, and the contents of the cuvette are then gently removed with a sterile transfer pipette. The cells are added directly to 10 ml of prewarmed nutrient media (DMEM with 15% calf serum) in a 10 cm dish and incubated at 37 degree C. The following day, the media is aspirated and replaced with 10 ml of fresh media and incubated for a further 16-24 hours.

[1081] The engineered fibroblasts are then injected into the host, either alone or after having been grown to confluence on cytodex 3 microcarrier beads. The fibroblasts now produce the protein product. The fibroblasts can then be introduced into a patient as described above.

Example 29

Method of Treatment Using Gene Therapy—In Vivo

[1082] Another aspect of the present invention is using in vivo gene therapy methods to treat disorders, diseases and conditions. The gene therapy method relates to the introduction of naked nucleic acid (DNA, RNA, and antisense DNA or RNA) sequences into an animal to increase or decrease the expression of the polypeptide. The polynucleotide of the present invention may be operatively linked to a promoter or any other genetic elements necessary for the expression of the polypeptide by the target tissue. Such gene therapy and delivery techniques and methods are known in the art, see, for example, WO90/11092, WO98/11779; U.S. Pat. Nos. 5,693,622, 5,705,151, 5,580,859; Tabata et al., Cardiovasc. Res. 35(3):470-479 (1997); Chao et al., Pharmacol. Res. 35(6):517-522 (1997); Wolff, Neuromuscul. Disord. 7(5):314-318 (1997); Schwartz et al., Gene Ther. 3(5):405-411 (1996); Tsurumi et al., Circulation 94(12):3281-3290 (1996) (incorporated herein by reference).

[1083] The polynucleotide constructs may be delivered by any method that delivers injectable materials to the cells of an animal, such as, injection into the interstitial space of tissues (heart, muscle, skin, lung, liver, intestine and the like). The polynucleotide constructs can be delivered in a pharmaceutically acceptable liquid or aqueous carrier.

[1084] The term "naked" polynucleotide, DNA or RNA, refers to sequences that are free from any delivery vehicle that acts to assist, promote, or facilitate entry into the cell, including viral sequences, viral particles, liposome formulations, lipofectin or precipitating agents and the like. However, the polynucleotides of the present invention may also be delivered in liposome formulations (such as those taught in Feigner P. L. et al. (1995) Ann. NY Acad. Sci. 772:126-139 and Abdallah B. et al. (1995) Biol. Cell 85(1):1-7) which can be prepared by methods well known to those skilled in the art.

[1085] The polynucleotide vector constructs used in the gene therapy method are preferably constructs that will not integrate into the host genome nor will they contain sequences that allow for replication. Any strong promoter known to those skilled in the art can be used for driving the expression of DNA. Unlike other gene therapies techniques, one major advantage of introducing naked nucleic acid sequences into target cells is the transitory nature of the polynucleotide synthesis in the cells. Studies have shown that non-replicating DNA sequences can be introduced into cells to provide production of the desired polypeptide for periods of up to six months.

[1086] The polynucleotide construct can be delivered to the interstitial space of tissues within the an animal, including of muscle, skin, brain, lung, liver, spleen, bone marrow, thymus, heart, lymph, blood, bone, cartilage, pancreas, kidney, gall bladder, stomach, intestine, testis, ovary, uterus, rectum, nervous system, eye, gland, and connective tissue. Interstitial space of the tissues comprises the intercellular fluid, mucopolysaccharide matrix among the reticular fibers of organ tissues, elastic fibers in the walls of vessels or chambers, collagen fibers of fibrous tissues, or that same matrix within connective tissue ensheathing muscle cells or in the lacunae of bone. It is similarly the space occupied by the plasma of the circulation and the lymph fluid of the lymphatic channels. Delivery to the interstitial space of muscle tissue is preferred for the reasons discussed below. They may be conveniently delivered by injection into the tissues comprising these cells. They are preferably delivered to and expressed in persistent, non-dividing cells which are differentiated, although delivery and expression may be achieved in non-differentiated or less completely differentiated cells, such as, for example, stem cells of blood or skin fibroblasts. In vivo muscle cells are particularly competent in their ability to take up and express polynucleotides.

[1087] For the naked polynucleotide injection, an effective dosage amount of DNA or RNA will be in the range of from about 0.05 g/kg body weight to about 50 mg/kg body weight. Preferably the dosage will be from about 0.005 mg/kg to about 20 mg/kg and more preferably from about 0.05 mg/kg to about 5 mg/kg. Of course, as the artisan of ordinary skill

will appreciate, this dosage will vary according to the tissue site of injection. The appropriate and effective dosage of nucleic acid sequence can readily be determined by those of ordinary skill in the art and may depend on the condition being treated and the route of administration. The preferred route of administration is by the parenteral route of injection into the interstitial space of tissues. However, other parenteral routes may also be used, such as, inhalation of an aerosol formulation particularly for delivery to lungs or bronchial tissues, throat or mucous membranes of the nose. In addition, naked polynucleotide constructs can be delivered to arteries during angioplasty by the catheter used in the procedure.

[1088] The dose response effects of injected polynucleotide in muscle in vivo is determined as follows. Suitable template DNA for production of mRNA coding for polypeptide of the present invention is prepared in accordance with a standard recombinant DNA methodology. The template DNA, which may be either circular or linear, is either used as naked DNA or complexed with liposomes. The quadriceps muscles of mice are then injected with various amounts of the template DNA.

[1089] Five to six week old female and male Balb/C mice are anesthetized by intraperitoneal injection with 0.3 ml of 2.5% Avertin. A 1.5 cm incision is made on the anterior thigh, and the quadriceps muscle is directly visualized. The template DNA is injected in 0.1 ml of carrier in a I cc syringe through a 27 gauge needle over one minute, approximately 0.5 cm from the distal insertion site of the muscle into the knee and about 0.2 cm deep. A suture is placed over the injection site for future localization, and the skin is closed with stainless steel clips.

[1090] After an appropriate incubation time (e.g., 7 days) muscle extracts are prepared by excising the entire quadriceps. Every fifth 15 um cross-section of the individual quadriceps muscles is histochemically stained for protein expression. A time course for protein expression may be done in a similar fashion except that quadriceps from different mice are harvested at different times. Persistence of DNA in muscle following injection may be determined by Southern blot analysis after preparing total cellular DNA and HIRT supernatants from injected and control mice. The results of the above experimentation in mice can be use to extrapolate proper dosages and other treatment parameters in humans and other animals using naked DNA.

Example 30

Transgenic Animals

[1091] The polypeptides of the invention can also be expressed in transgenic animals. Animals of any species, including, but not limited to, mice, rats, rabbits, hamsters, guinea pigs, pigs, micro-pigs, goats, sheep, cows and non-human primates, e.g., baboons, monkeys, and chimpanzees may be used to generate transgenic animals. In a specific embodiment, techniques described herein or otherwise known in the art, are used to express polypeptides of the invention in humans, as part of a gene therapy protocol.

[1092] Any technique known in the art may be used to introduce the transgene (i.e., polynucleotides of the invention) into animals to produce the founder lines of transgenic animals. Such techniques include, but are not limited to, pronuclear microinjection (Paterson et al., Appl. Microbiol. Biotechnol. 40:691-698 (1994); Carver et al., Biotechnology (NY) 11: 1263-1270 (1993); Wright et al., Biotechnology (NY) 9:830-834 (1991); and Hoppe et al., U.S. Pat. No. 4,873,191 (1989)); retrovirus mediated gene transfer into germ lines (Van der Putten et al., Proc. Natl. Acad. Sci., USA 82:6148-6152 (1985)), blastocysts or embryos; gene targeting in embryonic stem cells (Thompson et al., Cell 56:313-321 (1989)); electroporation of cells or embryos (Lo, 1983, Mol Cell. Biol. 3:1803-1814 (1983)); introduction of the polynucleotides of the invention using a gene gun (see, e.g., Ulmer et al., Science 259:1745 (1993); introducing nucleic acid constructs into embryonic pleuripotent stem cells and transferring the stem cells back into the blastocyst; and sperm-mediated gene transfer (Lavitrano et al., Cell 57:717-723 (1989); etc. For a review of such techniques, see Gordon, "Transgenic Animals," Intl. Rev. Cytol. 115:171-229 (1989), which is incorporated by reference herein in its entirety.

[1093] Any technique known in the art may be used to produce transgenic clones containing polynucleotides of the invention, for example, nuclear transfer into enucleated oocytes of nuclei from cultured embryonic, fetal, or adult cells induced to quiescence (Campell et al., Nature 380:64-66 (1996); Wilmut et al., Nature 385:810-813 (1997)).

[1094] The present invention provides for transgenic animals that carry the transgene in all their cells, as well as animals which carry the transgene in some, but not all their cells, i.e., mosaic animals or chimeric. The transgene may be integrated as a single transgene or as multiple copies such as in concatamers, e.g., head-to-head tandems or head-to-tail tandems. The transgene may also be selectively introduced into and activated in a particular cell type by following, for example, the teaching of Lasko et al. (Lasko et al., Proc. Natl. Acad. Sci. USA 89:6232-6236 (1992)). The regulatory sequences required for such a cell-type specific activation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art. When it is desired that the polynucleotide transgene be integrated into the chromosomal site of the endogenous gene, gene targeting is preferred. Briefly, when such a technique is to be utilized, vectors containing some nucleotide sequences homologous to the endogenous gene are designed for the purpose of integrating, via homologous recombination with chromosomal sequences, into and disrupting the function of the nucleotide sequence of the endogenous gene. The transgene may also be selectively introduced into a particular cell type, thus inactivating the endogenous gene in only that cell type, by following, for example, the teaching of Gu et al. (Gu et al., Science 265:103-106 (1994)). The regulatory sequences required for such a cell-type specific inactivation will depend upon the particular cell type of interest, and will be apparent to those of skill in the art.

[1095] Once transgenic animals have been generated, the expression of the recombinant gene may be assayed utilizing standard techniques. Initial screening may be accomplished by Southern blot analysis or PCR techniques to analyze animal tissues to verify that integration of the transgene has taken place. The level of mRNA expression of the transgene in the tissues of the transgenic animals may also be assessed using techniques which include, but are not limited to, Northern blot analysis of tissue samples obtained from the animal, in situ hybridization analysis, and reverse tran-

scriptase-PCR(RT-PCR). Samples of transgenic gene-expressing tissue may also be evaluated immunocytochemically or immunohistochemically using antibodies specific for the transgene product.

[1096] Once the founder animals are produced, they may be bred, inbred, outbred, or crossbred to produce colonies of the particular animal. Examples of such breeding strategies include, but are not limited to: outbreeding of founder animals with more than one integration site in order to establish separate lines; inbreeding of separate lines in order to produce compound transgenics that express the transgene at higher levels because of the effects of additive expression of each transgene; crossing of heterozygous transgenic animals to produce animals homozygous for a given integration site in order to both augment expression and eliminate the need for screening of animals by DNA analysis; crossing of separate homozygous lines to produce compound heterozygous or homozygous lines; and breeding to place the transgene on a distinct background that is appropriate for an experimental model of interest.

[1097] Transgenic animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

Example 31

Knock-Out Animals

[1098] Endogenous gene expression can also be reduced by inactivating or "knocking out" the gene and/or its promoter using targeted homologous recombination. (E.g., see Smithies et al., Nature 317:230-234 (1985); Thomas & Capecchi, Cell 51:503-512 (1987); Thompson et al., Cell 5:313-321 (1989); each of which is incorporated by reference herein in its entirety). For example, a mutant, non-functional polynucleotide of the invention (or a completely unrelated DNA sequence) flanked by DNA homologous to the endogenous polynucleotide sequence (either the coding regions or regulatory regions of the gene) can be used, with or without a selectable marker and/or a negative selectable marker, to transfect cells that express polypeptides of the invention in vivo. In another embodiment, techniques known in the art are used to generate knockouts in cells that contain, but do not express the gene of interest. Insertion of the DNA construct, via targeted homologous recombination, results in inactivation of the targeted gene. Such approaches are particularly suited in research and agricultural fields where modifications to embryonic stem cells can be used to generate animal offspring with an inactive targeted gene (e.g., see Thomas & Capecchi 1987 and Thompson 1989, supra). However this approach can be routinely adapted for use in humans provided the recombinant DNA constructs are directly administered or targeted to the required site in vivo using appropriate viral vectors that will be apparent to those of skill in the art.

[1099] In further embodiments of the invention, cells that are genetically engineered to express the polypeptides of the invention, or alternatively, that are genetically engineered not to express the polypeptides of the invention (e.g., knockouts) are administered to a patient in vivo. Such cells may be obtained from the patient (i.e., animal, including human) or an MHC compatible donor and can include, but are not limited to fibroblasts, bone marrow cells, blood cells (e.g., lymphocytes), adipocytes, muscle cells, endothelial cells etc. The cells are genetically engineered in vitro using recombinant DNA techniques to introduce the coding sequence of polypeptides of the invention into the cells, or alternatively, to disrupt the coding sequence and/or endogenous regulatory sequence associated with the polypeptides of the invention, e.g., by transduction (using viral vectors, and preferably vectors that integrate the transgene into the cell genome) or transfection procedures, including, but not limited to, the use of plasmids, cosmids, YACs, naked DNA, electroporation, liposomes, etc. The coding sequence of the polypeptides of the invention can be placed under the control of a strong constitutive or inducible promoter or promoter/enhancer to achieve expression, and preferably secretion, of the polypeptides of the invention. The engineered cells which express and preferably secrete the polypeptides of the invention can be introduced into the patient systemically, e.g., in the circulation, or intraperitoneally.

[1100] Alternatively, the cells can be incorporated into a matrix and implanted in the body, e.g., genetically engineered fibroblasts can be implanted as part of a skin graft; genetically engineered endothelial cells can be implanted as part of a lymphatic or vascular graft. (See, for example, Anderson et al. U.S. Pat. No. 5,399,349; and Mulligan & Wilson, U.S. Pat. No. 5,460,959 each of which is incorporated by reference herein in its entirety).

[1101] When the cells to be administered are non-autologous or non-MHC compatible cells, they can be administered using well known techniques which prevent the development of a host immune response against the introduced cells. For example, the cells may be introduced in an encapsulated form which, while allowing for an exchange of components with the immediate extracellular environment, does not allow the introduced cells to be recognized by the host immune system.

[1102] Transgenic and "knock-out" animals of the invention have uses which include, but are not limited to, animal model systems useful in elaborating the biological function of polypeptides of the present invention, studying diseases, disorders, and/or conditions associated with aberrant expression, and in screening for compounds effective in ameliorating such diseases, disorders, and/or conditions.

Example 32

Method of Isolating Antibody Fragments Directed Against Protease-19 from a Library of scFvs

[1103] Naturally occurring V-genes isolated from human PBLs are constructed into a library of antibody fragments which contain reactivities against Protease-19 to which the donor may or may not have been exposed (see e.g., U.S. Pat. No. 5,885,793 incorporated herein by reference in its entirety).

[1104] Rescue of the Library. A library of scFvs is constructed from the RNA of human PBLs as described in PCT publication WO 92/01047. To rescue phage displaying antibody fragments, approximately 109 *E. coli* harboring the

phagemid are used to inoculate 50 ml of 2×TY containing 1% glucose and 100 μg/ml of ampicillin (2×TY-AMP-GLU) and grown to an O.D. of 0.8 with shaking. Five ml of this culture is used to inoculate 50 ml of 2×TY-AMP-GLU, 2×108 TU of delta gene 3 helper (M13 delta gene III, see PCT publication WO 92/01047) are added and the culture incubated at 37° C. for 45 minutes without shaking and then at 37° C. for 45 minutes with shaking. The culture is centrifuged at 4000 r.p.m. for 10 min. and the pellet resuspended in 2 liters of 2×TY containing 100 μg/ml ampicillin and 50 ug/ml kanamycin and grown overnight. Phage are prepared as described in PCT publication WO 92/01047.

[1105] M13 delta gene III is prepared as follows: M13 delta gene III helper phage does not encode gene III protein, hence the phage(mid) displaying antibody fragments have a greater avidity of binding to antigen. Infectious M13 delta gene III particles are made by growing the helper phage in cells harboring a pUC19 derivative supplying the wild type gene III protein during phage morphogenesis. The culture is incubated for 1 hour at 37° C. without shaking and then for a further hour at 37° C. with shaking. Cells are spun down (IEC-Centra 8,400 r.p.m. for 10 min), resuspended in 300 ml 2×TY broth containing 100 μg ampicillin/ml and 25 pg kanamycin/ml (2×TY-AMP-KAN) and grown overnight, shaking at 37° C. Phage particles are purified and concentrated from the culture medium by two PEG-precipitations (Sambrook et al., 1990), resuspended in 2 ml PBS and passed through a 0.45 μm filter (Minisart NML; Sartorius) to give a final concentration of approximately 1013 transducing units/ml (ampicillin-resistant clones).

[1106] Panning of the Library. Immunotubes (Nunc) are coated overnight in PBS with 4 ml of either 100 pg/ml or 10 pg/ml of a polypeptide of the present invention. Tubes are blocked with 2% Marvel-PBS for 2 hours at 37° C. and then washed 3 times in PBS. Approximately 1013 TU of phage is applied to the tube and incubated for 30 minutes at room temperature tumbling on an over and under turntable and then left to stand for another 1.5 hours. Tubes are washed 10 times with PBS 0.1% Tween-20 and 10 times with PBS. Phage are eluted by adding 1 ml of 100 mM triethylamine and rotating 15 minutes on an under and over turntable after which the solution is immediately neutralized with 0.5 ml of 1.0M Tris-HCl, pH 7.4. Phage are then used to infect 10 ml of mid-log *E. coli* TG1 by incubating eluted phage with bacteria for 30 minutes at 37° C. The *E. coli* are then plated on TYE plates containing 1% glucose and 100 μg/ml ampicillin. The resulting bacterial library is then rescued with delta gene 3 helper phage as described above to prepare phage for a subsequent round of selection. This process is then repeated for a total of 4 rounds of affinity purification with tube-washing increased to 20 times with PBS, 0.1% Tween-20 and 20 times with PBS for rounds 3 and 4.

[1107] Characterization of Binders. Eluted phage from the 3rd and 4th rounds of selection are used to infect *E. coli* HB 2151 and soluble scFv is produced (Marks, et al., 1991) from single colonies for assay. ELISAs are performed with microtitre plates coated with either 10 Pg/ml of the polypeptide of the present invention in 50 mM bicarbonate pH 9.6. Clones positive in ELISA are further characterized by PCR fingerprinting (see, e.g., PCT publication WO 92/01047) and then by sequencing. These ELISA positive clones may also be further characterized by techniques known in the art, such as, for example, epitope mapping, binding affinity, receptor signal transduction, ability to block or competitively inhibit antibody/antigen binding, and competitive agonistic or antagonistic activity.

[1108] Moreover, in another preferred method, the antibodies directed against the polypeptides of the present invention may be produced in plants. Specific methods are disclosed in U.S. Pat. Nos. 5,959,177, and 6,080,560, which are hereby incorporated in their entirety herein. The methods not only describe methods of expressing antibodies, but also the means of assembling foreign multimeric proteins in plants (i.e., antibodies, etc,), and the subsequent secretion of such antibodies from the plant.

Example 33

Identification and Cloning of VH and VL Domains of Antibodies Directed Against the Protease-42 Polypeptide

[1109] VH and VL domains may be identified and cloned from cell lines expressing an antibody directed against a Protease-42 epitope by performing PCR with VH and VL specific primers on cDNA made from the antibody expressing cell lines. Briefly, RNA is isolated from the cell lines and used as a template for RT-PCR designed to amplify the VH and VL domains of the antibodies expressed by the EBV cell lines. Cells may be lysed using the TRIzol reagent (Life Technologies, Rockville, Md.) and extracted with one fifth volume of chloroform. After addition of chloroform, the solution is allowed to incubate at room temperature for 10 minutes, and then centrifuged at 14, 000 rpm for 15 minutes at 4 C in a tabletop centrifuge. The supernatant is collected and RNA is precipitated using an equal volume of isopropanol. Precipitated RNA is pelleted by centrifuging at 14, 000 rpm for 15 minutes at 4 C in a tabletop centrifuge.

[1110] Following centrifugation, the supernatant is discarded and washed with 75% ethanol. Follwing the wash step, the RNA is centrifuged again at 800 rpm for 5 minutes at 4 C. The supernatant is discarded and the pellet allowed to air dry. RNA is the dissolved in DEPC water and heated to 60 C for 10 minutes. Quantities of RNA can be determined using optical density measurements. cDNA may be synthesized, according to methods well-known in the art and/or described herein, from 1.5-2.5 micrograms of RNA using reverse transciptase and random hexamer primers. cDNA is then used as a template for PCR amplification of VH and VL domains.

[1111] Primers used to amplify VH and VL genes are shown below. Typically a PCR reaction makes use of a single 5'primer and a single 3'primer. Sometimes, when the amount of available RNA template is limiting, or for greater efficiency, groups of 5' and/or 3'primers may be used. For example, sometimes all five VH-5'primers and all JH3'primers are used in a single PCR reaction. The PCR reaction is carried out in a 50 microliter volume containing 1×PCR buffer, 2 mM of each dNTP, 0.7 units of High Fidelity Taq polymerse, 5'primer mix, 3'primer mix and 7.5 microliters of cDNA. The 5'and 3'primer mix of both VH and VL can be made by pooling together 22 pmole and 28 pmole, respectively, of each of the individual primers. PCR conditions are: 96 C for 5 minutes; followed by 25 cycles of 94 C for 1 minute, 50 C for 1 minute, and 72 C for 1 minute; followed by an extension cycle of 72 C for 10 minutes. After

the reaction has been completed, sample tubes may be stored at 4 C.

Primer Sequences Used to Amplify VH domains

Primer name	PrimerSequence	SEQ ID NO:
Hu VH1-5'	CAGGTGCAGCTGGTGCAGTCTGG	68
Hu VH2-5'	CAGGTCAACTTAAGGGAGTCTGG	69
Hu VH3-5'	GAGGTGCAGCTGGTGGAGTCTGG	70
Hu VH4-5'	CAGGTGCAGCTGCAGGAGTCGGG	71
Hu VH5-5'	GAGGTGCAGCTGTTGCAGTCTGC	72
Hu VH6-5'	CAGGTACAGCTGCAGCAGTCAGG	73
Hu JH1-5'	TGAGGAGACGGTGACCAGGGTGCC	74
Hu JH3-5'	TGAAGAGACGGTGACCATTGTCCC	75
Hu JH4-5'	TGAGGAGACGGTGACCAGGGTTCC	76
Hu JH6-5'	TGAGGAGACGGTGACCGTGGTCCC	77

[1112]

Primer Sequences Used to Amplify VL domains

Primer name	Primer Sequence	SEQ ID NO:
Hu Vkappa1-5'	GACATCCAGATGACCCAGTCTCC	78
Hu Vkappa2a-5'	GATGTTGTGATGACTCAGTCTCC	79
Hu Vkappa2b-5'	GATATTGTGATGACTCAGTCTCC	80
Hu Vkappa3-5'	GAAATTGTGTTGACGCAGTCTCC	81
Hu Vkappa4-5'	GACATCGTGATGACCCAGTCTCC	82
Hu Vkappa5-5'	GAAACGACACTCACGCAGTCTCC	83
Hu Vkappa6-5'	GAAATTGTGCTGACTCAGTCTCC	84
Hu Vlambda1-5'	CAGTCTGTGTTGACGCAGCCGCC	85
Hu Vlambda2-5'	CAGTCTGCCCTGACTCAGCCTGC	86
Hu Vlambda3-5'	TCCTATGTGCTGACTCAGCCACC	87
Hu Vlambda3b-5'	TCTTCTGAGCTGACTCAGGACCC	88
Hu Vlambda4-5'	CACGTTATACTGACTCAACCGCC	89
Hu Vlambda5-5'	CAGGCTGTGCTCACTCAGCCGTC	90
Hu Vlambda6-5'	AATTTTATGCTGACTCAGCCCCA	91
Hu Jkappa1-3'	ACGTTTGATTTCCACCTTGGTCCC	92
Hu Jkappa2-3'	ACGTTTGATCTCCAGCTTGGTCCC	93
Hu Jkappa3-3'	ACGTTTGATATCCACTTTGGTCCC	94
Hu Jkappa4-3'	ACGTTTGATCTCCACCTTGGTCCC	95
Hu Jkappa5-3'	ACGTTTAATCTCCAGTCGTGTCCC	96

-continued

Primer Sequences Used to Amplify VL domains

Primer name	Primer Sequence	SEQ ID NO:
Hu Vlambda1-3'	CAGTCTGTGTTGACGCAGCCGCC	97
Hu Vlambda2-3'	CAGTCTGCCCTGACTCAGCCTGC	98
Hu Vlambda3-3'	TCCTATGTGCTGACTCAGCCACC	99
Hu Vlambda3b-3'	TCTTCTGAGCTGACTCAGGACCC	100
Hu Vlambda4-3'	CACGTTATACTGACTCAACCGCC	101
Hu Vlambda5-3'	CAGGCTGTGCTCACTCAGCCGTC	102
Hu Vlambda6-3'	AATTTTATGCTGACTCAGCCCCA	103

[1113] PCR samples are then electrophoresed on a 1.3% agarose gel. DNA bands of the expected sizes (—506 base pairs for VH domains, and 344 base pairs for VL domains) can be cut out of the gel and purified using methods well known in the art and/or described herein.

[1114] Purified PCR products can be ligated into a PCR cloning vector (TA vector from Invitrogen Inc., Carlsbad, Calif.). Individual cloned PCR products can be isolated after transfection of *E. coli* and blue/white color selection. Cloned PCR products may then be sequenced using methods commonly known in the art and/or described herein.

[1115] The PCR bands containing the VH domain and the VL domains can also be used to create full-length Ig expression vectors. VH and VL domains can be cloned into vectors containing the nucleotide sequences of a heavy (e.g., human IgG1 or human IgG4) or light chain (human kappa or human ambda) constant regions such that a complete heavy or light chain molecule could be expressed from these vectors when transfected into an appropriate host cell. Further, when cloned heavy and light chains are both expressed in one cell line (from either one or two vectors), they can assemble into a complete functional antibody molecule that is secreted into the cell culture medium. Methods using polynucleotides encoding VH and VL antibody domain to generate expression vectors that encode complete antibody molecules are well known within the art.

Example 34

Assays Detecting Stimulation or Inhibition of B Cell Proliferation and Differentiation

[1116] Generation of functional humoral immune responses requires both soluble and cognate signaling between B-lineage cells and their microenvironment. Signals may impart a positive stimulus that allows a B-lineage cell to continue its programmed development, or a negative stimulus that instructs the cell to arrest its current developmental pathway. To date, numerous stimulatory and inhibitory signals have been found to influence B cell responsiveness including IL-2, IL-4, IL-5, IL-6, IL-7, IL-10, IL-13, IL-14 and IL-15. Interestingly, these signals are by themselves weak effectors but can, in combination with various co-stimulatory proteins, induce activation, proliferation, differentiation, homing, tolerance and death among B cell populations.

[1117] One of the best studied classes of B-cell co-stimulatory proteins is the TNF-superfamily. Within this family CD40, CD27, and CD30 along with their respective ligands CD154, CD70, and CD153 have been found to regulate a variety of immune responses. Assays which allow for the detection and/or observation of the proliferation and differentiation of these B-cell populations and their precursors are valuable tools in determining the effects various proteins may have on these B-cell populations in terms of proliferation and differentiation. Listed below are two assays designed to allow for the detection of the differentiation, proliferation, or inhibition of B-cell populations and their precursors.

[1118] In Vitro Assay—Purified polypeptides of the invention, or truncated forms thereof, is assessed for its ability to induce activation, proliferation, differentiation, or inhibition and/or death in B-cell populations and their precursors. The activity of the polypeptides of the invention on purified human tonsillar B cells, measured qualitatively over the dose range from 0.1 to 10,000 ng/mL, is assessed in a standard B-lymphocyte co-stimulation assay in which purified tonsillar B cells are cultured in the presence of either formalin-fixed *Staphylococcus aureus* Cowan I (SAC) or immobilized anti-human IgM antibody as the priming agent. Second signals such as IL-2 and IL-15 synergize with SAC and IgM crosslinking to elicit B cell proliferation as measured by tritiated-thymidine incorporation. Novel synergizing agents can be readily identified using this assay. The assay involves isolating human tonsillar B cells by magnetic bead (MACS) depletion of CD3-positive cells. The resulting cell population is greater than 95% B cells as assessed by expression of CD45R(B220).

[1119] Various dilutions of each sample are placed into individual wells of a 96-well plate to which are added 105 B-cells suspended in culture medium (RPMI 1640 containing 10% FBS, 5×10^{-5}M 2ME, 100U/ml penicillin, 10 ug/ml streptomycin, and 10-5 dilution of SAC) in a total volume of 150 ul. Proliferation or inhibition is quantitated by a 20 h pulse (1 uCi/well) with 3H-thymidine (6.7 Ci/mM) beginning 72 h post factor addition. The positive and negative controls are IL2 and medium respectively.

[1120] In Vivo Assay—BALB/c mice are injected (i.p.) twice per day with buffer only, or 2 mg/Kg of a polypeptide of the invention, or truncated forms thereof. Mice receive this treatment for 4 consecutive days, at which time they are sacrificed and various tissues and serum collected for analyses. Comparison of H&E sections from normal spleens and spleens treated with polypeptides of the invention identify the results of the activity of the polypeptides on spleen cells, such as the diffusion of peri-arterial lymphatic sheaths, and/or significant increases in the nucleated cellularity of the red pulp regions, which may indicate the activation of the differentiation and proliferation of B-cell populations. Immunohistochemical studies using a B cell marker, anti-CD45R(B220), are used to determine whether any physiological changes to splenic cells, such as splenic disorganization, are due to increased B-cell representation within loosely defined B-cell zones that infiltrate established T-cell regions.

[1121] Flow cytometric analyses of the spleens from mice treated with polypeptide is used to indicate whether the polypeptide specifically increases the proportion of ThB+, CD45R(B220) dull B cells over that which is observed in control mice.

[1122] Likewise, a predicted consequence of increased mature B-cell representation in vivo is a relative increase in serum Ig titers. Accordingly, serum IgM and IgA levels are compared between buffer and polypeptide-treated mice.

[1123] One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

Example 35

T Cell Proliferation Assay

[1124] A CD3-induced proliferation assay is performed on PBMCs and is measured by the uptake of 3H-thymidine. The assay is performed as follows. Ninety-six well plates are coated with 100 (l/well of mAb to CD3 (HIT3a, Pharmingen) or isotype-matched control mAb (B33.1) overnight at 4 degrees C. (1 (g/ml in 0.05M bicarbonate buffer, pH 9.5), then washed three times with PBS. PBMC are isolated by F/H gradient centrifugation from human peripheral blood and added to quadruplicate wells (5×104/well) of mAb coated plates in RPMI containing 10% FCS and P/S in the presence of varying concentrations of polypeptides of the invention (total volume 200 ul). Relevant protein buffer and medium alone are controls. After 48 hr. culture at 37 degrees C., plates are spun for 2 min. at 1000 rpm and 100 (1 of supernatant is removed and stored −20 degrees C. for measurement of IL-2 (or other cytokines) if effect on proliferation is observed. Wells are supplemented with 100 ul of medium containing 0.5 uCi of 3H-thymidine and cultured at 37 degrees C. for 18-24 hr. Wells are harvested and incorporation of 3H-thymidine used as a measure of proliferation. Anti-CD3 alone is the positive control for proliferation. IL-2 (100 U/ml) is also used as a control which enhances proliferation. Control antibody which does not induce proliferation of T cells is used as the negative controls for the effects of polypeptides of the invention.

[1125] One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

Example 36

Effect of Polypeptides of the Invention on the Expression of MHC Class II, Costimulatory and Adhesion Molecules and Cell Differentiation of Monocytes and Monocyte-Derived Human Dendritic Cells

[1126] Dendritic cells are generated by the expansion of proliferating precursors found in the peripheral blood: adherent PBMC or elutriated monocytic fractions are cultured for 7-10 days with GM-CSF (50 ng/ml) and IL-4 (20 ng/ml). These dendritic cells have the characteristic phenotype of immature cells (expression of CD1, CD80, CD86, CD40 and MHC class II antigens). Treatment with activating factors, such as TNF-, causes a rapid change in surface phenotype (increased expression of MHC class I and II, costimulatory and adhesion molecules, downregulation of

FC(R11, upregulation of CD83). These changes correlate with increased antigen-presenting capacity and with functional maturation of the dendritic cells.

[1127] FACS analysis of surface antigens is performed as follows. Cells are treated 1-3 days with increasing concentrations of polypeptides of the invention or LPS (positive control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC- or PE-labeled monoclonal antibodies for 30 minutes at 4 degrees C. After an additional wash, the labeled cells are analyzed by flow cytometry on a FACScan (Becton Dickinson).

[1128] Effect on the production of cytokines. Cytokines generated by dendritic cells, in particular IL-12, are important in the initiation of T-cell dependent immune responses. IL-12 strongly influences the development of Th1 helper T-cell immune response, and induces cytotoxic T and NK cell function. An ELISA is used to measure the IL-12 release as follows. Dendritic cells (106/ml) are treated with increasing concentrations of polypeptides of the invention for 24 hours. LPS (100 ng/ml) is added to the cell culture as positive control. Supernatants from the cell cultures are then collected and analyzed for IL-12 content using commercial ELISA kit (e.g., R & D Systems (Minneapolis, Minn.)). The standard protocols provided with the kits are used.

[1129] Effect on the expression of MHC Class II, costimulatory and adhesion molecules. Three major families of cell surface antigens can be identified on monocytes: adhesion molecules, molecules involved in antigen presentation, and Fc receptor. Modulation of the expression of MHC class II antigens and other costimulatory molecules, such as B7 and ICAM-1, may result in changes in the antigen presenting capacity of monocytes and ability to induce T cell activation. Increase expression of Fc receptors may correlate with improved monocyte cytotoxic activity, cytokine release and phagocytosis.

[1130] FACS analysis is used to examine the surface antigens as follows. Monocytes are treated 1-5 days with increasing concentrations of polypeptides of the invention or LPS (positive control), washed with PBS containing 1% BSA and 0.02 mM sodium azide, and then incubated with 1:20 dilution of appropriate FITC-or PE-labeled monoclonal antibodies for 30 minutes at 4 degrees C. After an additional wash, the labeled cells are analyzed by flow cytometry on a FACScan (Becton Dickinson).

[1131] Monocyte activation and/or increased survival. Assays for molecules that activate (or alternatively, inactivate) monocytes and/or increase monocyte survival (or alternatively, decrease monocyte survival) are known in the art and may routinely be applied to determine whether a molecule of the invention functions as an inhibitor or activator of monocytes. Polypeptides, agonists, or antagonists of the invention can be screened using the three assays described below. For each of these assays, Peripheral blood mononuclear cells (PBMC) are purified from single donor leukopacks (American Red Cross, Baltimore, Md.) by centrifugation through a Histopaque gradient (Sigma). Monocytes are isolated from PBMC by counterflow centrifugal elutriation.

[1132] Monocyte Survival Assay. Human peripheral blood monocytes progressively lose viability when cultured in absence of serum or other stimuli. Their death results from internally regulated process (apoptosis). Addition to the culture of activating factors, such as TNF-alpha dramatically improves cell survival and prevents DNA fragmentation. Propidium iodide (PI) staining is used to measure apoptosis as follows. Monocytes are cultured for 48 hours in polypropylene tubes in serum-free medium (positive control), in the presence of 100 ng/ml TNF-alpha (negative control), and in the presence of varying concentrations of the compound to be tested. Cells are suspended at a concentration of 2×106/ml in PBS containing PI at a final concentration of 5 (g/ml, and then incubated at room temperature for 5 minutes before FACScan analysis. PI uptake has been demonstrated to correlate with DNA fragmentation in this experimental paradigm.

[1133] Effect on cytokine release. An important function of monocytes/macrophages is their regulatory activity on other cellular populations of the immune system through the release of cytokines after stimulation. An ELISA to measure cytokine release is performed as follows. Human monocytes are incubated at a density of 5×105 cells/ml with increasing concentrations of the a polypeptide of the invention and under the same conditions, but in the absence of the polypeptide. For IL-12 production, the cells are primed overnight with IFN (100 U/ml) in presence of a polypeptide of the invention. LPS (10 ng/ml) is then added. Conditioned media are collected after 24 h and kept frozen until use. Measurement of TNF-alpha, IL-10, MCP-1 and IL-8 is then performed using a commercially available ELISA kit (e.g., R & D Systems (Minneapolis, Minn.)) and applying the standard protocols provided with the kit.

[1134] Oxidative burst. Purified monocytes are plated in 96-w plate at 2-1×105 cell/well. Increasing concentrations of polypeptides of the invention are added to the wells in a total volume of 0.2 ml culture medium (RPMI 1640+10% FCS, glutamine and antibiotics). After 3 days incubation, the plates are centrifuged and the medium is removed from the wells. To the macrophage monolayers, 0.2 ml per well of phenol red solution (140 mM NaCl, 10 mM potassium phosphate buffer pH 7.0, 5.5 mM dextrose, 0.56 mM phenol red and 19 U/ml of HRPO) is added, together with the stimulant (200 nM PMA). The plates are incubated at 37(C for 2 hours and the reaction is stopped by adding 20 μl 1N NaOH per well. The absorbance is read at 610 nm. To calculate the amount of H2O2 produced by the macrophages, a standard curve of a H2O2 solution of known molarity is performed for each experiment.

[1135] One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

Example 37

Biological Effects of Polypeptides of the Invention

Astrocyte and Neuronal Assays

[1136] Recombinant polypeptides of the invention, expressed in *Escherichia coli* and purified as described above, can be tested for activity in promoting the survival, neurite outgrowth, or phenotypic differentiation of cortical neuronal cells and for inducing the proliferation of glial

fibrillary acidic protein immunopositive cells, astrocytes. The selection of cortical cells for the bioassay is based on the prevalent expression of FGF-1 and FGF-2 in cortical structures and on the previously reported enhancement of cortical neuronal survival resulting from FGF-2 treatment. A thymidine incorporation assay, for example, can be used to elucidate a polypeptide of the invention's activity on these cells.

[1137] Moreover, previous reports describing the biological effects of FGF-2 (basic FGF) on cortical or hippocampal neurons in vitro have demonstrated increases in both neuron survival and neurite outgrowth (Walicke et al., "Fibroblast growth factor promotes survival of dissociated hippocampal neurons and enhances neurite extension." Proc. Natl. Acad. Sci. USA 83:3012-3016. (1986), assay herein incorporated by reference in its entirety). However, reports from experiments done on PC-12 cells suggest that these two responses are not necessarily synonymous and may depend on not only which FGF is being tested but also on which receptor(s) are expressed on the target cells. Using the primary cortical neuronal culture paradigm, the ability of a polypeptide of the invention to induce neurite outgrowth can be compared to the response achieved with FGF-2 using, for example, a thymidine incorporation assay.

Fibroblast and Endothelial Cell Assays

[1138] Human lung fibroblasts are obtained from Clonetics (San Diego, Calif.) and maintained in growth media from Clonetics. Dermal microvascular endothelial cells are obtained from Cell Applications (San Diego, Calif.). For proliferation assays, the human lung fibroblasts and dermal microvascular endothelial cells can be cultured at 5,000 cells/well in a 96-well plate for one day in growth medium. The cells are then incubated for one day in 0.1% BSA basal medium. After replacing the medium with fresh 0.1% BSA medium, the cells are incubated with the test proteins for 3 days. Alamar Blue (Alamar Biosciences, Sacramento, Calif.) is added to each well to a final concentration of 10%. The cells are incubated for 4 hr. Cell viability is measured by reading in a CytoFluor fluorescence reader. For the PGE2 assays, the human lung fibroblasts are cultured at 5,000 cells/well in a 96-well plate for one day. After a medium change to 0.1% BSA basal medium, the cells are incubated with FGF-2 or polypeptides of the invention with or without IL-1 (for 24 hours. The supernatants are collected and assayed for PGE2 by EIA kit (Cayman, Ann Arbor, Mich.). For the IL-6 assays, the human lung fibroblasts are cultured at 5,000 cells/well in a 96-well plate for one day. After a medium change to 0.1% BSA basal medium, the cells are incubated with FGF-2 or with or without polypeptides of the invention IL-1(for 24 hours. The supernatants are collected and assayed for IL-6 by ELISA kit (Endogen, Cambridge, Mass.).

[1139] Human lung fibroblasts are cultured with FGF-2 or polypeptides of the invention for 3 days in basal medium before the addition of Alamar Blue to assess effects on growth of the fibroblasts. FGF-2 should show a stimulation at 10-2500 ng/ml which can be used to compare stimulation with polypeptides of the invention.

Parkinson Models

[1140] The loss of motor function in Parkinson's disease is attributed to a deficiency of striatal dopamine resulting from the degeneration of the nigrostriatal dopaminergic projection neurons. An animal model for Parkinson's that has been extensively characterized involves the systemic administration of 1-methyl-4 phenyl 1,2,3,6-tetrahydropyridine (MPTP). In the CNS, MPTP is taken-up by astrocytes and catabolized by monoamine oxidase B to 1-methyl-4-phenyl pyridine (MPP+) and released. Subsequently, MPP+ is actively accumulated in dopaminergic neurons by the high-affinity reuptake transporter for dopamine. MPP+ is then concentrated in mitochondria by the electrochemical gradient and selectively inhibits nicotidamide adenine disphosphate: ubiquinone oxidoreductionase (complex I), thereby interfering with electron transport and eventually generating oxygen radicals.

[1141] It has been demonstrated in tissue culture paradigms that FGF-2 (basic FGF) has trophic activity towards nigral dopaminergic neurons (Ferrari et al., Dev. Biol. 1989). Recently, Dr. Unsicker's group has demonstrated that administering FGF-2 in gel foam implants in the striatum results in the near complete protection of nigral dopaminergic neurons from the toxicity associated with MPTP exposure (Otto and Unsicker, J. Neuroscience, 1990).

[1142] Based on the data with FGF-2, polypeptides of the invention can be evaluated to determine whether it has an action similar to that of FGF-2 in enhancing dopaminergic neuronal survival in vitro and it can also be tested in vivo for protection of dopaminergic neurons in the striatum from the damage associated with MPTP treatment. The potential effect of a polypeptide of the invention is first examined in vitro in a dopaminergic neuronal cell culture paradigm. The cultures are prepared by dissecting the midbrain floor plate from gestation day 14 Wistar rat embryos. The tissue is dissociated with trypsin and seeded at a density of 200,000 cells/cm2 on polyorthinine-laminin coated glass coverslips. The cells are maintained in Dulbecco's Modified Eagle's medium and F12 medium containing hormonal supplements (Ni). The cultures are fixed with paraformaldehyde after 8 days in vitro and are processed for tyrosine hydroxylase, a specific marker for dopaminergic neurons, immunohistochemical staining. Dissociated cell cultures are prepared from embryonic rats. The culture medium is changed every third day and the factors are also added at that time.

[1143] Since the dopaminergic neurons are isolated from animals at gestation day 14, a developmental time which is past the stage when the dopaminergic precursor cells are proliferating, an increase in the number of tyrosine hydroxylase immunopositive neurons would represent an increase in the number of dopaminergic neurons surviving in vitro. Therefore, if a polypeptide of the invention acts to prolong the survival of dopaminergic neurons, it would suggest that the polypeptide may be involved in Parkinson's Disease.

[1144] One skilled in the art could easily modify the exemplified studies to test the activity of polynucleotides of the invention (e.g., gene therapy), agonists, and/or antagonists of polynucleotides or polypeptides of the invention.

Example 38

Site Directed/Site-Specific Mutagenesis

[1145] In vitro site-directed mutagenesis is an invaluable technique for studying protein structure-function relationships and gene expression, for example, as well as for vector

modification. Site-directed mutagenesis can also be used for creating any of one or more of the mutants of the present invention, particularly the conservative and/or non-conservative amino acid substitution mutants of the prsent invention. Approaches utilizing single stranded DNA (ssDNA) as the template have been reported (e.g., T. A. Kunkel et al., 1985, *Proc. Natl. Acad. Sci. USA*), 82:488-492; M. A. Vandeyar et al., 1988, *Gene,* 65(1):129-133; M. Sugimoto et al., 1989, *Anal. Biochem.,* 179(2):309-311; and J. W. Taylor et al., 1985, *Nuc. Acids. Res.,* 13(24):8765-8785).

[1146] The use of PCR in site-directed mutagenesis accomplishes strand separation by using a denaturing step to separate the complementary strands and to allow efficient polymerization of the PCR primers. PCR site-directed mutagenesis methods thus permit site specific mutations to be incorporated in virtually any double stranded plasmid, thus eliminating the need for re-subcloning into M13-based bacteriophage vectors or single-stranded rescue. (M. P. Weiner et al., 1995, *Molecular Biology: Current Innovations and Future Trends,* Eds. A. M. Griffin and H. G. Griffin, Horizon Scientific Press, Norfolk, UK; and C. Papworth et al., 1996, *Strategies,* 9(3):3-4).

[1147] A protocol for performing site-directed mutagenesis, particularly employing the QuikChange™ site-directed mutagenesis kit (Stratagene, La Jolla, Calif.; U.S. Pat. Nos. 5,789,166 and 5,923,419) is provided for making point mutations, to switch or substitute amino acids, and to delete or insert single or multiple amino acids in the RATL1d6 amino acid sequence of this invention.

Primer Design

[1148] For primer design using this protocol, the mutagenic oligonucleotide primers are designed individually according to the desired mutation. The following considerations should be made for designing mutagenic primers: 1) Both of the mutagenic primers must contain the desired mutation and anneal to the same sequence on opposite strands of the plasmid; 2) Primers should be between 25 and 45 bases in length, and the melting temperature (T_m) of the primers should be greater than, or equal to, 78° C. The following formula is commonly used for estimating the T_m of primers: T=81.5+0.41 (%GC)−675/N−%mismatch. For calculating T_m, N is the primer length in bases; and values for %GC and % mismatch are whole numbers. For calculating T_m for primers intended to introduce insertions or deletions, a modified version of the above formula is employed: T=81.5+0.41 (%GC)−675/N, where N does not include the bases which are being inserted or deleted; 3) The desired mutation (deletion or insertion) should be in the middle of the primer with approximately 10-15 bases of correct sequence on both sides; 4) The primers optimally should have a minimum GC content of 40%, and should terminate in one or more C or G bases; 5) Primers need not be 5'-phosphorylated, but must be purified either by fast polynucleotide liquid chromatography (FPLC) or by polyacrylamide gel electrophoresis (PAGE). Failure to purify the primers results in a significant decrease in mutation efficiency; and 6) It is important that primer concentration is in excess. It is suggested to vary the amount of template while keeping the concentration of the primers constantly in excess (QuikChange™ Site-Directed Mutagenesis Kit, Stratagene, La Jolla, Calif.).

Protocol for Setting up the Reactions

[1149] Using the above-described primer design, two complimentary oligonucleotides containing the desired mutation, flanked by unmodified nucleic acid sequence, are synthesized. The resulting oligonucleotide primers are purified.

[1150] A control reaction is prepared using 5 μl 10× reaction buffer (100 mM KCl; 100 mM $(NH_4)_2SO_4$; 200 mM Tris-HCl, pH 8.8; 20 mM $MgSO_4$; 1% Triton® X-100; 1 mg/ml nuclease-free bovine serum albumin, BSA); 2 μl (10 ng) of pWhitescript™, 4.5-kb control plasmid (5 ng/μl); 1.25 μl (125 ng) of oligonucleotide control primer #1 (34-mer, 100 ng/μl); 1.25 μl (125 ng) of oligonucleotide control primer #2 (34-mer, 100 ng/μl); 1 μl of dNTP mix; double distilled H_2O; to a final volume of 50 μl. Thereafter, 1 μl of DNA polymerase (PfuTurbo® DNA Polymerase, Stratagene), (2.5U/μl) is added. PfuTurbo® DNA Polymerase is stated to have 6-fold higher fidelity in DNA synthesis than does Taq polymerase. To maximize temperature cycling performance, use of thin-walled test tubes is suggested to ensure optimum contact with the heating blocks of the temperature cycler.

[1151] The sample reaction is prepared by combining 5 μl of 10×reaction buffer; ×μl (5-50 ng) of dsDNA template; ×μl (125 ng) of oligonucleotide primer #1; ×μl (5-50 ng) of dsDNA template; ×μl (125 ng) of oligonucleotide primer #2; 1 μl of dNTP mix; and ddH_2O to a final volume of 50 μl. Thereafter, 1 μl of DNA polymerase (PfuTurbo DNA Polymerase, Stratagene), (2.5U/μl) is added.

[1152] It is suggested that if the thermal cycler does not have a hot-top assembly, each reaction should be overlaid with approximately 30 μl of mineral oil.

Cycling the Reactions

[1153] Each reaction is cycled using the following cycling parameters:

Segment	Cycles	Temperature	Time
1	1	95° C.	30 seconds
2	12–18	95° C.	30 seconds
		55° C.	1 minute
		68° C.	2 minutes/kb of plasmid length

[1154] For the control reaction, a 12-minute extension time is used and the reaction is run for 12 cycles. Segment 2 of the above cycling parameters is adjusted in accordance with the type of mutation desired. For example, for point mutations, 12 cycles are used; for single amino acid changes, 16 cycles are used; and for multiple amino acid deletions or insertions, 18 cycles are used. Following the temperature cycling, the reaction is placed on ice for 2 minutes to cool the reaction to ≦37° C.

Digesting the Products and Transforming Competent Cells

[1155] One μl of the DpnI restriction enzyme (10U/μl) is added directly (below mineral oil overlay) to each amplification reaction using a small, pointed pipette tip. The

reaction mixture is gently and thoroughly mixed by pipetting the solution up and down several times. The reaction mixture is then centrifuged for 1 minute in a microcentrifuge. Immediately thereafter, each reaction is incubated at 37° C. for 1 hour to digest the parental (i.e., the non-mutated) supercoiled dsDNA.

[1156] Competent cells (i.e., XL1-Blue supercompetent cells, Stratagene) are thawed gently on ice. For each control and sample reaction to be transformed, 50 μl of the supercompetent cells are aliquotted to a prechilled test tube (Falcon 2059 polypropylene). Next, 1 μl of the DpnI-digested DNA is transferred from the control and the sample reactions to separate aliquots of the supercompetent cells. The transformation reactions are gently swirled to mix and incubated for 30 minutes on ice. Thereafter, the transformation reactions are heat-pulsed for 45 seconds at 42° C. for 2 minutes.

[1157] 0.5 ml of NZY+ broth, preheated to 42° C. is added to the transformation reactions which are then incubated at 37° C. for 1 hour with shaking at 225-250 rpm. An aliquot of each transformation reaction is plated on agar plates containing the appropriate antibiotic for the vector. For the mutagenesis and transformation controls, cells are spead on LB-ampicillin agar plates containing 80 μg/ml of X-gal and 20 mM MIPTG. Transformation plates are incubated for >16 hours at 37° C.

[1158] It will be clear that the invention may be practiced otherwise than as particularly described in the foregoing description and examples. Numerous modifications and variations of the present invention are possible in light of the above teachings and, therefore, are within the scope of the appended claims.

[1159] The entire disclosure of each document cited (including patents, patent applications, journal articles, abstracts, laboratory manuals, books, or other disclosures) in the Background of the Invention, Detailed Description, and Examples is hereby incorporated herein by reference. Further, the hard copy of the sequence listing submitted herewith and the corresponding computer readable form are both incorpporated herein by reference in their entireties.

TABLE IV

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1	MET1	N	−23.219	27.222	−6.947
2	MET1	CA	−24.314	28.157	−6.658
3	MET1	CB	−24.159	29.428	−7.493
4	MET1	CG	−23.868	29.17	−8.973
5	MET1	SD	−25.216	28.55	−10.012
6	MET1	CE	−24.349	28.472	−11.592
7	MET1	C	−25.672	27.5	−6.882
8	MET1	O	−26.713	28.12	−6.649
9	ALA2	N	−25.652	26.222	−7.228
10	ALA2	CA	−26.893	25.482	−7.493
11	ALA2	CB	−26.572	24.399	−8.519
12	ALA2	C	−27.492	24.845	−6.237
13	ALA2	O	−28.583	24.258	−6.293
14	SER3	N	−26.787	24.989	−5.124
15	SER3	CA	−27.201	24.453	−3.819
16	SER3	CB	−28.282	25.375	−3.263
17	SER3	OG	−27.76	26.698	−3.264
18	SER3	C	−27.747	23.028	−3.89
19	SER3	O	−27.383	22.236	−4.77
20	SER4	N	−28.751	22.778	−3.067
21	SER4	CA	−29.325	21.433	−2.959

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
22	SER4	CB	−30.19	21.378	−1.707
23	SER4	OG	−29.369	21.689	−0.591
24	SER4	C	−30.179	21.03	−4.161
25	SER4	O	−30.259	19.83	−4.443
26	SER5	N	−30.575	21.977	−4.998
27	SER5	CA	−31.382	21.596	−6.163
28	SER5	CB	−32.205	22.787	−6.652
29	SER5	OG	−31.332	23.842	−7.035
30	SER5	C	−30.482	21.062	−7.275
31	SER5	O	−30.784	20.009	−7.849
32	GLY6	N	−29.266	21.583	−7.323
33	GLY6	CA	−28.263	21.1	−8.268
34	GLY6	C	−27.743	19.75	−7.807
35	GLY6	O	−27.749	18.785	−8.58
36	ARG7	N	−27.517	19.64	−6.508
37	ARG7	CA	−27.02	18.392	−5.922
38	ARG7	CB	−26.716	18.67	−4.458
39	ARG7	CG	−25.49	19.566	−4.333
40	ARG7	CD	−25.391	20.185	−2.946
41	ARG7	NE	−25.564	19.174	−1.893
42	ARG7	CZ	−25.959	19.498	−0.661
43	ARG7	NH1	−26.167	20.779	−0.346
44	ARG7	NH2	−26.121	18.549	0.262
45	ARG7	C	−28.007	17.231	−6.038
46	ARG7	O	−27.595	16.157	−6.494
47	VAL8	N	−29.297	17.491	−5.899
48	VAL8	CA	−30.282	16.413	−6.044
49	VAL8	CB	−31.583	16.84	−5.367
50	VAL8	CG1	−32.726	15.878	−5.676
51	VAL8	CG2	−31.392	16.971	−3.86
52	VAL8	C	−30.535	16.062	−7.51
53	VAL8	O	−30.655	14.872	−7.83
54	THR9	N	−30.317	17.014	−8.402
55	THR9	CA	−30.462	16.724	−9.831
56	THR9	CB	−30.533	18.038	−10.6
57	THR9	OG1	−31.708	18.721	−10.185
58	THR9	CG2	−30.638	17.806	−12.104
59	THR9	C	−29.291	15.889	−10.339
60	THR9	O	−29.523	14.85	−10.969
61	ILE10	N	−28.105	16.155	−9.813
62	ILE10	CA	−26.923	15.379	−10.201
63	ILE10	CB	−25.676	16.157	−9.79
64	ILE10	CG2	−24.411	15.343	−10.046
65	ILE10	CG1	−25.609	17.487	−10.529
66	ILE10	CD1	−24.402	18.306	−10.085
67	ILE10	C	−26.916	14.005	−9.539
68	ILE10	O	−26.604	13.012	−10.21
69	GLN11	N	−27.539	13.904	−8.377
70	GLN11	CA	−27.626	12.609	−7.71
71	GLN11	CB	−27.967	12.841	−6.244
72	GLN11	CG	−27.89	11.552	−5.435
73	GLN11	CD	−28.093	11.869	−3.957
74	GLN11	OE1	−28.865	12.772	−3.608
75	GLN11	NE2	−27.338	11.188	−3.112
76	GLN11	C	−28.672	11.719	−8.378
77	GLN11	O	−28.369	10.542	−8.61
78	LEU12	N	−29.691	12.326	−8.968
79	LEU12	CA	−30.696	11.564	−9.726
80	LEU12	CB	−31.94	12.431	−9.922
81	LEU12	CG	−33.088	12.085	−8.971
82	LEU12	CD1	−32.719	12.238	−7.498
83	LEU12	CD2	−34.314	12.934	−9.29
84	LEU12	C	−30.159	11.144	−11.092
85	LEU12	O	−30.362	9.993	−11.506
86	VAL13	N	−29.266	11.955	−11.636
87	VAL13	CA	−28.581	11.602	−12.881
88	VAL13	CB	−27.833	12.832	−13.389
89	VAL13	CG1	−26.809	12.482	−14.463
90	VAL13	CG2	−28.803	13.894	−13.891
91	VAL13	C	−27.611	10.447	−12.663
92	VAL13	O	−27.697	9.444	−13.389
93	ASP14	N	−26.974	10.428	−11.503
94	ASP14	CA	−26.041	9.346	−11.181
95	ASP14	CB	−25.217	9.74	−9.957

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
96	ASP14	CG	−24.411	11.014	−10.206
97	ASP14	OD1	−24.109	11.297	−11.36
98	ASP14	OD2	−24.1	11.686	−9.23
99	ASP14	C	−26.791	8.05	−10.891
100	ASP14	O	−26.469	7.024	−11.502
101	GLU15	N	−27.972	8.191	−10.309
102	GLU15	CA	−28.817	7.038	−9.978
103	GLU15	CB	−29.831	7.486	−8.933
104	GLU15	CG	−29.147	7.799	−7.61
105	GLU15	CD	−30.081	8.604	−6.712
106	GLU15	OE1	−29.8	8.663	−5.522
107	GLU15	OE2	−30.933	9.295	−7.254
108	GLU15	C	−29.566	6.447	−11.173
109	GLU15	O	−30.143	5.362	−11.037
110	GLU16	N	−29.559	7.112	−12.317
111	GLU16	CA	−30.163	6.485	−13.492
112	GLU16	CB	−31.087	7.474	−14.205
113	GLU16	CG	−30.418	8.802	−14.523
114	GLU16	CD	−31.449	9.839	−14.954
115	GLU16	OE1	−31.251	11.003	−14.627
116	GLU16	OE2	−32.417	9.455	−15.595
117	GLU16	C	−29.115	5.857	−14.417
118	GLU16	O	−29.457	4.864	−15.07
119	ALA17	N	−27.875	6.349	−14.386
120	ALA17	CA	−26.72	5.73	−15.095
121	ALA17	CB	−27.073	5.274	−16.508
122	ALA17	C	−25.523	6.675	−15.218
123	ALA17	O	−24.506	6.323	−15.834
124	GLY18	N	−25.669	7.868	−14.668
125	GLY18	CA	−24.639	8.905	−14.78
126	GLY18	C	−23.425	8.599	−13.915
127	GLY18	O	−23.527	7.885	−12.911
128	VAL19	N	−22.288	9.103	−14.38
129	VAL19	CA	−20.939	8.947	−13.791
130	VAL19	CB	−20.773	9.665	−12.442
131	VAL19	CG1	−21.21	11.117	−12.548
132	VAL19	CG2	−21.444	9.003	−11.239
133	VAL19	C	−20.526	7.48	−13.716
134	VAL19	O	−19.763	7.069	−12.83
135	GLY20	N	−20.856	6.762	−14.773
136	GLY20	CA	−20.532	5.339	−14.848
137	GLY20	C	−19.659	5.09	−16.064
138	GLY20	O	−19.279	6.033	−16.763
139	ALA21	N	−19.411	3.824	−16.356
140	ALA21	CA	−18.594	3.477	−17.526
141	ALA21	CB	−18.048	2.065	−17.344
142	ALA21	C	−19.399	3.562	−18.824
143	ALA21	O	−18.826	3.693	−19.911
144	GLY22	N	−20.715	3.571	−18.692
145	GLY22	CA	−21.59	3.815	−19.838
146	GLY22	C	−21.728	5.319	−20.034
147	GLY22	O	−21.146	5.896	−20.961
148	ARG23	N	−22.453	5.949	−19.126
149	ARG23	CA	−22.61	7.404	−19.177
150	ARG23	CB	−24.003	7.763	−18.681
151	ARG23	CG	−25.057	7.411	−19.722
152	ARG23	CD	−26.464	7.538	−19.154
153	ARG23	NE	−26.633	8.796	−18.413
154	ARG23	CZ	−27.738	9.078	−17.72
155	ARG23	NH1	−28.773	8.235	−17.746
156	ARG23	NH2	−27.827	10.222	−17.041
157	ARG23	C	−21.55	8.108	−18.341
158	ARG23	O	−21.786	8.465	−17.179
159	LEU24	N	−20.397	8.307	−18.96
160	LEU24	CA	−19.281	9.032	−18.341
161	LEU24	CB	−18.12	9.048	−19.33
162	LEU24	CG	−17.637	7.648	−19.689
163	LEU24	CD1	−16.82	7.673	−20.975
164	LEU24	CD2	−16.837	7.024	−18.552
165	LEU24	C	−19.662	10.475	−18.042
166	LEU24	O	−20.497	11.077	−18.729
167	GLN25	N	−19.074	11.013	−16.992
168	GLN25	CA	−19.302	12.421	−16.665
169	GLN25	CB	−19.119	12.608	−15.165

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
170	GLN25	CG	−19.425	14.039	−14.733
171	GLN25	CD	−19.393	14.13	−13.211
172	GLN25	OE1	−18.466	13.627	−12.566
173	GLN25	NE2	−20.448	14.697	−12.651
174	GLN25	C	−18.324	13.306	−17.434
175	GLN25	O	−17.127	13.34	−17.124
176	LEU26	N	−18.824	13.951	−18.476
177	LEU26	CA	−17.998	14.862	−19.284
178	LEU26	CB	−18.845	15.397	−20.432
179	LEU26	CG	−19.276	14.277	−21.371
180	LEU26	CD1	−20.339	14.763	−22.35
181	LEU26	CD2	−18.077	13.689	−22.111
182	LEU26	C	−17.502	16.026	−18.437
183	LEU26	O	−18.304	16.775	−17.86
184	PHE27	N	−16.189	16.191	−18.398
185	PHE27	CA	−15.587	17.179	−17.496
186	PHE27	CB	−14.079	16.966	−17.431
187	PHE27	CG	−13.421	17.604	−16.207
188	PHE27	CD1	−14.148	17.756	−15.031
189	PHE27	CE1	−13.556	18.332	−13.914
190	PHE27	CZ	−12.233	18.75	−13.97
191	PHE27	CE2	−11.503	18.591	−15.142
192	PHE27	CD2	−12.097	18.018	−16.259
193	PHE27	C	−15.917	18.597	−17.951
194	PHE27	O	−15.666	18.993	−19.097
195	ARG28	N	−16.65	19.27	−17.074
196	ARG28	CA	−17.158	20.637	−17.272
197	ARG28	CB	−15.986	21.61	−17.295
198	ARG28	CG	−15.268	21.624	−15.953
199	ARG28	CD	−16.187	22.1	−14.834
200	ARG28	NE	−15.491	22.077	−13.538
201	ARG28	CZ	−15.097	23.182	−12.901
202	ARG28	NH1	−14.461	23.085	−11.731
203	ARG28	NH2	−15.332	24.382	−13.44
204	ARG28	C	−17.979	20.79	−18.551
205	ARG28	O	−17.834	21.791	−19.262
206	GLY29	N	−18.758	19.771	−18.884
207	GLY29	CA	−19.621	19.819	−20.068
208	GLY29	C	−18.903	19.662	−21.414
209	GLY29	O	−19.552	19.832	−22.454
210	GLN30	N	−17.611	19.365	−21.424
211	GLN30	CA	−16.916	19.194	−22.706
212	GLN30	CB	−15.406	19.393	−22.546
213	GLN30	CG	−15.034	20.851	−22.267
214	GLN30	CD	−13.535	21.07	−22.499
215	GLN30	OE1	−13.051	22.21	−22.553
216	GLN30	NE2	−12.841	19.971	−22.738
217	GLN30	C	−17.208	17.827	−23.32
218	GLN30	O	−16.664	16.798	−22.899
219	SER31	N	−18.143	17.833	−24.255
220	SER31	CA	−18.476	16.635	−25.03
221	SER31	CB	−19.826	16.845	−25.707
222	SER31	OG	−19.932	15.896	−26.767
223	SER31	C	−17.449	16.379	−26.115
224	SER31	O	−17.356	17.163	−27.066
225	TYR32	N	−16.888	15.181	−26.108
226	TYR32	CA	−15.912	14.817	−27.138
227	TYR32	CB	−15.343	13.441	−26.835
228	TYR32	CG	−14.357	12.958	−27.894
229	TYR32	CD1	−13.128	13.59	−28.018
230	TYR32	CE1	−12.217	13.153	−28.968
231	TYR32	CZ	−12.537	12.092	−29.8
232	TYR32	OH	−11.574	11.587	−30.648
233	TYR32	CE2	−13.775	11.472	−29.698
234	TYR32	CD2	−14.689	11.909	−28.746
235	TYR32	C	−16.541	14.765	−28.521
236	TYR32	O	−16.008	15.383	−29.446
237	GLU33	N	−17.774	14.292	−28.595
238	GLU33	CA	−18.443	14.191	−29.89
239	GLU33	CB	−19.678	13.318	−29.732
240	GLU33	CG	−19.29	11.889	−29.367
241	GLU33	CD	−20.544	11.061	−29.12
242	GLU33	OE1	−20.485	9.857	−29.325
243	GLU33	OE2	−21.499	11.636	−28.617

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
244	GLU33	C	−18.83	15.555	−30.456
245	GLU33	O	−18.511	15.818	−31.621
246	ALA34	N	−19.263	16.48	−29.611
247	ALA34	CA	−19.62	17.805	−30.129
248	ALA34	CB	−20.478	18.52	−29.093
249	ALA34	C	−18.398	18.66	−30.456
250	ALA34	O	−18.381	19.319	−31.504
251	ILE35	N	−17.315	18.452	−29.726
252	ILE35	CA	−16.096	19.216	−29.983
253	ILE35	CB	−15.21	19.132	−28.748
254	ILE35	CG2	−13.832	19.705	−29.04
255	ILE35	CG1	−15.85	19.855	−27.568
256	ILE35	CD1	−14.988	19.739	−26.317
257	ILE35	C	−15.348	18.68	−31.197
258	ILE35	O	−14.957	19.471	−32.065
259	ARG36	N	−15.42	17.374	−31.398
260	ARG36	CA	−14.774	16.754	−32.551
261	ARG36	CB	−14.793	15.247	−32.354
262	ARG36	CG	−14.127	14.516	−33.512
263	ARG36	CD	−14.626	13.08	−33.58
264	ARG36	NE	−16.078	13.077	−33.829
265	ARG36	CZ	−16.962	12.425	−33.071
266	ARG36	NH1	−16.544	11.698	−32.033
267	ARG36	NH2	−18.263	12.489	−33.36
268	ARG36	C	−15.531	17.086	−33.827
269	ARG36	O	−14.906	17.524	−34.8
270	ALA37	N	−16.849	17.146	−33.728
271	ALA37	CA	−17.665	17.501	−34.888
272	ALA37	CB	−19.132	17.258	−34.552
273	ALA37	C	−17.463	18.959	−35.284
274	ALA37	O	−17.076	19.213	−36.432
275	ALA38	N	−17.411	19.846	−34.304
276	ALA38	CA	−17.226	21.271	−34.602
277	ALA38	CB	−17.433	22.068	−33.319
278	ALA38	C	−15.84	21.574	−35.17
279	ALA38	O	−15.747	22.134	−36.273
280	CYS39	N	−14.821	20.957	−34.595
281	CYS39	CA	−13.452	21.201	−35.054
282	CYS39	CB	−12.489	20.623	−34.025
283	CYS39	SG	−12.471	21.467	−32.427
284	CYS39	C	−13.183	20.582	−36.42
285	CYS39	O	−12.77	21.313	−37.33
286	LEU40	N	−13.724	19.398	−36.655
287	LEU40	CA	−13.49	18.715	−37.929
288	LEU40	CB	−13.858	17.247	−37.743
289	LEU40	CG	−13.506	16.404	−38.962
290	LEU40	CD1	−12.008	16.459	−39.242
291	LEU40	CD2	−13.961	14.963	−38.769
292	LEU40	C	−14.322	19.306	−39.068
293	LEU40	O	−13.798	19.462	−40.176
294	ASP41	N	−15.474	19.871	−38.74
295	ASP41	CA	−16.336	20.463	−39.77
296	ASP41	CB	−17.78	20.483	−39.272
297	ASP41	CG	−18.328	19.073	−39.051
298	ASP41	OD1	−19.305	18.957	−38.32
299	ASP41	OD2	−17.863	18.163	−39.725
300	ASP41	C	−15.93	21.891	−40.13
301	ASP41	O	−16.357	22.403	−41.171
302	SER42	N	−15.11	22.52	−39.304
303	SER42	CA	−14.64	23.866	−39.636
304	SER42	CB	−14.818	24.745	−38.406
305	SER42	OG	−14.433	26.066	−38.757
306	SER42	C	−13.176	23.87	−40.082
307	SER42	O	−12.703	24.851	−40.669
308	GLY43	N	−12.477	22.781	−39.809
309	GLY43	CA	−11.063	22.66	−40.182
310	GLY43	C	−10.176	23.293	−39.111
311	GLY43	O	−9.09	23.812	−39.4
312	ILE44	N	−10.647	23.234	−37.878
313	ILE44	CA	−9.96	23.896	−36.765
314	ILE44	CB	−10.979	24.782	−36.044
315	ILE44	CG2	−10.407	25.399	−34.772
316	ILE44	CG1	−11.487	25.883	−36.967
317	ILE44	CD1	−12.422	26.832	−36.225

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
318	ILE44	C	−9.357	22.887	−35.793
319	ILE44	O	−10.041	21.97	−35.327
320	LEU45	N	−8.064	23.03	−35.547
321	LEU45	CA	−7.404	22.238	−34.502
322	LEU45	CB	−5.897	22.422	−34.611
323	LEU45	CG	−5.349	21.814	−35.897
324	LEU45	CD1	−3.88	22.172	−36.087
325	LEU45	CD2	−5.544	20.302	−35.917
326	LEU45	C	−7.893	22.684	−33.126
327	LEU45	O	−8.017	23.883	−32.846
328	PHE46	N	−8.213	21.702	−32.304
329	PHE46	CA	−8.801	21.949	−30.983
330	PHE46	CB	−9.252	20.593	−30.453
331	PHE46	CG	−9.927	20.596	−29.085
332	PHE46	CD1	−10.8	21.617	−28.734
333	PHE46	CE1	−11.413	21.607	−27.489
334	PHE46	CZ	−11.159	20.573	−26.597
335	PHE46	CE2	−10.291	19.548	−26.951
336	PHE46	CD2	−9.678	19.559	−28.196
337	PHE46	C	−7.856	22.589	−29.967
338	PHE46	O	−6.921	21.956	−29.469
339	ARG47	N	−8.105	23.852	−29.67
340	ARG47	CA	−7.498	24.46	−28.484
341	ARG47	CB	−7.164	25.925	−28.735
342	ARG47	CG	−6.451	26.51	−27.521
343	ARG47	CD	−6.074	27.974	−27.706
344	ARG47	NE	−5.288	28.435	−26.55
345	ARG47	CZ	−4.115	29.061	−26.668
346	ARG47	NH1	−3.371	29.288	−25.583
347	ARG47	NH2	−3.622	29.327	−27.88
348	ARG47	C	−8.505	24.332	−27.345
349	ARG47	O	−9.638	24.814	−27.456
350	ASP48	N	−8.114	23.63	−26.296
351	ASP48	CA	−9.039	23.309	−25.204
352	ASP48	CB	−8.329	22.323	−24.282
353	ASP48	CG	−9.336	21.58	−23.419
354	ASP48	OD1	−9.876	22.204	−22.518
355	ASP48	OD2	−9.685	20.472	−23.799
356	ASP48	C	−9.47	24.552	−24.422
357	ASP48	O	−8.65	25.217	−23.778
358	PRO49	N	−10.776	24.781	−24.384
359	PRO49	CA	−11.324	25.989	−23.753
360	PRO49	CB	−12.754	26.037	−24.198
361	PRO49	CG	−13.098	24.753	−24.939
362	PRO49	CD	−11.808	23.957	−25.016
363	PRO49	C	−11.243	25.979	−22.223
364	PRO49	O	−11.067	27.042	−21.617
365	TYR50	N	−11.23	24.8	−21.619
366	TYR50	CA	−11.064	24.697	−20.167
367	TYR50	CB	−11.918	23.528	−19.677
368	TYR50	CG	−11.935	23.344	−18.163
369	TYR50	CD1	−11.692	22.091	−17.619
370	TYR50	CE1	−11.681	21.922	−16.241
371	TYR50	CZ	−11.919	23.008	−15.412
372	TYR50	OH	−11.841	22.851	−14.047
373	TYR50	CE2	−12.177	24.261	−15.952
374	TYR50	CD2	−12.186	24.428	−17.331
375	TYR50	C	−9.591	24.498	−19.778
376	TYR50	O	−9.243	24.509	−18.591
377	PHE51	N	−8.724	24.411	−20.773
378	PHE51	CA	−7.299	24.186	−20.51
379	PHE51	CB	−7.104	22.685	−20.339
380	PHE51	CG	−5.782	22.252	−19.715
381	PHE51	CD1	−5.371	22.804	−18.508
382	PHE51	CE1	−4.178	22.393	−17.928
383	PHE51	CZ	−3.398	21.428	−18.552
384	PHE51	CE2	−3.809	20.876	−19.757
385	PHE51	CD2	−5.002	21.286	−20.338
386	PHE51	C	−6.438	24.672	−21.674
387	PHE51	O	−6.063	23.869	−22.54
388	PRO52	N	−6.137	25.962	−21.702
389	PRO52	CA	−5.186	26.487	−22.683
390	PRO52	CB	−5.291	27.976	−22.568
391	PRO52	CG	−6.118	28.321	−21.338

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
392	PRO52	CD	−6.577	26.994	−20.757
393	PRO52	C	−3.774	26.009	−22.36
394	PRO52	O	−3.399	25.882	−21.188
395	ALA53	N	−3.002	25.729	−23.393
396	ALA53	CA	−1.629	25.272	−23.168
397	ALA53	CB	−1.176	24.433	−24.353
398	ALA53	C	−0.675	26.443	−22.939
399	ALA53	O	−0.268	27.144	−23.872
400	GLY54	N	−0.351	26.653	−21.676
401	GLY54	CA	0.602	27.701	−21.294
402	GLY54	C	1.391	27.298	−20.053
403	GLY54	O	1.34	26.144	−19.612
404	PRO55	N	2.041	28.271	−19.437
405	PRO55	CA	2.861	27.998	−18.248
406	PRO55	CB	3.629	29.262	−18.016
407	PRO55	CG	3.131	30.338	−18.972
408	PRO55	CD	2.083	29.675	−19.852
409	PRO55	C	2.032	27.622	−17.013
410	PRO55	O	2.483	26.805	−16.204
411	ASP56	N	0.758	27.988	−17.016
412	ASP56	CA	−0.172	27.598	−15.948
413	ASP56	CB	−1.298	28.626	−15.855
414	ASP56	CG	−2.192	28.598	−17.096
415	ASP56	OD1	−1.756	29.099	−18.126
416	ASP56	OD2	−3.276	28.045	−17.002
417	ASP56	C	−0.763	26.2	−16.163
418	ASP56	O	−1.468	25.686	−15.289
419	ALA57	N	−0.438	25.574	−17.286
420	ALA57	CA	−0.823	24.182	−17.52
421	ALA57	CB	−0.965	23.956	−19.021
422	ALA57	C	0.277	23.289	−16.958
423	ALA57	O	0.041	22.146	−16.544
424	LEU58	N	1.453	23.881	−16.833
425	LEU58	CA	2.533	23.248	−16.09
426	LEU58	CB	3.839	23.956	−16.42
427	LEU58	CG	4.172	23.876	−17.903
428	LEU58	CD1	5.337	24.796	−18.25
429	LEU58	CD2	4.473	22.44	−18.318
430	LEU58	C	2.21	23.415	−14.615
431	LEU58	O	2.002	22.425	−13.907
432	GLY59	N	1.955	24.649	−14.221
433	GLY59	CA	1.501	24.921	−12.855
434	GLY59	C	1.7	26.383	−12.485
435	GLY59	O	1.318	27.296	−13.226
436	TYR60	N	2.292	26.586	−11.321
437	TYR60	CA	2.628	27.936	−10.857
438	TYR60	CB	1.388	28.623	−10.275
439	TYR60	CG	0.598	27.843	−9.22
440	TYR60	CD1	1.002	27.858	−7.89
441	TYR60	CE1	0.28	27.152	−6.936
442	TYR60	CZ	−0.85	26.44	−7.314
443	TYR60	OH	−1.569	25.742	−6.368
444	TYR60	CE2	−1.266	26.434	−8.638
445	TYR60	CD2	−0.543	27.14	−9.59
446	TYR60	C	3.765	27.9	−9.84
447	TYR60	O	3.854	26.971	−9.03
448	ASP61	N	4.667	28.862	−9.972
449	ASP61	CA	5.805	29.056	−9.054
450	ASP61	CB	5.302	29.696	−7.768
451	ASP61	CG	4.717	31.068	−8.087
452	ASP61	OD1	3.498	31.178	−8.085
453	ASP61	OD2	5.496	31.974	−8.353
454	ASP61	C	6.6	27.794	−8.731
455	ASP61	O	6.718	27.421	−7.555
456	GLN62	N	7.13	27.17	−9.777
457	GLN62	CA	7.936	25.941	−9.66
458	GLN62	CB	7.218	24.917	−8.788
459	GLN62	CG	8.096	23.733	−8.409
460	GLN62	CD	7.302	22.862	−7.448
461	GLN62	OE1	7.254	21.634	−7.586
462	GLN62	NE2	6.57	23.533	−6.577
463	GLN62	C	8.135	25.329	−11.04
464	GLN62	O	9.25	25.22	−11.56
465	LEU63	N	7.027	24.815	−11.547

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
466	LEU63	CA	6.982	24.227	−12.881
467	LEU63	CB	6.22	22.909	−12.772
468	LEU63	CG	6.106	22.172	−14.098
469	LEU63	CD1	7.479	21.869	−14.686
470	LEU63	CD2	5.305	20.889	−13.92
471	LEU63	C	6.257	25.194	−13.802
472	LEU63	O	6.527	25.264	−15.004
473	GLY64	N	5.444	26.033	−13.187
474	GLY64	CA	4.756	27.082	−13.936
475	GLY64	C	5.481	28.413	−13.82
476	GLY64	O	6.639	28.47	−13.383
477	PRO65	N	4.725	29.475	−14.043
478	PRO65	CA	5.299	30.797	−14.301
479	PRO65	CB	4.134	31.686	−14.614
480	PRO65	CG	2.852	30.873	−14.559
481	PRO65	CD	3.277	29.451	−14.243
482	PRO65	C	6.095	31.336	−13.121
483	PRO65	O	5.639	31.234	−11.973
484	ASP66	N	7.373	31.531	−13.431
485	ASP66	CA	8.414	32.232	−12.639
486	ASP66	CB	8.064	32.464	−11.162
487	ASP66	CG	8.269	31.22	−10.288
488	ASP66	OD1	7.932	30.128	−10.725
489	ASP66	OD2	8.749	31.392	−9.178
490	ASP66	C	9.707	31.425	−12.713
491	ASP66	O	10.777	31.872	−12.286
492	SER67	N	9.582	30.227	−13.258
493	SER67	CA	10.687	29.277	−13.228
494	SER67	CB	10.088	27.906	−12.931
495	SER67	OG	11.132	26.94	−12.911
496	SER67	C	11.458	29.222	−14.534
497	SER67	O	10.884	29.293	−15.628
498	GLU68	N	12.72	28.851	−14.396
499	GLU68	CA	13.574	28.564	−15.547
500	GLU68	CB	15.01	28.489	−15.031
501	GLU68	CG	16.024	28.198	−16.132
502	GLU68	CD	17.433	28.21	−15.55
503	GLU68	OE1	17.768	27.248	−14.874
504	GLU68	OE2	18.071	29.249	−15.639
505	GLU68	C	13.188	27.245	−16.236
506	GLU68	O	13.467	27.096	−17.428
507	LYS69	N	12.378	26.413	−15.589
508	LYS69	CA	11.91	25.177	−16.216
509	LYS69	CB	11.785	24.139	−15.103
510	LYS69	CG	11.52	22.733	−15.626
511	LYS69	CD	11.429	21.73	−14.482
512	LYS69	CE	11.194	20.32	−15.007
513	LYS69	NZ	12.28	19.911	−15.912
514	LYS69	C	10.555	25.401	−16.907
515	LYS69	O	10.064	24.531	−17.635
516	ALA70	N	10.004	26.593	−16.731
517	ALA70	CA	8.711	26.935	−17.327
518	ALA70	CB	7.907	27.72	−16.302
519	ALA70	C	8.85	27.795	−18.578
520	ALA70	O	7.866	28.01	−19.296
521	LYS71	N	10.031	28.339	−18.805
522	LYS71	CA	10.209	29.18	−19.989
523	LYS71	CB	11.152	30.335	−19.67
524	LYS71	CG	12.527	29.877	−19.205
525	LYS71	CD	13.388	31.082	−18.849
526	LYS71	CE	14.804	30.671	−18.469
527	LYS71	NZ	15.604	31.84	−18.074
528	LYS71	C	10.688	28.364	−21.186
529	LYS71	O	11.484	27.429	−21.06
530	GLY72	N	10.149	28.716	−22.34
531	GLY72	CA	10.457	27.999	−23.581
532	GLY72	C	9.408	26.918	−23.812
533	GLY72	O	9.731	25.742	−24.012
534	VAL73	N	8.155	27.34	−23.785
535	VAL73	CA	7.022	26.412	−23.878
536	VAL73	CB	5.889	26.964	−23.016
537	VAL73	CG1	6.09	26.623	−21.546
538	VAL73	CG2	5.707	28.467	−23.206
539	VAL73	C	6.54	26.162	−25.309

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
540	VAL73	O	5.549	26.744	−25.767
541	LYS74	N	7.195	25.224	−25.972
542	LYS74	CA	6.764	24.802	−27.307
543	LYS74	CB	7.985	24.378	−28.116
544	LYS74	CG	7.602	24.093	−29.563
545	LYS74	CD	7.07	25.355	−30.234
546	LYS74	CE	6.538	25.067	−31.632
547	LYS74	NZ	5.38	24.16	−31.574
548	LYS74	C	5.783	23.633	−27.196
549	LYS74	O	6.177	22.465	−27.081
550	TRP75	N	4.502	23.962	−27.248
551	TRP75	CA	3.442	22.945	−27.175
552	TRP75	CB	2.199	23.563	−26.552
553	TRP75	CG	2.389	24.062	−25.135
554	TRP75	CD1	2.665	25.354	−24.748
555	TRP75	NE1	2.741	25.391	−23.396
556	TRP75	CE2	2.52	24.176	−22.86
557	TRP75	CZ2	2.461	23.722	−21.552
558	TRP75	CH2	2.201	22.379	−21.299
559	TRP75	CZ3	1.997	21.493	−22.352
560	TRP75	CE3	2.048	21.944	−23.666
561	TRP75	CD2	2.304	23.281	−23.921
562	TRP75	C	3.102	22.399	−28.56
563	TRP75	O	2.059	22.717	−29.144
564	MET76	N	3.955	21.507	−29.026
565	MET76	CA	3.85	20.98	−30.387
566	MET76	CB	5.23	20.483	−30.796
567	MET76	CG	5.853	19.655	−29.678
568	MET76	SD	7.255	18.629	−30.162
569	MET76	CE	6.374	17.482	−31.247
570	MET76	C	2.851	19.839	−30.532
571	MET76	O	2.869	18.858	−29.783
572	ARG77	N	1.98	19.972	−31.513
573	ARG77	CA	1.131	18.842	−31.885
574	ARG77	CB	−0.002	19.332	−32.777
575	ARG77	CG	−0.802	20.424	−32.079
576	ARG77	CD	−2.031	20.811	−32.891
577	ARG77	NE	−2.709	21.975	−32.299
578	ARG77	CZ	−3.847	21.91	−31.606
579	ARG77	NH1	−4.429	20.729	−31.38
580	ARG77	NH2	−4.395	23.028	−31.125
581	ARG77	C	1.998	17.824	−32.617
582	ARG77	O	2.825	18.2	−33.459
583	PRO78	N	1.757	16.545	−32.364
584	PRO78	CA	2.754	15.509	−32.68
585	PRO78	CB	2.235	14.264	−32.037
586	PRO78	CG	0.904	14.552	−31.363
587	PRO78	CD	0.652	16.033	−31.555
588	PRO78	C	2.964	15.27	−34.175
589	PRO78	O	4.114	15.069	−34.59
590	HIS79	N	1.969	15.64	−34.968
591	HIS79	CA	2.011	15.483	−36.427
592	HIS79	CB	0.562	15.6	−36.901
593	HIS79	CG	0.314	15.364	−38.378
594	HIS79	ND1	0.14	14.175	−38.986
595	HIS79	CE1	−0.057	14.372	−40.305
596	HIS79	NE2	−0.013	15.705	−40.531
597	HIS79	CD2	0.209	16.33	−39.352
598	HIS79	C	2.897	16.532	−37.122
599	HIS79	O	3.215	16.391	−38.307
600	GLU80	N	3.4	17.497	−36.365
601	GLU80	CA	4.313	18.491	−36.929
602	GLU80	CB	4.312	19.71	−36.016
603	GLU80	CG	2.932	20.336	−35.882
604	GLU80	CD	2.967	21.364	−34.757
605	GLU80	OE1	4.064	21.731	−34.364
606	GLU80	OE2	1.919	21.58	−34.162
607	GLU80	C	5.746	17.968	−36.993
608	GLU80	O	6.524	18.413	−37.844
609	PHE81	N	6.083	17.021	−36.131
610	PHE81	CA	7.455	16.5	−36.128
611	PHE81	CB	8.089	16.687	−34.749
612	PHE81	CG	8.658	18.075	−34.437
613	PHE81	CD1	8.499	19.135	−35.323

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
614	PHE81	CE1	9.023	20.385	−35.02
615	PHE81	CZ	9.719	20.576	−33.834
616	PHE81	CE2	9.895	19.515	−32.957
617	PHE81	CD2	9.37	18.266	−33.26
618	PHE81	C	7.484	15.026	−36.508
619	PHE81	O	8.481	14.529	−37.047
620	CYS82	N	6.398	14.334	−36.221
621	CYS82	CA	6.31	12.927	−36.596
622	CYS82	CB	5.727	12.138	−35.43
623	CYS82	SG	5.806	10.34	−35.592
624	CYS82	C	5.433	12.787	−37.829
625	CYS82	O	4.218	13.011	−37.765
626	ALA83	N	6.028	12.253	−38.886
627	ALA83	CA	5.34	12.115	−40.18
628	ALA83	CB	6.392	11.917	−41.266
629	ALA83	C	4.342	10.956	−40.224
630	ALA83	O	3.53	10.862	−41.151
631	GLU84	N	4.404	10.086	−39.231
632	GLU84	CA	3.361	9.072	−39.05
633	GLU84	CB	3.832	7.759	−39.666
634	GLU84	CG	2.735	6.699	−39.642
635	GLU84	CD	3.304	5.356	−40.087
636	GLU84	OE1	2.837	4.348	−39.579
637	GLU84	OE2	4.268	5.371	−40.84
638	GLU84	C	3.089	8.862	−37.559
639	GLU84	O	3.667	7.944	−36.963
640	PRO85	N	2.305	9.743	−36.953
641	PRO85	CA	1.967	9.6	−35.536
642	PRO85	CB	1.36	10.912	−35.147
643	PRO85	CG	1.111	11.735	−36.399
644	PRO85	CD	1.67	10.92	−37.552
645	PRO85	C	0.979	8.464	−35.298
646	PRO85	O	0.093	8.199	−36.123
647	LYS86	N	1.212	7.733	−34.225
648	LYS86	CA	0.257	6.716	−33.797
649	LYS86	CB	0.879	5.327	−33.829
650	LYS86	CG	1.045	4.823	−35.255
651	LYS86	CD	1.563	3.391	−35.26
652	LYS86	CE	1.736	2.867	−36.68
653	LYS86	NZ	2.249	1.489	−36.667
654	LYS86	C	−0.24	7.011	−32.395
655	LYS86	O	0.429	7.643	−31.568
656	PHE87	N	−1.457	6.575	−32.15
657	PHE87	CA	−2.017	6.744	−30.82
658	PHE87	CB	−3.523	6.536	−30.909
659	PHE87	CG	−4.298	7.071	−29.713
660	PHE87	CD1	−3.871	8.228	−29.075
661	PHE87	CE1	−4.578	8.722	−27.988
662	PHE87	CZ	−5.716	8.061	−27.544
663	PHE87	CE2	−6.146	6.908	−28.186
664	PHE87	CD2	−5.438	6.413	−29.272
665	PHE87	C	−1.367	5.701	−29.931
666	PHE87	O	−0.396	5.991	−29.218
667	ILE88	N	−1.763	4.466	−30.183
668	ILE88	CA	−1.244	3.305	−29.457
669	ILE88	CB	−2.144	3.089	−28.237
670	ILE88	CG2	−3.62	3.068	−28.621
671	ILE88	CG1	−1.778	1.836	−27.457
672	ILE88	CD1	−2.72	1.645	−26.279
673	ILE88	C	−1.239	2.069	−30.36
674	ILE88	O	−2.213	1.801	−31.074
675	CYS89	N	−0.115	1.374	−30.404
676	CYS89	CA	−0.084	0.114	−31.153
677	CYS89	CB	1.14	0.035	−32.063
678	CYS89	SG	2.755	0.364	−31.335
679	CYS89	C	−0.186	−1.056	−30.181
680	CYS89	O	0.81	−1.533	−29.618
681	GLU90	N	−1.367	−1.654	−30.213
682	GLU90	CA	−1.825	−2.663	−29.232
683	GLU90	CB	−3.34	−2.783	−29.398
684	GLU90	CG	−3.702	−3.248	−30.805
685	GLU90	CD	−5.219	−3.3	−30.985
686	GLU90	OE1	−5.866	−2.401	−30.472
687	GLU90	OE2	−5.666	−4.134	−31.758

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
688	GLU90	C	−1.218	−4.07	−29.328
689	GLU90	O	−1.824	−5.019	−28.822
690	ASP91	N	−0.11	−4.235	−30.031
691	ASP91	CA	0.503	−5.56	−30.15
692	ASP91	CB	0.618	−5.895	−31.634
693	ASP91	CG	1.008	−7.356	−31.85
694	ASP91	OD1	0.119	−8.194	−31.832
695	ASP91	OD2	2.183	−7.607	−32.072
696	ASP91	C	1.877	−5.573	−29.479
697	ASP91	O	2.515	−6.628	−29.36
698	MET92	N	2.324	−4.415	−29.02
699	MET92	CA	3.662	−4.342	−28.415
700	MET92	CB	4.119	−2.888	−28.339
701	MET92	CG	4.419	−2.342	−29.731
702	MET92	SD	5.624	−3.305	−30.683
703	MET92	CE	7.019	−3.221	−29.528
704	MET92	C	3.732	−4.96	−27.029
705	MET92	O	3.024	−4.569	−26.094
706	SER93	N	4.632	−5.923	−26.921
707	SER93	CA	4.936	−6.545	−25.622
708	SER93	CB	5.176	−8.038	−25.818
709	SER93	OG	6.346	−8.188	−26.616
710	SER93	C	6.177	−5.905	−24.996
711	SER93	O	6.543	−6.214	−23.854
712	ARG94	N	6.841	−5.066	−25.774
713	ARG94	CA	7.962	−4.265	−25.277
714	ARG94	CB	9.058	−4.247	−26.339
715	ARG94	CG	9.996	−5.441	−26.232
716	ARG94	CD	10.773	−5.377	−24.925
717	ARG94	NE	11.797	−6.427	−24.836
718	ARG94	CZ	13.019	−6.194	−24.353
719	ARG94	NH1	13.345	−4.969	−23.937
720	ARG94	NH2	13.913	−7.184	−24.285
721	ARG94	C	7.514	−2.833	−25.009
722	ARG94	O	7.46	−2.023	−25.944
723	THR95	N	7.18	−2.519	−23.768
724	THR95	CA	6.776	−1.137	−23.477
725	THR95	CB	5.598	−1.071	−22.506
726	THR95	OG1	5.491	0.273	−22.051
727	THR95	CG2	5.771	−1.959	−21.282
728	THR95	C	7.965	−0.325	−22.987
729	THR95	O	8.294	−0.289	−21.797
730	ASP96	N	8.558	0.367	−23.949
731	ASP96	CA	9.767	1.184	−23.751
732	ASP96	CB	9.59	2.201	−22.621
733	ASP96	CG	8.466	3.166	−22.962
734	ASP96	OD1	8.688	3.984	−23.851
735	ASP96	OD2	7.35	2.914	−22.53
736	ASP96	C	10.993	0.306	−23.514
737	ASP96	O	10.921	−0.772	−22.914
738	VAL97	N	12.107	0.758	−24.062
739	VAL97	CA	13.387	0.058	−23.903
740	VAL97	CB	14.432	0.768	−24.758
741	VAL97	CG1	14.363	0.314	−26.212
742	VAL97	CG2	14.32	2.287	−24.633
743	VAL97	C	13.846	0.013	−22.448
744	VAL97	O	13.187	0.549	−21.551
745	CYS98	N	15.052	−0.495	−22.244
746	CYS98	CA	15.607	−0.62	−20.886
747	CYS98	CB	16.81	−1.558	−20.945
748	CYS98	SG	16.497	−3.163	−21.72
749	CYS98	C	16.039	0.74	−20.327
750	CYS98	O	16.093	0.941	−19.11
751	GLN99	N	16.205	1.701	−21.222
752	GLN99	CA	16.419	3.096	−20.834
753	GLN99	CB	17.525	3.695	−21.69
754	GLN99	CG	18.857	2.998	−21.46
755	GLN99	CD	19.937	3.727	−22.248
756	GLN99	OE1	20.773	3.103	−22.911
757	GLN99	NE2	19.875	5.046	−22.196
758	GLN99	C	15.152	3.931	−21.016
759	GLN99	O	15.246	5.113	−21.375
760	GLY100	N	13.988	3.335	−20.805
761	GLY100	CA	12.721	4.047	−21.006

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
762	GLY100	C	12.257	4.79	−19.755
763	GLY100	O	11.07	4.726	−19.416
764	SER101	N	13.093	5.72	−19.315
765	SER101	CA	12.902	6.439	−18.05
766	SER101	CB	14.276	6.861	−17.543
767	SER101	OG	15.071	5.688	−17.432
768	SER101	C	12.003	7.67	−18.173
769	SER101	O	11.77	8.364	−17.177
770	LEU102	N	11.524	7.951	−19.375
771	LEU102	CA	10.522	9.006	−19.543
772	LEU102	CB	10.591	9.56	−20.966
773	LEU102	CG	11.203	10.96	−21.039
774	LEU102	CD1	12.623	11.008	−20.481
775	LEU102	CD2	11.188	11.482	−22.471
776	LEU102	C	9.123	8.449	−19.279
777	LEU102	O	8.218	9.209	−18.912
778	GLY103	N	8.998	7.132	−19.332
779	GLY103	CA	7.722	6.47	−19.06
780	GLY103	C	7.735	5.838	−17.674
781	GLY103	O	8.654	5.097	−17.308
782	ASN104	N	6.721	6.176	−16.898
783	ASN104	CA	6.61	5.664	−15.529
784	ASN104	CB	5.805	6.653	−14.69
785	ASN104	CG	4.439	6.902	−15.32
786	ASN104	OD1	3.655	5.966	−15.533
787	ASN104	ND2	4.152	8.168	−15.559
788	ASN104	C	5.972	4.28	−15.477
789	ASN104	O	5.386	3.794	−16.455
790	CYS105	N	5.896	3.765	−14.262
791	CYS105	CA	5.411	2.403	−14.046
792	CYS105	CB	5.901	1.942	−12.68
793	CYS105	SG	7.699	1.954	−12.479
794	CYS105	C	3.891	2.238	−14.137
795	CYS105	O	3.45	1.106	−14.363
796	TRP106	N	3.113	3.31	−14.198
797	TRP106	CA	1.681	3.084	−14.401
798	TRP106	CB	0.794	4.045	−13.606
799	TRP106	CG	0.897	5.542	−13.829
800	TRP106	CD1	1.57	6.44	−13.03
801	TRP106	NE1	1.36	7.69	−13.515
802	TRP106	CE2	0.574	7.661	−14.606
803	TRP106	CZ2	0.029	8.664	−15.399
804	TRP106	CH2	−0.779	8.325	−16.478
805	TRP106	CZ3	−1.049	6.992	−16.763
806	TRP106	CE3	−0.52	5.982	−15.965
807	TRP106	CD2	0.281	6.316	−14.884
808	TRP106	C	1.345	3.056	−15.89
809	TRP106	O	0.395	2.368	−16.277
810	PHE107	N	2.265	3.532	−16.716
811	PHE107	CA	2.142	3.31	−18.161
812	PHE107	CB	2.936	4.363	−18.923
813	PHE107	CG	2.208	5.686	−19.116
814	PHE107	CD1	1.129	5.751	−19.987
815	PHE107	CE1	0.462	6.953	−20.177
816	PHE107	CZ	0.874	8.09	−19.496
817	PHE107	CE2	1.952	8.024	−18.623
818	PHE107	CD2	2.618	6.821	−18.433
819	PHE107	C	2.669	1.93	−18.53
820	PHE107	O	2.117	1.268	−19.418
821	LEU108	N	3.545	1.413	−17.686
822	LEU108	CA	4.072	0.058	−17.875
823	LEU108	CB	5.354	−0.08	−17.052
824	LEU108	CG	6.656	0.145	−17.834
825	LEU108	CD1	6.694	1.439	−18.644
826	LEU108	CD2	7.854	0.086	−16.894
827	LEU108	C	3.038	−0.977	−17.429
828	LEU108	O	2.778	−1.941	−18.161
829	ALA109	N	2.268	−0.627	−16.409
830	ALA109	CA	1.172	−1.493	−15.961
831	ALA109	CB	0.796	−1.105	−14.535
832	ALA109	C	−0.055	−1.377	−16.863
833	ALA109	O	−0.736	−2.383	−17.094
834	ALA110	N	−0.169	−0.26	−17.565
835	ALA110	CA	−1.238	−0.09	−18.55

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
836	ALA110	CB	−1.377	1.398	−18.84
837	ALA110	C	−0.933	−0.838	−19.846
838	ALA110	O	−1.847	−1.393	−20.464
839	ALA111	N	0.345	−1.073	−20.098
840	ALA111	CA	0.749	−1.897	−21.239
841	ALA111	CB	2.161	−1.511	−21.644
842	ALA111	C	0.683	−3.389	−20.915
843	ALA111	O	0.399	−4.199	−21.805
844	ALA112	N	0.688	−3.715	−19.631
845	ALA112	CA	0.417	−5.094	−19.203
846	ALA112	CB	1.023	−5.309	−17.82
847	ALA112	C	−1.092	−5.351	−19.157
848	ALA112	O	−1.548	−6.499	−19.205
849	SER113	N	−1.845	−4.265	−19.194
850	SER113	CA	−3.301	−4.288	−19.326
851	SER113	CB	−3.853	−3.165	−18.466
852	SER113	OG	−3.459	−3.417	−17.126
853	SER113	C	−3.751	−4.097	−20.778
854	SER113	O	−4.912	−3.736	−21.021
855	LEU114	N	−2.798	−4.185	−21.697
856	LEU114	CA	−3.055	−4.154	−23.144
857	LEU114	CB	−1.975	−3.335	−23.838
858	LEU114	CG	−2.284	−1.85	−23.772
859	LEU114	CD1	−1.128	−1.03	−24.332
860	LEU114	CD2	−3.576	−1.561	−24.527
861	LEU114	C	−3.074	−5.549	−23.76
862	LEU114	O	−2.955	−5.691	−24.981
863	THR115	N	−3.214	−6.563	−22.925
864	THR115	CA	−3.232	−7.946	−23.412
865	THR115	CB	−2.846	−8.874	−22.249
866	THR115	OG1	−1.894	−8.203	−21.435
867	THR115	CG2	−2.205	−10.191	−22.69
868	THR115	C	−4.644	−8.226	−23.948
869	THR115	O	−5.46	−7.305	−24.064
870	LEU116	N	−4.927	−9.471	−24.297
871	LEU116	CA	−6.275	−9.837	−24.752
872	LEU116	CB	−6.203	−11.217	−25.398
873	LEU116	CG	−5.206	−11.256	−26.551
874	LEU116	CD1	−4.949	−12.688	−27.006
875	LEU116	CD2	−5.666	−10.388	−27.718
876	LEU116	C	−7.217	−9.892	−23.551
877	LEU116	O	−8.428	−9.674	−23.662
878	TYR117	N	−6.618	−10.133	−22.398
879	TYR117	CA	−7.309	−10.051	−21.113
880	TYR117	CB	−8.165	−11.302	−20.889
881	TYR117	CG	−7.527	−12.648	−21.237
882	TYR117	CD1	−6.389	−13.091	−20.574
883	TYR117	CE1	−5.825	−14.316	−20.904
884	TYR117	CZ	−6.406	−15.099	−21.89
885	TYR117	OH	−5.946	−16.38	−22.095
886	TYR117	CE2	−7.544	−14.663	−22.556
887	TYR117	CD2	−8.104	−13.435	−22.227
888	TYR117	C	−6.291	−9.854	−19.99
889	TYR117	O	−5.132	−10.266	−20.121
890	PRO118	N	−6.669	−9.07	−18.995
891	PRO118	CA	−7.746	−8.09	−19.143
892	PRO118	CB	−8.071	−7.689	−17.738
893	PRO118	CG	−6.918	−8.113	−16.838
894	PRO118	CD	−5.961	−8.894	−17.727
895	PRO118	C	−7.287	−6.88	−19.956
896	PRO118	O	−6.106	−6.503	−19.927
897	ARG119	N	−8.231	−6.269	−20.648
898	ARG119	CA	−7.962	−5.027	−21.378
899	ARG119	CB	−8.86	−4.954	−22.605
900	ARG119	CG	−8.724	−6.163	−23.518
901	ARG119	CD	−9.644	−5.997	−24.72
902	ARG119	NE	−9.644	−7.191	−25.575
903	ARG119	CZ	−10.75	−7.64	−26.17
904	ARG119	NH1	−11.904	−6.989	−26.006
905	ARG119	NH2	−10.702	−8.732	−26.936
906	ARG119	C	−8.273	−3.817	−20.503
907	ARG119	O	−9.111	−2.982	−20.868
908	LEU120	N	−7.485	−3.628	−19.457
909	LEU120	CA	−7.774	−2.555	−18.499

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
910	LEU120	CB	−7.059	−2.84	−17.182
911	LEU120	CG	−7.55	−4.121	−16.52
912	LEU120	CD1	−6.671	−4.492	−15.33
913	LEU120	CD2	−9.008	−3.994	−16.094
914	LEU120	C	−7.339	−1.191	−19.028
915	LEU120	O	−8.002	−0.192	−18.722
916	LEU121	N	−6.469	−1.18	−20.027
917	LEU121	CA	−6.06	0.098	−20.609
918	LEU121	CB	−4.698	−0.08	−21.278
919	LEU121	CG	−4.032	1.24	−21.673
920	LEU121	CD1	−4.41	1.721	−23.069
921	LEU121	CD2	−4.245	2.327	−20.624
922	LEU121	C	−7.116	0.607	−21.592
923	LEU121	O	−7.3	1.827	−21.665
924	ARG122	N	−8.018	−0.27	−22.008
925	ARG122	CA	−9.115	0.114	−22.905
926	ARG122	CB	−9.6	−1.173	−23.565
927	ARG122	CG	−10.821	−0.997	−24.459
928	ARG122	CD	−11.211	−2.338	−25.071
929	ARG122	NE	−12.464	−2.253	−25.835
930	ARG122	CZ	−13.51	−3.046	−25.588
931	ARG122	NH1	−13.452	−3.935	−24.593
932	ARG122	NH2	−14.62	−2.936	−26.321
933	ARG122	C	−10.251	0.794	−22.13
934	ARG122	O	−11.01	1.595	−22.691
935	ARG123	N	−10.221	0.644	−20.814
936	ARG123	CA	−11.181	1.331	−19.95
937	ARG123	CB	−11.391	0.5	−18.692
938	ARG123	CG	−12.148	−0.78	−19.017
939	ARG123	CD	−12.443	−1.599	−17.765
940	ARG123	NE	−13.446	−2.631	−18.069
941	ARG123	CZ	−14.721	−2.531	−17.684
942	ARG123	NH1	−15.097	−1.539	−16.873
943	ARG123	NH2	−15.6	−3.47	−18.04
944	ARG123	C	−10.717	2.737	−19.571
945	ARG123	O	−11.539	3.558	−19.144
946	VAL124	N	−9.462	3.05	−19.842
947	VAL124	CA	−8.966	4.403	−19.594
948	VAL124	CB	−7.594	4.293	−18.945
949	VAL124	CG1	−7.085	5.664	−18.518
950	VAL124	CG2	−7.633	3.348	−17.754
951	VAL124	C	−8.835	5.148	−20.915
952	VAL124	O	−9.232	6.312	−21.05
953	VAL125	N	−8.348	4.426	−21.907
954	VAL125	CA	−8.16	4.979	−23.246
955	VAL125	CB	−6.725	4.689	−23.676
956	VAL125	CG1	−6.452	5.169	−25.096
957	VAL125	CG2	−5.724	5.305	−22.704
958	VAL125	C	−9.134	4.342	−24.231
959	VAL125	O	−8.955	3.193	−24.653
960	PRO126	N	−10.147	5.11	−24.596
961	PRO126	CA	−11.059	4.712	−25.67
962	PRO126	CB	−12.094	5.793	−25.722
963	PRO126	CG	−11.719	6.892	−24.74
964	PRO126	CD	−10.425	6.452	−24.085
965	PRO126	C	−10.308	4.583	−26.994
966	PRO126	O	−9.608	5.507	−27.428
967	PRO127	N	−10.471	3.432	−27.626
968	PRO127	CA	−9.571	3.011	−28.712
969	PRO127	CB	−9.745	1.524	−28.787
970	PRO127	CG	−10.915	1.103	−27.911
971	PRO127	CD	−11.371	2.36	−27.19
972	PRO127	C	−9.84	3.63	−30.093
973	PRO127	O	−8.993	3.49	−30.982
974	GLY128	N	−10.906	4.398	−30.255
975	GLY128	CA	−11.249	4.913	−31.59
976	GLY128	C	−10.819	6.364	−31.815
977	GLY128	O	−11.644	7.215	−32.165
978	GLN129	N	−9.537	6.629	−31.624
979	GLN129	CA	−8.996	7.987	−31.8
980	GLN129	CB	−8.687	8.571	−30.425
981	GLN129	CG	−9.894	8.479	−29.499
982	GLN129	CD	−9.627	9.19	−28.183
983	GLN129	OE1	−9.81	10.408	−28.089

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
984	GLN129	NE2	−9.313	8.42	−27.157
985	GLN129	C	−7.715	7.922	−32.629
986	GLN129	O	−6.719	7.343	−32.18
987	ASP130	N	−7.724	8.515	−33.81
988	ASP130	CA	−6.574	8.323	−34.699
989	ASP130	CB	−6.969	7.255	−35.719
990	ASP130	CG	−5.758	6.758	−36.499
991	ASP130	OD1	−5.358	7.46	−37.421
992	ASP130	OD2	−5.127	5.823	−36.031
993	ASP130	C	−6.137	9.62	−35.384
994	ASP130	O	−6.901	10.222	−36.145
995	PHE131	N	−4.829	9.837	−35.37
996	PHE131	CA	−4.217	11.067	−35.909
997	PHE131	CB	−2.722	11.012	−35.601
998	PHE131	CG	−2.337	11.142	−34.129
999	PHE131	CD1	−2.26	10.023	−33.309
1000	PHE131	CE1	−1.91	10.163	−31.973
1001	PHE131	CZ	−1.623	11.419	−31.458
1002	PHE131	CE2	−1.682	12.534	−32.281
1003	PHE131	CD2	−2.033	12.394	−33.617
1004	PHE131	C	−4.374	11.263	−37.423
1005	PHE131	O	−4.479	12.404	−37.884
1006	GLN132	N	−4.529	10.183	−38.17
1007	GLN132	CA	−4.725	10.299	−39.619
1008	GLN132	CB	−3.817	9.303	−40.35
1009	GLN132	CG	−2.352	9.745	−40.479
1010	GLN132	CD	−1.573	9.67	−39.164
1011	GLN132	OE1	−1.249	10.7	−38.56
1012	GLN132	NE2	−1.324	8.453	−38.711
1013	GLN132	C	−6.184	10.037	−39.995
1014	GLN132	O	−6.628	10.369	−41.102
1015	HIS133	N	−6.935	9.514	−39.041
1016	HIS133	CA	−8.335	9.126	−39.26
1017	HIS133	CB	−8.433	7.613	−39.455
1018	HIS133	CG	−7.755	7.044	−40.688
1019	HIS133	ND1	−7.661	7.619	−41.902
1020	HIS133	CE1	−6.991	6.794	−42.733
1021	HIS133	NE2	−6.657	5.688	−42.032
1022	HIS133	CD2	−7.122	5.827	−40.77
1023	HIS133	C	−9.171	9.501	−38.043
1024	HIS133	O	−9.428	8.665	−37.166
1025	GLY134	N	−9.599	10.749	−38.007
1026	GLY134	CA	−10.365	11.246	−36.862
1027	GLY134	C	−9.426	11.927	−35.876
1028	GLY134	O	−9.179	11.423	−34.769
1029	TYR135	N	−8.897	13.058	−36.315
1030	TYR135	CA	−7.926	13.823	−35.532
1031	TYR135	CB	−6.712	14.067	−36.432
1032	TYR135	CG	−5.463	14.658	−35.773
1033	TYR135	CD1	−5.203	14.431	−34.428
1034	TYR135	CE1	−4.069	14.976	−33.842
1035	TYR135	CZ	−3.196	15.744	−34.599
1036	TYR135	OH	−2.041	16.232	−34.027
1037	TYR135	CE2	−3.455	15.977	−35.942
1038	TYR135	CD2	−4.591	15.434	−36.528
1039	TYR135	C	−8.523	15.15	−35.057
1040	TYR135	O	−9.083	15.23	−33.957
1041	ALA136	N	−8.261	16.198	−35.832
1042	ALA136	CA	−8.621	17.602	−35.527
1043	ALA136	CB	−10.136	17.76	−35.624
1044	ALA136	C	−8.131	18.134	−34.17
1045	ALA136	O	−8.655	19.134	−33.669
1046	GLY137	N	−7.102	17.513	−33.613
1047	GLY137	CA	−6.583	17.89	−32.3
1048	GLY137	C	−7.457	17.447	−31.12
1049	GLY137	O	−7.201	17.896	−29.999
1050	VAL138	N	−8.449	16.596	−31.343
1051	VAL138	CA	−9.417	16.293	−30.279
1052	VAL138	CB	−10.829	16.437	−30.848
1053	VAL138	CG1	−11.856	16.524	−29.725
1054	VAL138	CG2	−10.962	17.661	−31.745
1055	VAL138	C	−9.257	14.872	−29.739
1056	VAL138	O	−9.487	13.893	−30.46
1057	PHE139	N	−8.837	14.767	−28.489

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1058	PHE139	CA	−8.785	13.461	−27.82
1059	PHE139	CB	−7.331	13.061	−27.591
1060	PHE139	CG	−6.585	12.766	−28.89
1061	PHE139	CD1	−5.674	13.679	−29.406
1062	PHE139	CE1	−5.012	13.401	−30.594
1063	PHE139	CZ	−5.258	12.211	−31.267
1064	PHE139	CE2	−6.166	11.298	−30.751
1065	PHE139	CD2	−6.828	11.576	−29.562
1066	PHE139	C	−9.567	13.485	−26.507
1067	PHE139	O	−10.021	14.546	−26.066
1068	HIS140	N	−9.881	12.307	−26.001
1069	HIS140	CA	−10.587	12.184	−24.714
1070	HIS140	CB	−12.106	12.207	−24.902
1071	HIS140	CG	−12.789	10.971	−25.476
1072	HIS140	ND1	−12.429	10.244	−26.551
1073	HIS140	CE1	−13.309	9.24	−26.734
1074	HIS140	NE2	−14.246	9.338	−25.765
1075	HIS140	CD2	−13.942	10.402	−24.988
1076	HIS140	C	−10.179	10.923	−23.959
1077	HIS140	O	−9.838	9.888	−24.551
1078	PHE141	N	−10.118	11.066	−22.646
1079	PHE141	CA	−9.733	9.944	−21.784
1080	PHE141	CB	−8.3	10.149	−21.305
1081	PHE141	CG	−7.275	10.084	−22.437
1082	PHE141	CD1	−6.749	11.251	−22.978
1083	PHE141	CE1	−5.829	11.183	−24.015
1084	PHE141	CZ	−5.431	9.949	−24.51
1085	PHE141	CE2	−5.955	8.783	−23.969
1086	PHE141	CD2	−6.878	8.85	−22.934
1087	PHE141	C	−10.678	9.752	−20.599
1088	PHE141	O	−11.061	10.696	−19.893
1089	GLN142	N	−10.962	8.486	−20.349
1090	GLN142	CA	−11.888	8.065	−19.294
1091	GLN142	CB	−12.552	6.785	−19.781
1092	GLN142	CG	−13.174	7.041	−21.148
1093	GLN142	CD	−13.645	5.75	−21.808
1094	GLN142	OE1	−14.399	5.785	−22.788
1095	GLN142	NE2	−13.197	4.627	−21.277
1096	GLN142	C	−11.137	7.832	−17.986
1097	GLN142	O	−10.384	6.864	−17.825
1098	LEU143	N	−11.373	8.731	−17.047
1099	LEU143	CA	−10.608	8.75	−15.792
1100	LEU143	CB	−9.872	10.085	−15.714
1101	LEU143	CG	−8.947	10.302	−16.911
1102	LEU143	CD1	−8.387	11.717	−16.934
1103	LEU143	CD2	−7.816	9.281	−16.943
1104	LEU143	C	−11.51	8.58	−14.569
1105	LEU143	O	−12.486	9.318	−14.383
1106	TRP144	N	−11.148	7.641	−13.712
1107	TRP144	CA	−11.946	7.364	−12.512
1108	TRP144	CB	−11.696	5.913	−12.118
1109	TRP144	CG	−12.57	5.36	−11.015
1110	TRP144	CD1	−13.834	4.843	−11.161
1111	TRP144	NE1	−14.268	4.408	−9.951
1112	TRP144	CE2	−13.341	4.626	−9.002
1113	TRP144	CZ2	−13.278	4.292	−7.654
1114	TRP144	CH2	−12.172	4.67	−6.904
1115	TRP144	CZ3	−11.121	5.358	−7.499
1116	TRP144	CE3	−11.157	5.656	−8.856
1117	TRP144	CD2	−12.257	5.281	−9.606
1118	TRP144	C	−11.57	8.304	−11.366
1119	TRP144	O	−10.388	8.505	−11.05
1120	GLN145	N	−12.591	8.903	−10.777
1121	GLN145	CA	−12.423	9.827	−9.649
1122	GLN145	CB	−12.609	11.266	−10.124
1123	GLN145	CG	−11.437	11.78	−10.955
1124	GLN145	CD	−10.195	12.021	−10.093
1125	GLN145	OE1	−10.054	13.077	−9.465
1126	GLN145	NE2	−9.303	11.046	−10.084
1127	GLN145	C	−13.421	9.539	−8.535
1128	GLN145	O	−14.566	10.008	−8.573
1129	PHE146	N	−12.923	8.839	−7.524
1130	PHE146	CA	−13.683	8.463	−6.316
1131	PHE146	CB	−13.654	9.638	−5.343

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1132	PHE146	CG	−14.266	9.342	−3.976
1133	PHE146	CD1	−13.996	8.137	−3.338
1134	PHE146	CE1	−14.555	7.87	−2.094
1135	PHE146	CZ	−15.38	8.808	−1.487
1136	PHE146	CE2	−15.646	10.014	−2.123
1137	PHE146	CD2	−15.088	10.281	−3.366
1138	PHE146	C	−15.127	8.071	−6.622
1139	PHE146	O	−16.058	8.875	−6.482
1140	GLY147	N	−15.288	6.867	−7.137
1141	GLY147	CA	−16.614	6.363	−7.506
1142	GLY147	C	−16.94	6.638	−8.972
1143	GLY147	O	−17.106	5.705	−9.767
1144	ARG148	N	−16.974	7.915	−9.314
1145	ARG148	CA	−17.38	8.371	−10.644
1146	ARG148	CB	−17.549	9.882	−10.567
1147	ARG148	CG	−18.557	10.257	−9.49
1148	ARG148	CD	−18.8	11.76	−9.438
1149	ARG148	NE	−19.893	12.071	−8.503
1150	ARG148	CZ	−20.399	13.296	−8.345
1151	ARG148	NH1	−21.421	13.492	−7.508
1152	ARG148	NH2	−19.904	14.318	−9.046
1153	ARG148	C	−16.374	8.056	−11.743
1154	ARG148	O	−15.158	8.025	−11.52
1155	TRP149	N	−16.909	7.731	−12.904
1156	TRP149	CA	−16.103	7.677	−14.127
1157	TRP149	CB	−16.488	6.467	−14.969
1158	TRP149	CG	−15.818	5.164	−14.58
1159	TRP149	CD1	−16.347	4.15	−13.813
1160	TRP149	NE1	−15.424	3.159	−13.714
1161	TRP149	CE2	−14.3	3.47	−14.39
1162	TRP149	CZ2	−13.101	2.8	−14.583
1163	TRP149	CH2	−12.099	3.385	−15.35
1164	TRP149	CZ3	−12.294	4.639	−15.922
1165	TRP149	CE3	−13.491	5.318	−15.731
1166	TRP149	CD2	−14.492	4.74	−14.965
1167	TRP149	C	−16.315	8.948	−14.942
1168	TRP149	O	−17.413	9.211	−15.452
1169	MET150	N	−15.263	9.739	−15.042
1170	MET150	CA	−15.318	10.988	−15.804
1171	MET150	CB	−14.615	12.083	−15.012
1172	MET150	CG	−15.261	12.284	−13.647
1173	MET150	SD	−14.55	13.61	−12.647
1174	MET150	CE	−15.55	13.414	−11.154
1175	MET150	C	−14.649	10.84	−17.165
1176	MET150	O	−14.038	9.811	−17.476
1177	ASP151	N	−14.879	11.824	−18.014
1178	ASP151	CA	−14.209	11.869	−19.315
1179	ASP151	CB	−15.191	11.504	−20.42
1180	ASP151	CG	−14.475	11.575	−21.765
1181	ASP151	OD1	−13.548	10.799	−21.95
1182	ASP151	OD2	−14.737	12.53	−22.485
1183	ASP151	C	−13.614	13.25	−19.575
1184	ASP151	O	−14.331	14.245	−19.751
1185	VAL152	N	−12.295	13.28	−19.638
1186	VAL152	CA	−11.568	14.53	−19.866
1187	VAL152	CB	−10.353	14.548	−18.945
1188	VAL152	CG1	−9.49	15.783	−19.176
1189	VAL152	CG2	−10.787	14.467	−17.487
1190	VAL152	C	−11.13	14.658	−21.323
1191	VAL152	O	−10.292	13.888	−21.812
1192	VAL153	N	−11.799	15.554	−22.03
1193	VAL153	CA	−11.428	15.872	−23.414
1194	VAL153	CB	−12.639	16.521	−24.08
1195	VAL153	CG1	−12.407	16.798	−25.562
1196	VAL153	CG2	−13.864	15.634	−23.903
1197	VAL153	C	−10.218	16.81	−23.411
1198	VAL153	O	−10.209	17.813	−22.693
1199	VAL154	N	−9.207	16.455	−24.185
1200	VAL154	CA	−7.929	17.177	−24.182
1201	VAL154	CB	−7.048	16.501	−23.124
1202	VAL154	CG1	−6.823	15.024	−23.436
1203	VAL154	CG2	−5.717	17.213	−22.91
1204	VAL154	C	−7.266	17.152	−25.569
1205	VAL154	O	−7.321	16.152	−26.294

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1206	ASP155	N	−6.716	18.287	−25.972
1207	ASP155	CA	−6.002	18.353	−27.255
1208	ASP155	CB	−5.987	19.775	−27.798
1209	ASP155	CG	−5.278	20.759	−26.88
1210	ASP155	OD1	−4.063	20.666	−26.773
1211	ASP155	OD2	−5.968	21.536	−26.234
1212	ASP155	C	−4.575	17.816	−27.158
1213	ASP155	O	−4.06	17.551	−26.066
1214	ASP156	N	−3.925	17.705	−28.306
1215	ASP156	CA	−2.584	17.098	−28.362
1216	ASP156	CB	−2.542	16.116	−29.527
1217	ASP156	CG	−3.042	16.749	−30.82
1218	ASP156	OD1	−2.276	17.46	−31.45
1219	ASP156	OD2	−4.16	16.442	−31.199
1220	ASP156	C	−1.4	18.068	−28.462
1221	ASP156	O	−0.502	17.842	−29.279
1222	ARG157	N	−1.36	19.101	−27.637
1223	ARG157	CA	−0.188	19.991	−27.635
1224	ARG157	CB	−0.661	21.409	−27.359
1225	ARG157	CG	−1.6	21.855	−28.472
1226	ARG157	CD	−2.264	23.188	−28.166
1227	ARG157	NE	−1.288	24.283	−28.104
1228	ARG157	CZ	−1.636	25.516	−27.732
1229	ARG157	NH1	−2.885	25.755	−27.329
1230	ARG157	NH2	−0.723	26.489	−27.695
1231	ARG157	C	0.838	19.516	−26.601
1232	ARG157	O	0.757	19.808	−25.403
1233	LEU158	N	1.846	18.838	−27.119
1234	LEU158	CA	2.816	18.097	−26.309
1235	LEU158	CB	3.428	17.041	−27.225
1236	LEU158	CG	2.367	16.055	−27.704
1237	LEU158	CD1	3.005	14.963	−28.545
1238	LEU158	CD2	1.608	15.43	−26.536
1239	LEU158	C	3.89	18.971	−25.646
1240	LEU158	O	4.196	20.082	−26.099
1241	PRO159	N	4.41	18.463	−24.535
1242	PRO159	CA	5.011	19.305	−23.487
1243	PRO159	CB	5.045	18.457	−22.257
1244	PRO159	CG	4.578	17.057	−22.594
1245	PRO159	CD	4.14	17.107	−24.04
1246	PRO159	C	6.408	19.9	−23.712
1247	PRO159	O	7.443	19.287	−23.409
1248	VAL160	N	6.361	21.125	−24.215
1249	VAL160	CA	7.337	22.214	−23.986
1250	VAL160	CB	7.392	22.44	−22.47
1251	VAL160	CG1	8.397	23.516	−22.074
1252	VAL160	CG2	6.017	22.809	−21.926
1253	VAL160	C	8.769	22.177	−24.552
1254	VAL160	O	9.103	23.114	−25.289
1255	ARG161	N	9.603	21.183	−24.29
1256	ARG161	CA	11.043	21.465	−24.453
1257	ARG161	CB	11.833	20.816	−23.327
1258	ARG161	CG	11.575	21.575	−22.03
1259	ARG161	CD	12.608	21.243	−20.966
1260	ARG161	NE	13.958	21.563	−21.457
1261	ARG161	CZ	15.025	21.578	−20.657
1262	ARG161	NH1	14.882	21.298	−19.36
1263	ARG161	NH2	16.232	21.871	−21.149
1264	ARG161	C	11.679	21.151	−25.805
1265	ARG161	O	12.483	20.219	−25.925
1266	GLU162	N	11.522	22.113	−26.707
1267	GLU162	CA	12.16	22.132	−28.039
1268	GLU162	CB	13.611	22.593	−27.882
1269	GLU162	CG	13.725	23.98	−27.258
1270	GLU162	CD	15.192	24.343	−27.015
1271	GLU162	OE1	15.888	23.532	−26.423
1272	GLU162	OE2	15.54	25.481	−27.294
1273	GLU162	C	12.17	20.762	−28.704
1274	GLU162	O	13.235	20.146	−28.838
1275	GLY163	N	10.998	20.25	−29.035
1276	GLY163	CA	10.927	18.908	−29.624
1277	GLY163	C	10.77	17.844	−28.543
1278	GLY163	O	9.77	17.119	−28.523
1279	LYS164	N	11.758	17.753	−27.666

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1280	LYS164	CA	11.727	16.797	−26.561
1281	LYS164	CB	13.034	16.901	−25.786
1282	LYS164	CG	14.223	16.549	−26.671
1283	LYS164	CD	15.528	16.602	−25.887
1284	LYS164	CE	16.707	16.177	−26.75
1285	LYS164	NZ	16.512	14.811	−27.262
1286	LYS164	C	10.55	17.061	−25.632
1287	LYS164	O	10.216	18.202	−25.281
1288	LEU165	N	9.84	15.985	−25.356
1289	LEU165	CA	8.696	16.07	−24.457
1290	LEU165	CB	7.566	15.208	−25
1291	LEU165	CG	7.278	15.532	−26.463
1292	LEU165	CD1	6.209	14.601	−27.017
1293	LEU165	CD2	6.884	16.992	−26.665
1294	LEU165	C	9.137	15.583	−23.093
1295	LEU165	O	9.521	14.42	−22.921
1296	MET166	N	9.088	16.481	−22.126
1297	MET166	CA	9.591	16.138	−20.793
1298	MET166	CB	10.078	17.404	−20.106
1299	MET166	CG	11.309	17.927	−20.834
1300	MET166	SD	12.699	16.772	−20.916
1301	MET166	CE	13.83	17.744	−21.934
1302	MET166	C	8.553	15.42	−19.94
1303	MET166	O	8.896	14.805	−18.923
1304	PHE167	N	7.313	15.433	−20.389
1305	PHE167	CA	6.29	14.63	−19.73
1306	PHE167	CB	5.056	15.483	−19.482
1307	PHE167	CG	5.279	16.567	−18.432
1308	PHE167	CD1	5.342	17.903	−18.806
1309	PHE167	CE1	5.55	18.883	−17.846
1310	PHE167	CZ	5.694	18.528	−16.512
1311	PHE167	CE2	5.628	17.193	−16.136
1312	PHE167	CD2	5.42	16.212	−17.097
1313	PHE167	C	5.982	13.395	−20.569
1314	PHE167	O	5.52	13.499	−21.714
1315	VAL168	N	6.315	12.261	−19.965
1316	VAL168	CA	6.198	10.891	−20.507
1317	VAL168	CB	5.034	10.182	−19.83
1318	VAL168	CG1	4.982	8.721	−20.263
1319	VAL168	CG2	5.179	10.26	−18.316
1320	VAL168	C	6.054	10.77	−22.018
1321	VAL168	O	4.948	10.725	−22.57
1322	ARG169	N	7.19	10.719	−22.684
1323	ARG169	CA	7.183	10.457	−24.12
1324	ARG169	CB	8.238	11.347	−24.77
1325	ARG169	CG	8.099	11.385	−26.287
1326	ARG169	CD	9.212	12.194	−26.936
1327	ARG169	NE	8.906	12.42	−28.356
1328	ARG169	CZ	9.579	13.284	−29.117
1329	ARG169	NH1	9.141	13.573	−30.344
1330	ARG169	NH2	10.61	13.956	−28.602
1331	ARG169	C	7.511	8.984	−24.356
1332	ARG169	O	8.246	8.386	−23.558
1333	SER170	N	6.875	8.379	−25.349
1334	SER170	CA	7.298	7.045	−25.789
1335	SER170	CB	6.476	6.62	−27
1336	SER170	OG	7.002	5.385	−27.473
1337	SER170	C	8.773	7.117	−26.163
1338	SER170	O	9.178	7.947	−26.981
1339	GLU171	N	9.555	6.198	−25.621
1340	GLU171	CA	11.017	6.31	−25.704
1341	GLU171	CB	11.616	5.481	−24.57
1342	GLU171	CG	12.975	6.011	−24.12
1343	GLU171	CD	12.796	7.279	−23.286
1344	GLU171	OE1	12.537	7.127	−22.096
1345	GLU171	OE2	12.869	8.363	−23.844
1346	GLU171	C	11.625	5.849	−27.038
1347	GLU171	O	12.841	5.979	−27.22
1348	GLN172	N	10.824	5.321	−27.949
1349	GLN172	CA	11.386	4.889	−29.233
1350	GLN172	CB	11.869	3.445	−29.091
1351	GLN172	CG	12.761	3.004	−30.256
1352	GLN172	CD	11.99	2.179	−31.288
1353	GLN172	OE1	11.584	1.047	−31.006

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1354	GLN172	NE2	11.73	2.778	−32.436
1355	GLN172	C	10.354	5.009	−30.347
1356	GLN172	O	10.658	5.476	−31.452
1357	ARG173	N	9.172	4.488	−30.073
1358	ARG173	CA	8.088	4.481	−31.057
1359	ARG173	CB	7.193	3.289	−30.751
1360	ARG173	CG	7.879	1.993	−31.165
1361	ARG173	CD	7.418	0.816	−30.315
1362	ARG173	NE	7.912	0.986	−28.94
1363	ARG173	CZ	8.949	0.308	−28.444
1364	ARG173	NH1	9.518	0.714	−27.309
1365	ARG173	NH2	9.549	−0.628	−29.182
1366	ARG173	C	7.28	5.771	−31.055
1367	ARG173	O	7.292	6.546	−30.093
1368	ASN174	N	6.482	5.901	−32.101
1369	ASN174	CA	5.599	7.055	−32.364
1370	ASN174	CB	5.235	7.009	−33.846
1371	ASN174	CG	4.672	5.645	−34.275
1372	ASN174	OD1	4.162	4.849	−33.475
1373	ASN174	ND2	4.642	5.464	−35.581
1374	ASN174	C	4.291	7.097	−31.563
1375	ASN174	O	3.289	7.607	−32.079
1376	GLU175	N	4.279	6.53	−30.369
1377	GLU175	CA	3.04	6.404	−29.601
1378	GLU175	CB	3.115	5.137	−28.778
1379	GLU175	CG	3.319	3.936	−29.681
1380	GLU175	CD	3.281	2.684	−28.828
1381	GLU175	OE1	4.337	2.258	−28.385
1382	GLU175	OE2	2.183	2.178	−28.629
1383	GLU175	C	2.817	7.594	−28.685
1384	GLU175	O	3.377	7.689	−27.584
1385	PHE176	N	1.848	8.394	−29.083
1386	PHE176	CA	1.554	9.64	−28.384
1387	PHE176	CB	1.279	10.715	−29.42
1388	PHE176	CG	2.513	10.961	−30.283
1389	PHE176	CD1	2.49	10.685	−31.644
1390	PHE176	CE1	3.627	10.897	−32.413
1391	PHE176	CZ	4.787	11.383	−31.823
1392	PHE176	CE2	4.81	11.657	−30.462
1393	PHE176	CD2	3.674	11.443	−29.692
1394	PHE176	C	0.428	9.509	−27.365
1395	PHE176	O	0.125	10.481	−26.66
1396	TRP177	N	−0.033	8.287	−27.139
1397	TRP177	CA	−1.02	8.054	−26.078
1398	TRP177	CB	−1.606	6.644	−26.208
1399	TRP177	CG	−0.705	5.504	−25.758
1400	TRP177	CD1	0.339	4.936	−26.455
1401	TRP177	NE1	0.884	3.952	−25.696
1402	TRP177	CE2	0.249	3.845	−24.514
1403	TRP177	CZ2	0.434	3.01	−23.421
1404	TRP177	CH2	−0.393	3.136	−22.311
1405	TRP177	CZ3	−1.403	4.092	−22.292
1406	TRP177	CE3	−1.597	4.93	−23.384
1407	TRP177	CD2	−0.773	4.807	−24.493
1408	TRP177	C	−0.39	8.223	−24.691
1409	TRP177	O	−1.078	8.692	−23.777
1410	ALA178	N	0.926	8.092	−24.6
1411	ALA178	CA	1.622	8.326	−23.33
1412	ALA178	CB	3.066	7.834	−23.424
1413	ALA178	C	1.52	9.788	−22.855
1414	ALA178	O	0.795	9.998	−21.869
1415	PRO179	N	2.011	10.789	−23.586
1416	PRO179	CA	1.86	12.158	−23.078
1417	PRO179	CB	2.745	13.004	−23.94
1418	PRO179	CG	3.271	12.171	−25.097
1419	PRO179	CD	2.749	10.766	−24.865
1420	PRO179	C	0.417	12.678	−23.111
1421	PRO179	O	0.063	13.483	−22.242
1422	LEU180	N	−0.457	12.083	−23.912
1423	LEU180	CA	−1.85	12.541	−23.941
1424	LEU180	CB	−2.487	12.1	−25.249
1425	LEU180	CG	−1.919	12.9	−26.41
1426	LEU180	CD1	−2.501	12.427	−27.737
1427	LEU180	CD2	−2.185	14.383	−26.194

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1428	LEU180	C	−2.675	12.022	−22.767
1429	LEU180	O	−3.426	12.804	−22.17
1430	LEU181	N	−2.363	10.829	−22.286
1431	LEU181	CA	−3.077	10.298	−21.121
1432	LEU181	CB	−2.863	8.789	−21.063
1433	LEU181	CG	−3.546	8.152	−19.856
1434	LEU181	CD1	−5.037	8.468	−19.828
1435	LEU181	CD2	−3.318	6.645	−19.838
1436	LEU181	C	−2.557	10.955	−19.849
1437	LEU181	O	−3.354	11.352	−18.988
1438	GLU182	N	−1.293	11.342	−19.886
1439	GLU182	CA	−0.721	12.068	−18.756
1440	GLU182	CB	0.789	12.064	−18.906
1441	GLU182	CG	1.454	12.828	−17.771
1442	GLU182	CD	2.919	13	−18.119
1443	GLU182	OE1	3.74	13.02	−17.212
1444	GLU182	OE2	3.203	12.988	−19.307
1445	GLU182	C	−1.208	13.514	−18.713
1446	GLU182	O	−1.519	14.012	−17.625
1447	LYS183	N	−1.494	14.089	−19.871
1448	LYS183	CA	−2.004	15.459	−19.916
1449	LYS183	CB	−1.811	16.003	−21.327
1450	LYS183	CG	−2.24	17.463	−21.411
1451	LYS183	CD	−2.002	18.054	−22.795
1452	LYS183	CE	−2.436	19.515	−22.826
1453	LYS183	NZ	−2.231	20.112	−24.153
1454	LYS183	C	−3.48	15.524	−19.541
1455	LYS183	O	−3.877	16.45	−18.823
1456	ALA184	N	−4.211	14.448	−19.782
1457	ALA184	CA	−5.615	14.409	−19.367
1458	ALA184	CB	−6.322	13.296	−20.125
1459	ALA184	C	−5.751	14.162	−17.868
1460	ALA184	O	−6.565	14.828	−17.216
1461	TYR185	N	−4.801	13.436	−17.298
1462	TYR185	CA	−4.82	13.182	−15.854
1463	TYR185	CB	−3.921	11.98	−15.585
1464	TYR185	CG	−4.138	11.279	−14.246
1465	TYR185	CD1	−3.046	10.913	−13.469
1466	TYR185	CE1	−3.241	10.256	−12.26
1467	TYR185	CZ	−4.53	9.966	−11.833
1468	TYR185	OH	−4.724	9.283	−10.65
1469	TYR185	CE2	−5.625	10.337	−12.603
1470	TYR185	CD2	−5.427	10.996	−13.811
1471	TYR185	C	−4.305	14.401	−15.088
1472	TYR185	O	−4.892	14.782	−14.067
1473	ALA186	N	−3.416	15.148	−15.725
1474	ALA186	CA	−2.905	16.386	−15.136
1475	ALA186	CB	−1.656	16.807	−15.897
1476	ALA186	C	−3.93	17.508	−15.214
1477	ALA186	O	−4.137	18.196	−14.211
1478	LYS187	N	−4.738	17.504	−16.262
1479	LYS187	CA	−5.812	18.493	−16.399
1480	LYS187	CB	−6.3	18.478	−17.84
1481	LYS187	CG	−7.476	19.428	−17.98
1482	LYS187	CD	−8.253	19.211	−19.266
1483	LYS187	CE	−9.5	20.078	−19.234
1484	LYS187	NZ	−10.385	19.783	−20.361
1485	LYS187	C	−7	18.174	−15.492
1486	LYS187	O	−7.617	19.088	−14.93
1487	LEU188	N	−7.127	16.905	−15.142
1488	LEU188	CA	−8.153	16.458	−14.197
1489	LEU188	CB	−8.179	14.937	−14.297
1490	LEU188	CG	−9.17	14.28	−13.349
1491	LEU188	CD1	−10.604	14.669	−13.693
1492	LEU188	CD2	−8.997	12.77	−13.4
1493	LEU188	C	−7.81	16.875	−12.765
1494	LEU188	O	−8.702	17.039	−11.923
1495	HIS189	N	−6.539	17.146	−12.524
1496	HIS189	CA	−6.108	17.656	−11.223
1497	HIS189	CB	−4.94	16.808	−10.738
1498	HIS189	CG	−5.293	15.347	−10.525
1499	HIS189	ND1	−6.46	14.853	−10.065
1500	HIS189	CE1	−6.387	13.508	−10.015
1501	HIS189	NE2	−5.158	13.151	−10.448

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1502	HIS189	CD2	−4.471	14.272	−10.764
1503	HIS189	C	−5.712	19.132	−11.306
1504	HIS189	O	−5.223	19.708	−10.327
1505	GLY190	N	−5.95	19.744	−12.453
1506	GLY190	CA	−5.548	21.135	−12.675
1507	GLY190	C	−4.382	21.222	−13.659
1508	GLY190	O	−4.567	21.486	−14.853
1509	SER191	N	−3.183	21.039	−13.133
1510	SER191	CA	−1.973	21.151	−13.957
1511	SER191	CB	−1.275	22.459	−13.61
1512	SER191	OG	−0.737	22.33	−12.301
1513	SER191	C	−1.012	19.993	−13.706
1514	SER191	O	−1.197	19.209	−12.766
1515	TYR192	N	0.102	20.004	−14.424
1516	TYR192	CA	1.115	18.94	−14.288
1517	TYR192	CB	2.175	19.114	−15.371
1518	TYR192	CG	1.799	18.695	−16.789
1519	TYR192	CD1	1.571	19.656	−17.766
1520	TYR192	CE1	1.26	19.268	−19.063
1521	TYR192	CZ	1.186	17.918	−19.38
1522	TYR192	OH	1.024	17.533	−20.693
1523	TYR192	CE2	1.409	16.955	−18.406
1524	TYR192	CD2	1.721	17.345	−17.11
1525	TYR192	C	1.836	18.964	−12.938
1526	TYR192	O	2.125	17.9	−12.376
1527	GLU193	N	1.893	20.133	−12.32
1528	GLU193	CA	2.528	20.291	−11.008
1529	GLU193	CB	2.848	21.773	−10.815
1530	GLU193	CG	3.701	22.027	−9.575
1531	GLU193	CD	3.971	23.519	−9.421
1532	GLU193	OE1	3.825	24.012	−8.315
1533	GLU193	OE2	4.212	24.168	−10.439
1534	GLU193	C	1.651	19.786	−9.857
1535	GLU193	O	2.174	19.527	−8.769
1536	VAL194	N	0.409	19.422	−10.143
1537	VAL194	CA	−0.445	18.834	−9.107
1538	VAL194	CB	−1.9	19.153	−9.438
1539	VAL194	CG1	−2.837	18.683	−8.331
1540	VAL194	CG2	−2.069	20.653	−9.648
1541	VAL194	C	−0.204	17.32	−9.012
1542	VAL194	O	−0.588	16.675	−8.03
1543	MET195	N	0.578	16.797	−9.947
1544	MET195	CA	1.024	15.405	−9.875
1545	MET195	CB	1.244	14.882	−11.289
1546	MET195	CG	−0.034	14.945	−12.119
1547	MET195	SD	0.114	14.31	−13.804
1548	MET195	CE	0.58	12.604	−13.431
1549	MET195	C	2.322	15.269	−9.072
1550	MET195	O	2.783	14.142	−8.841
1551	ARG196	N	2.867	16.384	−8.602
1552	ARG196	CA	4.085	16.369	−7.777
1553	ARG196	CB	4.684	17.771	−7.767
1554	ARG196	CG	5.258	18.147	−9.129
1555	ARG196	CD	6.477	17.294	−9.466
1556	ARG196	NE	7.549	17.508	−8.477
1557	ARG196	CZ	8.084	16.529	−7.742
1558	ARG196	NH1	8.993	16.821	−6.809
1559	ARG196	NH2	7.658	15.272	−7.891
1560	ARG196	C	3.79	15.934	−6.344
1561	ARG196	O	3.565	16.758	−5.451
1562	GLY197	N	3.851	14.63	−6.136
1563	GLY197	CA	3.542	14.04	−4.836
1564	GLY197	C	2.377	13.064	−4.977
1565	GLY197	O	1.804	12.611	−3.978
1566	GLY198	N	2.021	12.771	−6.218
1567	GLY198	CA	0.931	11.831	−6.491
1568	GLY198	C	1.414	10.394	−6.333
1569	GLY198	O	2.366	9.962	−6.994
1570	HIS199	N	0.786	9.683	−5.413
1571	HIS199	CA	1.164	8.293	−5.151
1572	HIS199	CB	0.494	7.825	−3.865
1573	HIS199	CG	0.906	8.627	−2.646
1574	HIS199	ND1	0.087	9.213	−1.752
1575	HIS199	CE1	0.826	9.838	−0.813

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1576	HIS199	NE2	2.129	9.644	−1.121
1577	HIS199	CD2	2.194	8.899	−2.248
1578	HIS199	C	0.761	7.387	−6.305
1579	HIS199	O	−0.35	7.488	−6.843
1580	MET200	N	1.579	6.368	−6.519
1581	MET200	CA	1.362	5.426	−7.626
1582	MET200	CB	2.599	4.548	−7.748
1583	MET200	CG	3.845	5.376	−8.037
1584	MET200	SD	5.389	4.439	−8.097
1585	MET200	CE	4.919	3.212	−9.338
1586	MET200	C	0.144	4.533	−7.411
1587	MET200	O	−0.577	4.258	−8.376
1588	ASN201	N	−0.251	4.36	−6.158
1589	ASN201	CA	−1.457	3.588	−5.862
1590	ASN201	CB	−1.439	3.201	−4.389
1591	ASN201	CG	−2.614	2.276	−4.097
1592	ASN201	OD1	−2.996	1.459	−4.942
1593	ASN201	ND2	−3.18	2.42	−2.912
1594	ASN201	C	−2.725	4.389	−6.152
1595	ASN201	O	−3.684	3.814	−6.673
1596	GLU202	N	−2.611	5.708	−6.157
1597	GLU202	CA	−3.776	6.547	−6.441
1598	GLU202	CB	−3.538	7.919	−5.828
1599	GLU202	CG	−3.37	7.801	−4.319
1600	GLU202	CD	−2.893	9.124	−3.734
1601	GLU202	OE1	−2.172	9.81	−4.446
1602	GLU202	OE2	−2.946	9.236	−2.517
1603	GLU202	C	−3.968	6.67	−7.947
1604	GLU202	O	−5.109	6.64	−8.428
1605	ALA203	N	−2.873	6.52	−8.674
1606	ALA203	CA	−2.955	6.474	−10.131
1607	ALA203	CB	−1.558	6.652	−10.711
1608	ALA203	C	−3.536	5.134	−10.565
1609	ALA203	O	−4.572	5.127	−11.241
1610	PHE204	N	−3.103	4.075	−9.896
1611	PHE204	CA	−3.589	2.716	−10.183
1612	PHE204	CB	−2.844	1.726	−9.289
1613	PHE204	CG	−1.45	1.267	−9.725
1614	PHE204	CD1	−0.885	0.166	−9.093
1615	PHE204	CE1	0.377	−0.28	−9.463
1616	PHE204	CZ	1.077	0.374	−10.467
1617	PHE204	CE2	0.513	1.472	−11.102
1618	PHE204	CD2	−0.75	1.917	−10.733
1619	PHE204	C	−5.085	2.57	−9.905
1620	PHE204	O	−5.847	2.193	−10.808
1621	VAL205	N	−5.526	3.094	−8.772
1622	VAL205	CA	−6.943	3.009	−8.414
1623	VAL205	CB	−7.096	3.394	−6.946
1624	VAL205	CG1	−8.556	3.412	−6.522
1625	VAL205	CG2	−6.318	2.442	−6.049
1626	VAL205	C	−7.82	3.904	−9.291
1627	VAL205	O	−8.837	3.403	−9.786
1628	ASP206	N	−7.275	5.014	−9.767
1629	ASP206	CA	−8.03	5.919	−10.648
1630	ASP206	CB	−7.431	7.319	−10.566
1631	ASP206	CG	−7.656	7.959	−9.196
1632	ASP206	OD1	−8.513	7.478	−8.464
1633	ASP206	OD2	−7.097	9.028	−8.984
1634	ASP206	C	−8.04	5.479	−12.115
1635	ASP206	O	−8.715	6.101	−12.946
1636	PHE207	N	−7.301	4.436	−12.447
1637	PHE207	CA	−7.366	3.901	−13.805
1638	PHE207	CB	−5.953	3.792	−14.366
1639	PHE207	CG	−5.264	5.142	−14.529
1640	PHE207	CD1	−5.955	6.212	−15.082
1641	PHE207	CE1	−5.332	7.444	−15.22
1642	PHE207	CZ	−4.018	7.605	−14.806
1643	PHE207	CE2	−3.325	6.533	−14.263
1644	PHE207	CD2	−3.944	5.299	−14.133
1645	PHE207	C	−8.046	2.538	−13.852
1646	PHE207	O	−8.265	1.997	−14.941
1647	THR208	N	−8.347	1.965	−12.7
1648	THR208	CA	−8.946	0.626	−12.71
1649	THR208	CB	−7.962	−0.313	−12.024

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1650	THR208	OG1	−6.678	−0.099	−12.595
1651	THR208	CG2	−8.336	−1.781	−12.2
1652	THR208	C	−10.295	0.575	−11.998
1653	THR208	O	−11.123	−0.302	−12.28
1654	GLY209	N	−10.507	1.512	−11.089
1655	GLY209	CA	−11.723	1.535	−10.275
1656	GLY209	C	−11.728	0.36	−9.303
1657	GLY209	O	−12.757	−0.299	−9.118
1658	GLY210	N	−10.587	0.108	−8.684
1659	GLY210	CA	−10.466	−1.096	−7.861
1660	GLY210	C	−9.703	−0.903	−6.556
1661	GLY210	O	−9.308	0.206	−6.174
1662	VAL211	N	−9.441	−2.033	−5.922
1663	VAL211	CA	−8.828	−2.06	−4.588
1664	VAL211	CB	−9.303	−3.323	−3.874
1665	VAL211	CG1	−8.737	−3.407	−2.459
1666	VAL211	CG2	−10.826	−3.387	−3.842
1667	VAL211	C	−7.305	−2.056	−4.664
1668	VAL211	O	−6.671	−3.101	−4.86
1669	GLY212	N	−6.73	−0.879	−4.493
1670	GLY212	CA	−5.269	−0.739	−4.494
1671	GLY212	C	−4.677	−0.95	−3.105
1672	GLY212	O	−4.645	−0.039	−2.266
1673	GLU213	N	−4.213	−2.164	−2.87
1674	GLU213	CA	−3.64	−2.506	−1.564
1675	GLU213	CB	−4.159	−3.881	−1.152
1676	GLU213	CG	−4.026	−4.88	−2.293
1677	GLU213	CD	−4.544	−6.252	−1.895
1678	GLU213	OE1	−5.454	−6.31	−1.081
1679	GLU213	OE2	−3.976	−7.224	−2.378
1680	GLU213	C	−2.112	−2.469	−1.581
1681	GLU213	O	−1.446	−3.255	−2.266
1682	VAL214	N	−1.567	−1.542	−0.813
1683	VAL214	CA	−0.11	−1.408	−0.708
1684	VAL214	CB	0.24	0.055	−0.441
1685	VAL214	CG1	0.176	0.878	−1.721
1686	VAL214	CG2	−0.645	0.67	0.64
1687	VAL214	C	0.472	−2.315	0.377
1688	VAL214	O	0.378	−2.047	1.58
1689	LEU215	N	1.075	−3.399	−0.076
1690	LEU215	CA	1.743	−4.337	0.826
1691	LEU215	CB	1.727	−5.73	0.207
1692	LEU215	CG	0.309	−6.255	0.013
1693	LEU215	CD1	0.332	−7.592	−0.716
1694	LEU215	CD2	−0.42	−6.387	1.346
1695	LEU215	C	3.185	−3.918	1.067
1696	LEU215	O	3.715	−3.013	0.407
1697	TYR216	N	3.772	−4.508	2.09
1698	TYR216	CA	5.19	−4.291	2.376
1699	TYR216	CB	5.329	−3.543	3.7
1700	TYR216	CG	4.651	−2.175	3.726
1701	TYR216	CD1	5.21	−1.116	3.022
1702	TYR216	CE1	4.591	0.128	3.033
1703	TYR216	CZ	3.416	0.308	3.749
1704	TYR216	OH	2.773	1.526	3.712
1705	TYR216	CE2	2.86	−0.746	4.462
1706	TYR216	CD2	3.48	−1.988	4.451
1707	TYR216	C	5.907	−5.633	2.455
1708	TYR216	O	5.672	−6.393	3.401
1709	LEU217	N	6.936	−5.794	1.637
1710	LEU217	CA	7.622	−7.095	1.52
1711	LEU217	CB	8.355	−7.205	0.193
1712	LEU217	CG	7.409	−7.232	−0.997
1713	LEU217	CD1	8.194	−7.5	−2.276
1714	LEU217	CD2	6.323	−8.288	−0.818
1715	LEU217	C	8.622	−7.373	2.636
1716	LEU217	O	9.222	−8.453	2.688
1717	ARG218	N	8.814	−6.413	3.521
1718	ARG218	CA	9.625	−6.664	4.708
1719	ARG218	CB	10.662	−5.552	4.798
1720	ARG218	CG	11.472	−5.563	3.503
1721	ARG218	CD	12.542	−4.482	3.425
1722	ARG218	NE	13.214	−4.554	2.115
1723	ARG218	CZ	13.955	−3.568	1.605

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1724	ARG218	NH1	14.153	−2.45	2.306
1725	ARG218	NH2	14.512	−3.707	0.4
1726	ARG218	C	8.739	−6.771	5.952
1727	ARG218	O	9.17	−7.296	6.984
1728	GLN219	N	7.457	−6.487	5.766
1729	GLN219	CA	6.483	−6.514	6.866
1730	GLN219	CB	5.876	−5.123	7.051
1731	GLN219	CG	6.907	−3.999	7.141
1732	GLN219	CD	7.851	−4.167	8.33
1733	GLN219	OE1	9.072	−4.055	8.17
1734	GLN219	NE2	7.282	−4.356	9.508
1735	GLN219	C	5.345	−7.484	6.549
1736	GLN219	O	4.196	−7.242	6.94
1737	ASN220	N	5.683	−8.562	5.861
1738	ASN220	CA	4.704	−9.52	5.324
1739	ASN220	CB	5.468	−10.684	4.708
1740	ASN220	CG	6.469	−10.173	3.688
1741	ASN220	OD1	6.106	−9.487	2.725
1742	ASN220	ND2	7.721	−10.534	3.905
1743	ASN220	C	3.757	−10.12	6.352
1744	ASN220	O	4.107	−10.363	7.513
1745	SER221	N	2.542	−10.352	5.887
1746	SER221	CA	1.57	−11.131	6.655
1747	SER221	CB	0.175	−10.937	6.072
1748	SER221	OG	0.158	−11.536	4.783
1749	SER221	C	1.951	−12.602	6.561
1750	SER221	O	2.72	−12.996	5.674
1751	MET222	N	1.347	−13.421	7.405
1752	MET222	CA	1.675	−14.853	7.407
1753	MET222	CB	1.078	−15.48	8.662
1754	MET222	CG	1.431	−16.96	8.772
1755	MET222	SD	0.795	−17.802	10.238
1756	MET222	CE	−0.974	−17.537	9.972
1757	MET222	C	1.135	−15.557	6.16
1758	MET222	O	1.84	−16.373	5.556
1759	GLY223	N	0.025	−15.058	5.64
1760	GLY223	CA	−0.504	−15.579	4.378
1761	GLY223	C	−0.231	−14.606	3.234
1762	GLY223	O	−1.094	−14.399	2.371
1763	LEU224	N	1.019	−14.182	3.115
1764	LEU224	CA	1.372	−13.17	2.114
1765	LEU224	CB	2.752	−12.615	2.47
1766	LEU224	CG	3.058	−11.259	1.83
1767	LEU224	CD1	3.56	−11.356	0.392
1768	LEU224	CD2	1.883	−10.295	1.957
1769	LEU224	C	1.366	−13.776	0.713
1770	LEU224	O	0.729	−13.204	−0.178
1771	PHE225	N	1.733	−15.045	0.618
1772	PHE225	CA	1.721	−15.722	−0.683
1773	PHE225	CB	2.683	−16.901	−0.624
1774	PHE225	CG	3.012	−17.54	−1.971
1775	PHE225	CD1	2.309	−18.653	−2.413
1776	PHE225	CE1	2.624	−19.236	−3.633
1777	PHE225	CZ	3.645	−18.707	−4.411
1778	PHE225	CE2	4.349	−17.595	−3.97
1779	PHE225	CD2	4.034	−17.013	−2.75
1780	PHE225	C	0.316	−16.2	−1.052
1781	PHE225	O	−0.014	−16.261	−2.243
1782	SER226	N	−0.568	−16.216	−0.066
1783	SER226	CA	−1.958	−16.584	−0.31
1784	SER226	CB	−2.621	−16.933	1.017
1785	SER226	OG	−1.811	−17.897	1.674
1786	SER226	C	−2.668	−15.388	−0.921
1787	SER226	O	−3.298	−15.53	−1.974
1788	ALA227	N	−2.272	−14.207	−0.472
1789	ALA227	CA	−2.811	−12.962	−1.025
1790	ALA227	CB	−2.56	−11.838	−0.027
1791	ALA227	C	−2.178	−12.606	−2.37
1792	ALA227	O	−2.813	−11.909	−3.173
1793	LEU228	N	−1.041	−13.208	−2.682
1794	LEU228	CA	−0.464	−13.054	−4.019
1795	LEU228	CB	0.987	−13.529	−4.024
1796	LEU228	CG	1.874	−12.738	−3.069
1797	LEU228	CD1	3.293	−13.292	−3.068

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1798	LEU228	CD2	1.885	−11.254	−3.409
1799	LEU228	C	−1.253	−13.894	−5.014
1800	LEU228	O	−1.722	−13.359	−6.027
1801	ARG229	N	−1.635	−15.089	−4.59
1802	ARG229	CA	−2.425	−15.971	−5.455
1803	ARG229	CB	−2.46	−17.365	−4.836
1804	ARG229	CG	−1.073	−17.992	−4.815
1805	ARG229	CD	−1.092	−19.408	−4.253
1806	ARG229	NE	−1.491	−19.417	−2.837
1807	ARG229	CZ	−1.033	−20.317	−1.963
1808	ARG229	NH1	−0.177	−21.261	−2.363
1809	ARG229	NH2	−1.432	−20.277	−0.691
1810	ARG229	C	−3.851	−15.464	−5.635
1811	ARG229	O	−4.306	−15.338	−6.782
1812	HIS230	N	−4.418	−14.914	−4.574
1813	HIS230	CA	−5.797	−14.421	−4.641
1814	HIS230	CB	−6.316	−14.129	−3.236
1815	HIS230	CG	−6.362	−15.313	−2.287
1816	HIS230	ND1	−6.351	−15.245	−0.944
1817	HIS230	CE1	−6.395	−16.491	−0.43
1818	HIS230	NE2	−6.449	−17.357	−1.468
1819	HIS230	CD2	−6.441	−16.647	−2.618
1820	HIS230	C	−5.902	−13.145	−5.468
1821	HIS230	O	−6.79	−13.057	−6.323
1822	ALA231	N	−4.89	−12.293	−5.404
1823	ALA231	CA	−4.94	−11.057	−6.185
1824	ALA231	CB	−4.032	−10.02	−5.546
1825	ALA231	C	−4.533	−11.271	−7.637
1826	ALA231	O	−5.138	−10.657	−8.525
1827	LEU232	N	−3.754	−12.304	−7.909
1828	LEU232	CA	−3.418	−12.579	−9.305
1829	LEU232	CB	−2.226	−13.525	−9.369
1830	LEU232	CG	−1.662	−13.577	−10.784
1831	LEU232	CD1	−1.186	−12.194	−11.221
1832	LEU232	CD2	−0.529	−14.589	−10.889
1833	LEU232	C	−4.619	−13.194	−10.021
1834	LEU232	O	−4.971	−12.719	−11.108
1835	ALA233	N	−5.417	−13.941	−9.271
1836	ALA233	CA	−6.662	−14.513	−9.801
1837	ALA233	CB	−7.008	−15.752	−8.981
1838	ALA233	C	−7.842	−13.535	−9.766
1839	ALA233	O	−8.922	−13.857	−10.275
1840	LYS234	N	−7.627	−12.364	−9.183
1841	LYS234	CA	−8.608	−11.272	−9.15
1842	LYS234	CB	−8.573	−10.683	−7.74
1843	LYS234	CG	−9.715	−9.713	−7.466
1844	LYS234	CD	−11.065	−10.412	−7.551
1845	LYS234	CE	−11.161	−11.548	−6.539
1846	LYS234	NZ	−12.477	−12.201	−6.605
1847	LYS234	C	−8.243	−10.2	−10.187
1848	LYS234	O	−8.839	−9.113	−10.239
1849	GLU235	N	−7.218	−10.511	−10.969
1850	GLU235	CA	−6.699	−9.626	−12.021
1851	GLU235	CB	−7.782	−9.293	−13.051
1852	GLU235	CG	−8.522	−10.52	−13.591
1853	GLU235	CD	−7.567	−11.545	−14.188
1854	GLU235	OE1	−6.837	−11.171	−15.1
1855	GLU235	OE2	−7.372	−12.563	−13.53
1856	GLU235	C	−6.167	−8.352	−11.376
1857	GLU235	O	−6.882	−7.348	−11.264
1858	SER236	N	−4.97	−8.451	−10.828
1859	SER236	CA	−4.383	−7.299	−10.145
1860	SER236	CB	−4.214	−7.632	−8.672
1861	SER236	OG	−5.506	−7.921	−8.169
1862	SER236	C	−3.039	−6.877	−10.716
1863	SER236	O	−2.139	−7.697	−10.936
1864	LEU237	N	−2.892	−5.572	−10.867
1865	LEU237	CA	−1.61	−5	−11.292
1866	LEU237	CB	−1.842	−3.667	−11.992
1867	LEU237	CG	−2.628	−3.826	−13.289
1868	LEU237	CD1	−2.99	−2.462	−13.866
1869	LEU237	CD2	−1.85	−4.655	−14.308
1870	LEU237	C	−0.716	−4.803	−10.071
1871	LEU237	O	−0.986	−3.952	−9.214

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1872	VAL238	N	0.286	−5.657	−9.968
1873	VAL238	CA	1.185	−5.649	−8.811
1874	VAL238	CB	1.358	−7.092	−8.364
1875	VAL238	CG1	2.183	−7.167	−7.09
1876	VAL238	CG2	0.007	−7.77	−8.162
1877	VAL238	C	2.554	−5.054	−9.138
1878	VAL238	O	3.279	−5.581	−9.989
1879	GLY239	N	2.922	−4.012	−8.411
1880	GLY239	CA	4.238	−3.375	−8.588
1881	GLY239	C	5.053	−3.312	−7.291
1882	GLY239	O	4.72	−2.566	−6.357
1883	ALA240	N	6.114	−4.105	−7.254
1884	ALA240	CA	7.057	−4.105	−6.123
1885	ALA240	CB	7.61	−5.513	−5.941
1886	ALA240	C	8.219	−3.129	−6.343
1887	ALA240	O	9.065	−3.327	−7.22
1888	THR241	N	8.31	−2.151	−5.459
1889	THR241	CA	9.303	−1.071	−5.562
1890	THR241	CB	8.58	0.268	−5.498
1891	THR241	OG1	7.923	0.342	−4.245
1892	THR241	CG2	7.528	0.42	−6.586
1893	THR241	C	10.32	−1.121	−4.42
1894	THR241	O	9.964	−1.319	−3.252
1895	ALA242	N	11.587	−0.945	−4.764
1896	ALA242	CA	12.661	−0.865	−3.75
1897	ALA242	CB	13.838	−1.718	−4.203
1898	ALA242	C	13.088	0.585	−3.485
1899	ALA242	O	12.372	1.512	−3.885
1900	LEU243	N	14.223	0.795	−2.829
1901	LEU243	CA	14.553	2.163	−2.389
1902	LEU243	CB	14.154	2.319	−0.924
1903	LEU243	CG	14.163	3.786	−0.494
1904	LEU243	CD1	13.222	4.616	−1.357
1905	LEU243	CD2	13.825	3.941	0.984
1906	LEU243	C	16.03	2.558	−2.551
1907	LEU243	O	16.751	2.748	−1.56
1908	SER244	N	16.485	2.51	−3.794
1909	SER244	CA	17.744	3.137	−4.253
1910	SER244	CB	17.682	4.619	−3.869
1911	SER244	OG	18.741	5.318	−4.514
1912	SER244	C	19.038	2.502	−3.714
1913	SER244	O	20.129	3.031	−3.95
1914	ASP245	N	18.943	1.318	−3.139
1915	ASP245	CA	20.132	0.66	−2.598
1916	ASP245	CB	19.678	−0.213	−1.425
1917	ASP245	CG	20.865	−0.835	−0.703
1918	ASP245	OD1	21.717	−0.065	−0.285
1919	ASP245	OD2	21.075	−2.021	−0.923
1920	ASP245	C	20.8	−0.143	−3.714
1921	ASP245	O	20.14	−0.501	−4.694
1922	ARG246	N	22.116	−0.262	−3.672
1923	ARG246	CA	22.817	−1.05	−4.697
1924	ARG246	CB	23.693	−0.122	−5.53
1925	ARG246	CG	22.803	0.795	−6.358
1926	ARG246	CD	23.574	1.845	−7.141
1927	ARG246	NE	22.635	2.697	−7.893
1928	ARG246	CZ	22.184	3.879	−7.464
1929	ARG246	NH1	21.259	4.532	−8.171
1930	ARG246	NH2	22.605	4.379	−6.298
1931	ARG246	C	23.619	−2.218	−4.123
1932	ARG246	O	24.597	−2.673	−4.728
1933	GLY247	N	23.197	−2.707	−2.969
1934	GLY247	CA	23.848	−3.872	−2.367
1935	GLY247	C	23.388	−5.139	−3.079
1936	GLY247	O	22.3	−5.178	−3.667
1937	GLU248	N	24.207	−6.177	−3.013
1938	GLU248	CA	23.847	−7.419	−3.706
1939	GLU248	CB	25.105	−8.209	−4.082
1940	GLU248	CG	25.841	−8.775	−2.871
1941	GLU248	CD	27.009	−9.663	−3.3
1942	GLU248	OE1	27.502	−10.358	−2.42
1943	GLU248	OE2	27.588	−9.344	−4.328
1944	GLU248	C	22.909	−8.304	−2.891
1945	GLU248	O	22.111	−9.013	−3.508

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
1946	TYR249	N	22.795	−8.042	−1.594
1947	TYR249	CA	22.057	−8.924	−0.673
1948	TYR249	CB	20.573	−8.569	−0.686
1949	TYR249	CG	20.278	−7.183	−0.116
1950	TYR249	CD1	19.347	−6.361	−0.737
1951	TYR249	CE1	19.086	−5.097	−0.223
1952	TYR249	CZ	19.758	−4.658	0.908
1953	TYR249	OH	19.579	−3.365	1.348
1954	TYR249	CE2	20.681	−5.481	1.538
1955	TYR249	CD2	20.94	−6.746	1.026
1956	TYR249	C	22.278	−10.381	−1.058
1957	TYR249	O	21.46	−10.984	−1.767
1958	ARG250	N	23.335	−10.943	−0.489
1959	ARG250	CA	23.879	−12.249	−0.891
1960	ARG250	CB	23.085	−13.354	−0.209
1961	ARG250	CG	23.274	−13.286	1.3
1962	ARG250	CD	22.586	−14.451	2.001
1963	ARG250	NE	22.849	−14.412	3.449
1964	ARG250	CZ	23.721	−15.216	4.063
1965	ARG250	NH1	24.393	−16.136	3.366
1966	ARG250	NH2	23.907	−15.113	5.382
1967	ARG250	C	23.902	−12.471	−2.403
1968	ARG250	O	24.686	−11.833	−3.114
1969	THR251	N	22.98	−13.293	−2.881
1970	THR251	CA	23.01	−13.788	−4.265
1971	THR251	CB	22.222	−15.092	−4.314
1972	THR251	OG1	20.839	−14.779	−4.185
1973	THR251	CG2	22.625	−16.046	−3.195
1974	THR251	C	22.407	−12.858	−5.322
1975	THR251	O	22.397	−13.239	−6.497
1976	GLU252	N	21.887	−11.699	−4.956
1977	GLU252	CA	21.278	−10.853	−5.991
1978	GLU252	CB	20.148	−10.011	−5.409
1979	GLU252	CG	19.129	−10.898	−4.701
1980	GLU252	CD	18.584	−11.953	−5.663
1981	GLU252	OE1	18.219	−11.585	−6.774
1982	GLU252	OE2	18.512	−13.104	−5.253
1983	GLU252	C	22.33	−9.997	−6.697
1984	GLU252	O	23.49	−9.915	−6.281
1985	GLU253	N	21.893	−9.346	−7.762
1986	GLU253	CA	22.81	−8.686	−8.706
1987	GLU253	CB	22.092	−8.697	−10.054
1988	GLU253	CG	23.011	−8.47	−11.25
1989	GLU253	CD	22.306	−9.023	−12.481
1990	GLU253	OE1	22.805	−8.808	−13.577
1991	GLU253	OE2	21.463	−9.884	−12.259
1992	GLU253	C	23.253	−7.266	−8.313
1993	GLU253	O	24.013	−6.635	−9.058
1994	GLY254	N	22.83	−6.78	−7.157
1995	GLY254	CA	23.269	−5.454	−6.702
1996	GLY254	C	22.297	−4.366	−7.133
1997	GLY254	O	22.689	−3.229	−7.419
1998	LEU255	N	21.026	−4.723	−7.161
1999	LEU255	CA	20.006	−3.801	−7.667
2000	LEU255	CB	19.59	−4.219	−9.074
2001	LEU255	CG	20.652	−3.871	−10.113
2002	LEU255	CD1	20.291	−4.439	−11.48
2003	LEU255	CD2	20.864	−2.361	−10.193
2004	LEU255	C	18.785	−3.756	−6.763
2005	LEU255	O	17.86	−4.567	−6.882
2006	VAL256	N	18.808	−2.787	−5.865
2007	VAL256	CA	17.69	−2.525	−4.956
2008	VAL256	CB	18.196	−2.622	−3.519
2009	VAL256	CG1	17.062	−2.907	−2.538
2010	VAL256	CG2	19.271	−3.695	−3.395
2011	VAL256	C	17.14	−1.123	−5.251
2012	VAL256	O	16.569	−0.458	−4.373
2013	LYS257	N	17.448	−0.639	−6.444
2014	LYS257	CA	16.949	0.653	−6.947
2015	LYS257	CB	17.503	0.928	−8.354
2016	LYS257	CG	16.766	0.235	−9.511
2017	LYS257	CD	17.046	−1.259	−9.655
2018	LYS257	CE	16.253	−1.859	−10.811
2019	LYS257	NZ	16.474	−3.309	−10.912

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2020	LYS257	C	15.419	0.71	−6.963
2021	LYS257	O	14.741	−0.315	−6.824
2022	GLY258	N	14.884	1.918	−7.057
2023	GLY258	CA	13.427	2.125	−6.993
2024	GLY258	C	12.688	1.852	−8.308
2025	GLY258	O	12.244	2.777	−8.998
2026	HIS259	N	12.6	0.581	−8.66
2027	HIS259	CA	11.832	0.157	−9.832
2028	HIS259	CB	12.776	−0.529	−10.818
2029	HIS259	CG	12.14	−0.9	−12.145
2030	HIS259	ND1	11.851	−0.063	−13.159
2031	HIS259	CE1	11.284	−0.758	−14.166
2032	HIS259	NE2	11.213	−2.052	−13.78
2033	HIS259	CD2	11.735	−2.154	−12.537
2034	HIS259	C	10.718	−0.787	−9.387
2035	HIS259	O	10.872	−1.503	−8.39
2036	ALA260	N	9.576	−0.675	−10.047
2037	ALA260	CA	8.426	−1.534	−9.752
2038	ALA260	CB	7.143	−0.75	−9.994
2039	ALA260	C	8.41	−2.792	−10.612
2040	ALA260	O	8.03	−2.764	−11.788
2041	TYR261	N	8.795	−3.892	−9.995
2042	TYR261	CA	8.73	−5.201	−10.645
2043	TYR261	CB	9.817	−6.093	−10.069
2044	TYR261	CG	11.082	−6.017	−10.909
2045	TYR261	CD1	12.063	−5.063	−10.662
2046	TYR261	CE1	13.189	−5.001	−11.474
2047	TYR261	CZ	13.327	−5.894	−12.529
2048	TYR261	OH	14.303	−5.696	−13.484
2049	TYR261	CE2	12.361	−6.86	−12.754
2050	TYR261	CD2	11.238	−6.923	−11.946
2051	TYR261	C	7.352	−5.831	−10.489
2052	TYR261	O	6.792	−5.893	−9.387
2053	SER262	N	6.815	−6.29	−11.604
2054	SER262	CA	5.452	−6.822	−11.607
2055	SER262	CB	4.854	−6.636	−12.994
2056	SER262	OG	3.521	−7.127	−12.949
2057	SER262	C	5.389	−8.295	−11.226
2058	SER262	O	6.173	−9.111	−11.718
2059	ILE263	N	4.484	−8.625	−10.323
2060	ILE263	CA	4.24	−10.042	−10.02
2061	ILE263	CB	3.669	−10.197	−8.613
2062	ILE263	CG2	3.137	−11.608	−8.382
2063	ILE263	CG1	4.715	−9.84	−7.563
2064	ILE263	CD1	4.176	−10.017	−6.148
2065	ILE263	C	3.28	−10.612	−11.057
2066	ILE263	O	2.125	−10.18	−11.161
2067	THR264	N	3.798	−11.509	−11.879
2068	THR264	CA	2.985	−12.101	−12.945
2069	THR264	CB	3.69	−11.906	−14.284
2070	THR264	OG1	5.038	−12.343	−14.154
2071	THR264	CG2	3.713	−10.436	−14.682
2072	THR264	C	2.7	−13.58	−12.706
2073	THR264	O	1.791	−14.147	−13.323
2074	GLY265	N	3.446	−14.189	−11.801
2075	GLY265	CA	3.194	−15.599	−11.477
2076	GLY265	C	3.424	−15.914	−10.003
2077	GLY265	O	4.371	−15.415	−9.386
2078	THR266	N	2.576	−16.771	−9.464
2079	THR266	CA	2.704	−17.196	−8.064
2080	THR266	CB	1.8	−16.307	−7.211
2081	THR266	OG1	1.703	−16.874	−5.914
2082	THR266	CG2	0.394	−16.189	−7.785
2083	THR266	C	2.382	−18.689	−7.923
2084	THR266	O	1.233	−19.101	−7.719
2085	HIS267	N	3.43	−19.493	−8.001
2086	HIS267	CA	3.254	−20.953	−8.073
2087	HIS267	CB	3.584	−21.407	−9.492
2088	HIS267	CG	2.66	−20.847	−10.556
2089	HIS267	ND1	1.39	−21.226	−10.791
2090	HIS267	CE1	0.888	−20.502	−11.811
2091	HIS267	NE2	1.855	−19.651	−12.224
2092	HIS267	CD2	2.952	−19.851	−11.46
2093	HIS267	C	4.133	−21.715	−7.084

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2094	HIS267	O	5.365	−21.646	−7.136
2095	LYS268	N	3.494	−22.487	−6.223
2096	LYS268	CA	4.227	−23.288	−5.232
2097	LYS268	CB	3.341	−23.405	−3.997
2098	LYS268	CG	4.077	−24.019	−2.814
2099	LYS268	CD	3.193	−24.016	−1.575
2100	LYS268	CE	3.981	−24.416	−0.335
2101	LYS268	NZ	3.112	−24.45	0.851
2102	LYS268	C	4.558	−24.675	−5.788
2103	LYS268	O	3.658	−25.489	−6.02
2104	VAL269	N	5.84	−24.937	−5.995
2105	VAL269	CA	6.251	−26.208	−6.608
2106	VAL269	CB	7.261	−25.928	−7.719
2107	VAL269	CG1	6.705	−24.916	−8.712
2108	VAL269	CG2	8.597	−25.442	−7.166
2109	VAL269	C	6.865	−27.176	−5.599
2110	VAL269	O	7.439	−26.778	−4.578
2111	PHE270	N	6.741	−28.459	−5.894
2112	PHE270	CA	7.389	−29.484	−5.064
2113	PHE270	CB	6.663	−30.816	−5.215
2114	PHE270	CG	5.355	−30.906	−4.438
2115	PHE270	CD1	4.15	−30.567	−5.04
2116	PHE270	CE1	2.966	−30.652	−4.319
2117	PHE270	CZ	2.987	−31.082	−2.999
2118	PHE270	CE2	4.191	−31.427	−2.399
2119	PHE270	CD2	5.375	−31.339	−3.119
2120	PHE270	C	8.853	−29.649	−5.45
2121	PHE270	O	9.191	−30.347	−6.409
2122	LEU271	N	9.704	−28.961	−4.711
2123	LEU271	CA	11.152	−29.025	−4.942
2124	LEU271	CB	11.643	−27.582	−5.091
2125	LEU271	CG	13.084	−27.438	−5.586
2126	LEU271	CD1	14.144	−27.477	−4.485
2127	LEU271	CD2	13.397	−28.387	−6.738
2128	LEU271	C	11.86	−29.729	−3.784
2129	LEU271	O	11.618	−29.387	−2.62
2130	GLY272	N	12.606	−30.772	−4.134
2131	GLY272	CA	13.524	−31.531	−3.255
2132	GLY272	C	13.168	−31.602	−1.773
2133	GLY272	O	13.88	−31.032	−0.94
2134	PHE273	N	12.044	−32.247	−1.483
2135	PHE273	CA	11.509	−32.428	−0.116
2136	PHE273	CB	12.342	−33.497	0.587
2137	PHE273	CG	12.381	−34.842	−0.133
2138	PHE273	CD1	11.205	−35.545	−0.363
2139	PHE273	CE1	11.243	−36.768	−1.02
2140	PHE273	CZ	12.458	−37.289	−1.446
2141	PHE273	CE2	13.634	−36.589	−1.215
2142	PHE273	CD2	13.596	−35.366	−0.557
2143	PHE273	C	11.499	−31.159	0.747
2144	PHE273	O	11.865	−31.208	1.926
2145	THR274	N	11.138	−30.035	0.151
2146	THR274	CA	11.051	−28.771	0.881
2147	THR274	CB	12.314	−27.972	0.576
2148	THR274	OG1	13.437	−28.776	0.911
2149	THR274	CG2	12.389	−26.678	1.379
2150	THR274	C	9.827	−27.997	0.41
2151	THR274	O	9.144	−27.336	1.202
2152	LYS275	N	9.546	−28.157	−0.875
2153	LYS275	CA	8.399	−27.522	−1.548
2154	LYS275	CB	7.098	−28.056	−0.963
2155	LYS275	CG	5.883	−27.473	−1.671
2156	LYS275	CD	4.601	−28.076	−1.117
2157	LYS275	CE	4.551	−27.958	0.402
2158	LYS275	NZ	3.289	−28.501	0.926
2159	LYS275	C	8.475	−25.998	−1.459
2160	LYS275	O	8.041	−25.367	−0.487
2161	VAL276	N	8.984	−25.421	−2.531
2162	VAL276	CA	9.29	−23.994	−2.553
2163	VAL276	CB	10.673	−23.833	−3.179
2164	VAL276	CG1	11.116	−22.377	−3.179
2165	VAL276	CG2	11.704	−24.69	−2.451
2166	VAL276	C	8.253	−23.209	−3.353
2167	VAL276	O	8.023	−23.468	−4.541

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2168	ARG277	N	7.59	−22.29	−2.673
2169	ARG277	CA	6.697	−21.357	−3.366
2170	ARG277	CB	5.743	−20.725	−2.362
2171	ARG277	CG	6.393	−20.494	−1.007
2172	ARG277	CD	5.409	−19.842	−0.045
2173	ARG277	NE	4.135	−20.579	−0.024
2174	ARG277	CZ	3.169	−20.346	0.868
2175	ARG277	NH1	3.373	−19.477	1.86
2176	ARG277	NH2	2.019	−21.02	0.799
2177	ARG277	C	7.496	−20.309	−4.136
2178	ARG277	O	8.477	−19.741	−3.635
2179	LEU278	N	7.136	−20.172	−5.4
2180	LEU278	CA	7.872	−19.308	−6.325
2181	LEU278	CB	8.149	−20.086	−7.603
2182	LEU278	CG	8.817	−21.427	−7.34
2183	LEU278	CD1	9.027	−22.164	−8.651
2184	LEU278	CD2	10.142	−21.259	−6.612
2185	LEU278	C	7.104	−18.052	−6.711
2186	LEU278	O	5.872	−18.049	−6.844
2187	LEU279	N	7.887	−17.043	−7.044
2188	LEU279	CA	7.365	−15.749	−7.474
2189	LEU279	CB	7.71	−14.711	−6.414
2190	LEU279	CG	6.893	−13.443	−6.619
2191	LEU279	CD1	5.406	−13.769	−6.548
2192	LEU279	CD2	7.257	−12.384	−5.587
2193	LEU279	C	7.981	−15.347	−8.816
2194	LEU279	O	9.177	−15.043	−8.928
2195	ARG280	N	7.151	−15.343	−9.84
2196	ARG280	CA	7.607	−14.939	−11.169
2197	ARG280	CB	6.765	−15.623	−12.238
2198	ARG280	CG	7.268	−15.273	−13.635
2199	ARG280	CD	6.151	−15.363	−14.67
2200	ARG280	NE	5.458	−16.656	−14.603
2201	ARG280	CZ	4.224	−16.847	−15.074
2202	ARG280	NH1	3.558	−15.832	−15.627
2203	ARG280	NH2	3.655	−18.05	−14.985
2204	ARG280	C	7.462	−13.433	−11.341
2205	ARG280	O	6.36	−12.925	−11.596
2206	LEU281	N	8.556	−12.73	−11.103
2207	LEU281	CA	8.6	−11.3	−11.406
2208	LEU281	CB	9.711	−10.618	−10.615
2209	LEU281	CG	9.368	−10.503	−9.135
2210	LEU281	CD1	10.484	−9.794	−8.378
2211	LEU281	CD2	8.058	−9.75	−8.944
2212	LEU281	C	8.82	−11.068	−12.894
2213	LEU281	O	9.3	−11.952	−13.616
2214	ARG282	N	8.276	−9.954	−13.346
2215	ARG282	CA	8.47	−9.477	−14.713
2216	ARG282	CB	7.143	−9.573	−15.46
2217	ARG282	CG	7.267	−9.025	−16.879
2218	ARG282	CD	5.937	−8.987	−17.617
2219	ARG282	NE	6.12	−8.374	−18.942
2220	ARG282	CZ	5.615	−7.187	−19.283
2221	ARG282	NH1	4.835	−6.521	−18.428
2222	ARG282	NH2	5.85	−6.689	−20.499
2223	ARG282	C	8.952	−8.03	−14.741
2224	ARG282	O	8.376	−7.15	−14.083
2225	ASN283	N	10.066	−7.808	−15.42
2226	ASN283	CA	10.477	−6.436	−15.729
2227	ASN283	CB	11.934	−6.412	−16.188
2228	ASN283	CG	12.342	−5.001	−16.623
2229	ASN283	OD1	11.981	−4.551	−17.725
2230	ASN283	ND2	13.179	−4.368	−15.822
2231	ASN283	C	9.595	−5.928	−16.855
2232	ASN283	O	9.717	−6.408	−17.987
2233	PRO284	N	8.835	−4.879	−16.583
2234	PRO284	CA	7.765	−4.459	−17.494
2235	PRO284	CB	6.934	−3.503	−16.697
2236	PRO284	CG	7.586	−3.269	−15.342
2237	PRO284	CD	8.819	−4.152	−15.313
2238	PRO284	C	8.23	−3.826	−18.815
2239	PRO284	O	7.458	−3.855	−19.778
2240	TRP285	N	9.489	−3.426	−18.93
2241	TRP285	CA	9.994	−2.971	−20.231

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2242	TRP285	CB	11.333	−2.259	−20.051
2243	TRP285	CG	11.312	−0.895	−19.385
2244	TRP285	CD1	10.44	0.142	−19.632
2245	TRP285	NE1	10.786	1.195	−18.848
2246	TRP285	CE2	11.859	0.903	−18.089
2247	TRP285	CZ2	12.585	1.637	−17.164
2248	TRP285	CH2	13.675	1.056	−16.526
2249	TRP285	CZ3	14.042	−0.255	−16.813
2250	TRP285	CE3	13.322	−0.996	−17.742
2251	TRP285	CD2	12.234	−0.422	−18.379
2252	TRP285	C	10.226	−4.202	−21.099
2253	TRP285	O	9.72	−4.303	−22.225
2254	GLY286	N	10.817	−5.2	−20.46
2255	GLY286	CA	11.091	−6.501	−21.077
2256	GLY286	C	12.496	−6.973	−20.71
2257	GLY286	O	12.877	−8.125	−20.97
2258	CYS287	N	13.19	−6.115	−19.983
2259	CYS287	CA	14.623	−6.285	−19.715
2260	CYS287	CB	15.203	−4.886	−19.566
2261	CYS287	SG	14.754	−3.778	−20.919
2262	CYS287	C	14.883	−7.089	−18.447
2263	CYS287	O	14.887	−6.546	−17.337
2264	VAL288	N	15.221	−8.354	−18.635
2265	VAL288	CA	15.436	−9.258	−17.498
2266	VAL288	CB	15.536	−10.688	−18.024
2267	VAL288	CG1	15.495	−11.705	−16.886
2268	VAL288	CG2	14.41	−10.976	−19.009
2269	VAL288	C	16.71	−8.884	−16.737
2270	VAL288	O	16.636	−8.231	−15.688
2271	GLU289	N	17.851	−9.179	−17.345
2272	GLU289	CA	19.175	−8.927	−16.745
2273	GLU289	CB	19.505	−7.429	−16.742
2274	GLU289	CG	19.9	−6.862	−18.11
2275	GLU289	CD	18.699	−6.452	−18.964
2276	GLU289	OE1	18.175	−7.315	−19.661
2277	GLU289	OE2	18.398	−5.268	−18.994
2278	GLU289	C	19.254	−9.464	−15.317
2279	GLU289	O	19.415	−8.695	−14.362
2280	TRP290	N	19.151	−10.776	−15.19
2281	TRP290	CA	19.132	−11.397	−13.867
2282	TRP290	CB	17.699	−11.818	−13.566
2283	TRP290	CG	17.46	−12.272	−12.142
2284	TRP290	CD1	17.784	−11.602	−10.983
2285	TRP290	NE1	17.405	−12.362	−9.925
2286	TRP290	CE2	16.833	−13.511	−10.337
2287	TRP290	CZ2	16.289	−14.588	−9.655
2288	TRP290	CH2	15.756	−15.654	−10.37
2289	TRP290	CZ3	15.766	−15.65	−11.758
2290	TRP290	CE3	16.307	−14.575	−12.45
2291	TRP290	CD2	16.837	−13.509	−11.74
2292	TRP290	C	20.085	−12.589	−13.806
2293	TRP290	O	20.326	−13.278	−14.807
2294	THR291	N	20.677	−12.765	−12.639
2295	THR291	CA	21.647	−13.835	−12.409
2296	THR291	CB	22.463	−13.454	−11.176
2297	THR291	OG1	23.074	−12.194	−11.426
2298	THR291	CG2	23.564	−14.466	−10.876
2299	THR291	C	20.97	−15.189	−12.19
2300	THR291	O	20.244	−15.397	−11.211
2301	GLY292	N	21.195	−16.087	−13.135
2302	GLY292	CA	20.753	−17.48	−12.994
2303	GLY292	C	19.306	−17.705	−13.417
2304	GLY292	O	18.485	−16.782	−13.421
2305	ALA293	N	18.977	−18.975	−13.597
2306	ALA293	CA	17.63	−19.385	−14.019
2307	ALA293	CB	17.747	−20.668	−14.835
2308	ALA293	C	16.679	−19.623	−12.846
2309	ALA293	O	15.55	−20.083	−13.055
2310	TRP294	N	17.167	−19.364	−11.641
2311	TRP294	CA	16.396	−19.559	−10.41
2312	TRP294	CB	16.062	−21.037	−10.261
2313	TRP294	CG	15.007	−21.339	−9.221
2314	TRP294	CD1	14.223	−20.43	−8.547
2315	TRP294	NE1	13.39	−21.111	−7.724

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2316	TRP294	CE2	13.577	−22.438	−7.831
2317	TRP294	CZ2	12.948	−23.526	−7.249
2318	TRP294	CH2	13.358	−24.812	−7.575
2319	TRP294	CZ3	14.392	−25.016	−8.482
2320	TRP294	CE3	15.021	−23.931	−9.077
2321	TRP294	CD2	14.613	−22.647	−8.76
2322	TRP294	C	17.224	−19.121	−9.205
2323	TRP294	O	18.361	−19.579	−9.024
2324	SER295	N	16.696	−18.169	−8.452
2325	SER295	CA	17.315	−17.801	−7.173
2326	SER295	CB	16.566	−16.63	−6.529
2327	SER295	OG	15.219	−17.008	−6.253
2328	SER295	C	17.266	−19.032	−6.282
2329	SER295	O	16.19	−19.611	−6.104
2330	ASP296	N	18.418	−19.423	−5.758
2331	ASP296	CA	18.579	−20.71	−5.06
2332	ASP296	CB	17.617	−20.813	−3.875
2333	ASP296	CG	17.753	−22.161	−3.175
2334	ASP296	OD1	16.727	−22.742	−2.851
2335	ASP296	OD2	18.884	−22.599	−3.004
2336	ASP296	C	18.376	−21.873	−6.034
2337	ASP296	O	17.283	−22.068	−6.581
2338	SER297	N	19.44	−22.651	−6.179
2339	SER297	CA	19.519	−23.823	−7.071
2340	SER297	CB	18.802	−25.017	−6.431
2341	SER297	OG	17.412	−24.765	−6.277
2342	SER297	C	19.007	−23.54	−8.483
2343	SER297	O	18.073	−24.193	−8.964
2344	CYS298	N	19.806	−22.768	−9.206
2345	CYS298	CA	19.449	−22.281	−10.553
2346	CYS298	CB	20.62	−21.481	−11.127
2347	CYS298	SG	21.234	−20.118	−10.112
2348	CYS298	C	19.003	−23.365	−11.556
2349	CYS298	O	17.904	−23.209	−12.098
2350	PRO299	N	19.793	−24.394	−11.864
2351	PRO299	CA	19.32	−25.416	−12.81
2352	PRO299	CB	20.579	−25.994	−13.38
2353	PRO299	CG	21.733	−25.664	−12.445
2354	PRO299	CD	21.165	−24.701	−11.417
2355	PRO299	C	18.493	−26.55	−12.186
2356	PRO299	O	18.054	−27.446	−12.919
2357	ARG300	N	18.175	−26.476	−10.902
2358	ARG300	CA	17.675	−27.661	−10.197
2359	ARG300	CB	18.239	−27.675	−8.788
2360	ARG300	CG	19.759	−27.786	−8.865
2361	ARG300	CD	20.391	−28.019	−7.496
2362	ARG300	NE	21.85	−28.151	−7.626
2363	ARG300	CZ	22.48	−29.327	−7.67
2364	ARG300	NH1	21.79	−30.463	−7.538
2365	ARG300	NH2	23.806	−29.365	−7.812
2366	ARG300	C	16.159	−27.841	−10.19
2367	ARG300	O	15.632	−28.578	−9.348
2368	TRP301	N	15.524	−27.426	−11.274
2369	TRP301	CA	14.092	−27.665	−11.467
2370	TRP301	CB	13.63	−26.879	−12.692
2371	TRP301	CG	13.87	−25.38	−12.66
2372	TRP301	CD1	14.944	−24.694	−13.187
2373	TRP301	NE1	14.753	−23.368	−12.976
2374	TRP301	CE2	13.59	−23.139	−12.337
2375	TRP301	CZ2	12.936	−21.976	−11.962
2376	TRP301	CH2	11.718	−22.052	−11.3
2377	TRP301	CZ3	11.149	−23.291	−11.022
2378	TRP301	CE3	11.785	−24.461	−11.414
2379	TRP301	CD2	12.998	−24.389	−12.077
2380	TRP301	C	13.858	−29.15	−11.744
2381	TRP301	O	12.959	−29.755	−11.15
2382	ASP302	N	14.884	−29.773	−12.304
2383	ASP302	CA	14.844	−31.199	−12.653
2384	ASP302	CB	15.93	−31.466	−13.693
2385	ASP302	CG	15.73	−30.612	−14.942
2386	ASP302	OD1	14.585	−30.345	−15.283
2387	ASP302	OD2	16.731	−30.218	−15.522
2388	ASP302	C	15.069	−32.151	−11.469
2389	ASP302	O	15.12	−33.367	−11.686

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2390	THR303	N	15.193	−31.643	−10.251
2391	THR303	CA	15.388	−32.541	−9.11
2392	THR303	CB	16.216	−31.853	−8.031
2393	THR303	OG1	15.475	−30.751	−7.526
2394	THR303	CG2	17.548	−31.347	−8.572
2395	THR303	C	14.07	−33.016	−8.495
2396	THR303	O	14.09	−33.929	−7.662
2397	LEU304	N	12.954	−32.4	−8.861
2398	LEU304	CA	11.65	−32.904	−8.388
2399	LEU304	CB	11.425	−32.587	−6.909
2400	LEU304	CG	10.292	−33.433	−6.335
2401	LEU304	CD1	10.569	−34.915	−6.556
2402	LEU304	CD2	10.053	−33.14	−4.858
2403	LEU304	C	10.44	−32.464	−9.242
2404	LEU304	O	9.673	−33.362	−9.61
2405	PRO305	N	10.189	−31.174	−9.485
2406	PRO305	CA	9.052	−30.811	−10.343
2407	PRO305	CB	9.039	−29.314	−10.413
2408	PRO305	CG	10.174	−28.762	−9.572
2409	PRO305	CD	10.876	−29.971	−8.98
2410	PRO305	C	9.163	−31.433	−11.732
2411	PRO305	O	10.125	−31.215	−12.474
2412	THR306	N	8.133	−32.19	−12.071
2413	THR306	CA	8.087	−32.943	−13.33
2414	THR306	CB	6.849	−33.829	−13.326
2415	THR306	OG1	5.713	−32.99	−13.515
2416	THR306	CG2	6.703	−34.585	−12.009
2417	THR306	C	7.991	−32.044	−14.553
2418	THR306	O	7.775	−30.829	−14.446
2419	GLU307	N	7.85	−32.706	−15.691
2420	GLU307	CA	7.745	−32.024	−16.986
2421	GLU307	CB	7.892	−33.082	−18.075
2422	GLU307	CG	7.888	−32.477	−19.475
2423	GLU307	CD	8.046	−33.582	−20.517
2424	GLU307	OE1	9.184	−33.942	−20.788
2425	GLU307	OE2	7.03	−34.14	−20.908
2426	GLU307	C	6.428	−31.257	−17.174
2427	GLU307	O	6.442	−30.253	−17.894
2428	CYS308	N	5.434	−31.492	−16.328
2429	CYS308	CA	4.183	−30.742	−16.455
2430	CYS308	CB	3.067	−31.536	−15.788
2431	CYS308	SG	1.437	−30.758	−15.821
2432	CYS308	C	4.33	−29.382	−15.777
2433	CYS308	O	3.958	−28.358	−16.361
2434	ARG309	N	5.171	−29.358	−14.754
2435	ARG309	CA	5.471	−28.109	−14.06
2436	ARG309	CB	5.947	−28.425	−12.647
2437	ARG309	CG	6.263	−27.146	−11.879
2438	ARG309	CD	5.02	−26.286	−11.684
2439	ARG309	NE	4.068	−26.945	−10.778
2440	ARG309	CZ	3.23	−26.271	−9.988
2441	ARG309	NH1	3.191	−24.938	−10.042
2442	ARG309	NH2	2.407	−26.93	−9.17
2443	ARG309	C	6.554	−27.35	−14.815
2444	ARG309	O	6.479	−26.122	−14.91
2445	ASP310	N	7.331	−28.083	−15.598
2446	ASP310	CA	8.324	−27.455	−16.474
2447	ASP310	CB	9.299	−28.517	−16.971
2448	ASP310	CG	10.07	−29.145	−15.812
2449	ASP310	OD1	10.426	−28.411	−14.9
2450	ASP310	OD2	10.369	−30.328	−15.915
2451	ASP310	C	7.662	−26.775	−17.673
2452	ASP310	O	8.097	−25.681	−18.056
2453	ALA311	N	6.485	−27.253	−18.055
2454	ALA311	CA	5.707	−26.614	−19.123
2455	ALA311	CB	4.77	−27.656	−19.726
2456	ALA311	C	4.894	−25.423	−18.609
2457	ALA311	O	4.414	−24.603	−19.399
2458	LEU312	N	4.828	−25.289	−17.294
2459	LEU312	CA	4.235	−24.115	−16.652
2460	LEU312	CB	3.599	−24.558	−15.34
2461	LEU312	CG	2.08	−24.528	−15.404
2462	LEU312	CD1	1.483	−25.072	−14.111
2463	LEU312	CD2	1.584	−23.111	−15.674

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2464	LEU312	C	5.284	−23.051	−16.332
2465	LEU312	O	4.931	−21.944	−15.905
2466	LEU313	N	6.546	−23.359	−16.578
2467	LEU313	CA	7.628	−22.449	−16.191
2468	LEU313	CB	8.902	−23.238	−15.922
2469	LEU313	CG	8.782	−24.11	−14.683
2470	LEU313	CD1	10.056	−24.916	−14.466
2471	LEU313	CD2	8.449	−23.28	−13.449
2472	LEU313	C	7.945	−21.393	−17.238
2473	LEU313	O	7.259	−20.37	−17.344
2474	VAL314	N	9.104	−21.55	−17.848
2475	VAL314	CA	9.621	−20.546	−18.778
2476	VAL314	CB	11.134	−20.715	−18.897
2477	VAL314	CG1	11.732	−19.697	−19.866
2478	VAL314	CG2	11.795	−20.592	−17.527
2479	VAL314	C	8.958	−20.662	−20.143
2480	VAL314	O	9.219	−21.591	−20.916
2481	LYS315	N	8.015	−19.765	−20.361
2482	LYS315	CA	7.334	−19.654	−21.648
2483	LYS315	CB	6.076	−20.521	−21.624
2484	LYS315	CG	5.42	−20.592	−23
2485	LYS315	CD	4.271	−21.593	−23.024
2486	LYS315	CE	4.759	−22.998	−22.683
2487	LYS315	NZ	3.655	−23.97	−22.741
2488	LYS315	C	6.983	−18.191	−21.899
2489	LYS315	O	6.796	−17.757	−23.044
2490	LYS316	N	6.984	−17.425	−20.821
2491	LYS316	CA	6.653	−16.001	−20.908
2492	LYS316	CB	6.246	−15.5	−19.527
2493	LYS316	CG	5.022	−16.242	−19.003
2494	LYS316	CD	3.799	−15.998	−19.883
2495	LYS316	CE	3.425	−14.52	−19.927
2496	LYS316	NZ	2.273	−14.299	−20.816
2497	LYS316	C	7.837	−15.195	−21.42
2498	LYS316	O	8.931	−15.211	−20.842
2499	GLU317	N	7.615	−14.548	−22.549
2500	GLU317	CA	8.638	−13.698	−23.156
2501	GLU317	CB	8.408	−13.676	−24.663
2502	GLU317	CG	8.493	−15.09	−25.236
2503	GLU317	CD	8.206	−15.086	−26.735
2504	GLU317	OE1	8.856	−15.843	−27.441
2505	GLU317	OE2	7.384	−14.282	−27.15
2506	GLU317	C	8.571	−12.292	−22.565
2507	GLU317	O	7.515	−11.861	−22.092
2508	ASP318	N	9.71	−11.614	−22.582
2509	ASP318	CA	9.857	−10.248	−22.037
2510	ASP318	CB	8.888	−9.278	−22.713
2511	ASP318	CG	9.177	−9.215	−24.209
2512	ASP318	OD1	8.245	−9.432	−24.969
2513	ASP318	OD2	10.353	−9.242	−24.547
2514	ASP318	C	9.667	−10.211	−20.522
2515	ASP318	O	8.546	−10.249	−20.003
2516	GLY319	N	10.791	−10.138	−19.831
2517	GLY319	CA	10.798	−10.041	−18.368
2518	GLY319	C	10.5	−11.357	−17.648
2519	GLY319	O	9.85	−11.344	−16.6
2520	GLU320	N	11.036	−12.455	−18.151
2521	GLU320	CA	10.831	−13.766	−17.514
2522	GLU320	CB	11.014	−14.837	−18.59
2523	GLU320	CG	10.796	−16.265	−18.085
2524	GLU320	CD	9.329	−16.539	−17.763
2525	GLU320	OE1	8.827	−15.932	−16.826
2526	GLU320	OE2	8.77	−17.448	−18.37
2527	GLU320	C	11.843	−13.998	−16.39
2528	GLU320	O	13.006	−14.313	−16.671
2529	PHE321	N	11.412	−13.864	−15.145
2530	PHE321	CA	12.362	−14.039	−14.034
2531	PHE321	CB	12.172	−12.939	−12.996
2532	PHE321	CG	12.867	−11.627	−13.353
2533	PHE321	CD1	12.409	−10.846	−14.407
2534	PHE321	CE1	13.062	−9.668	−14.736
2535	PHE321	CZ	14.171	−9.262	−14.008
2536	PHE321	CE2	14.623	−10.033	−12.947
2537	PHE321	CD2	13.971	−11.215	−12.62

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2538	PHE321	C	12.29	−15.402	−13.355
2539	PHE321	O	13.259	−16.166	−13.446
2540	TRP322	N	11.161	−15.706	−12.732
2541	TRP322	CA	11.002	−16.943	−11.94
2542	TRP322	CB	11.085	−18.176	−12.832
2543	TRP322	CG	9.775	−18.557	−13.477
2544	TRP322	CD1	9.44	−18.447	−14.807
2545	TRP322	NE1	8.178	−18.918	−14.967
2546	TRP322	CE2	7.661	−19.339	−13.798
2547	TRP322	CZ2	6.449	−19.921	−13.463
2548	TRP322	CH2	6.19	−20.255	−12.14
2549	TRP322	CZ3	7.141	−20.019	−11.154
2550	TRP322	CE3	8.365	−19.451	−11.485
2551	TRP322	CD2	8.627	−19.116	−12.804
2552	TRP322	C	12.025	−17.089	−10.813
2553	TRP322	O	13.103	−17.672	−10.981
2554	MET323	N	11.672	−16.552	−9.662
2555	MET323	CA	12.503	−16.738	−8.474
2556	MET323	CB	13.051	−15.388	−8.034
2557	MET323	CG	11.953	−14.374	−7.754
2558	MET323	SD	12.542	−12.692	−7.479
2559	MET323	CE	13.295	−12.396	−9.094
2560	MET323	C	11.687	−17.391	−7.369
2561	MET323	O	10.544	−17.801	−7.587
2562	GLU324	N	12.327	−17.631	−6.242
2563	GLU324	CA	11.572	−18.162	−5.09
2564	GLU324	CB	12.326	−19.241	−4.294
2565	GLU324	CG	13.851	−19.276	−4.39
2566	GLU324	CD	14.55	−18.15	−3.631
2567	GLU324	OE1	13.952	−17.087	−3.552
2568	GLU324	OE2	15.715	−18.308	−3.299
2569	GLU324	C	11.053	−17.053	−4.174
2570	GLU324	O	11.246	−15.859	−4.434
2571	LEU325	N	10.428	−17.48	−3.086
2572	LEU325	CA	9.835	−16.589	−2.066
2573	LEU325	CB	8.883	−17.399	−1.177
2574	LEU325	CG	9.599	−18.25	−0.118
2575	LEU325	CD1	8.769	−18.36	1.159
2576	LEU325	CD2	10.039	−19.633	−0.606
2577	LEU325	C	10.838	−15.856	−1.156
2578	LEU325	O	10.424	−15.087	−0.282
2579	ARG326	N	12.125	−15.95	−1.452
2580	ARG326	CA	13.164	−15.288	−0.657
2581	ARG326	CB	14.38	−16.198	−0.652
2582	ARG326	CG	15.353	−15.917	0.481
2583	ARG326	CD	16.521	−16.887	0.374
2584	ARG326	NE	16.008	−18.246	0.138
2585	ARG326	CZ	16.446	−19.331	0.779
2586	ARG326	NH1	15.908	−20.524	0.511
2587	ARG326	NH2	17.409	−19.221	1.697
2588	ARG326	C	13.501	−13.921	−1.269
2589	ARG326	O	14.402	−13.206	−0.811
2590	ASP327	N	12.631	−13.497	−2.176
2591	ASP327	CA	12.731	−12.219	−2.897
2592	ASP327	CB	11.71	−12.254	−4.034
2593	ASP327	CG	10.27	−12.326	−3.517
2594	ASP327	OD1	9.656	−11.276	−3.407
2595	ASP327	OD2	9.797	−13.428	−3.282
2596	ASP327	C	12.462	−10.975	−2.043
2597	ASP327	O	12.783	−9.874	−2.501
2598	PHE328	N	12.167	−11.169	−0.762
2599	PHE328	CA	11.856	−10.085	0.186
2600	PHE328	CB	11.127	−10.703	1.375
2601	PHE328	CG	9.828	−11.432	1.033
2602	PHE328	CD1	9.542	−12.647	1.641
2603	PHE328	CE1	8.363	−13.317	1.339
2604	PHE328	CZ	7.468	−12.772	0.427
2605	PHE328	CE2	7.753	−11.557	−0.18
2606	PHE328	CD2	8.931	−10.886	0.124
2607	PHE328	C	13.102	−9.351	0.699
2608	PHE328	O	13.005	−8.488	1.58
2609	LEU329	N	14.267	−9.73	0.189
2610	LEU329	CA	15.49	−8.967	0.439
2611	LEU329	CB	16.685	−9.846	0.087

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2612	LEU329	CG	16.714	−11.138	0.896
2613	LEU329	CD1	17.774	−12.095	0.361
2614	LEU329	CD2	16.935	−10.857	2.38
2615	LEU329	C	15.518	−7.73	−0.459
2616	LEU329	O	16.242	−6.764	−0.187
2617	LEU330	N	14.733	−7.776	−1.523
2618	LEU330	CA	14.554	−6.628	−2.407
2619	LEU330	CB	14.821	−7.067	−3.844
2620	LEU330	CG	16.207	−7.669	−4.041
2621	LEU330	CD1	16.338	−8.272	−5.435
2622	LEU330	CD2	17.297	−6.633	−3.804
2623	LEU330	C	13.109	−6.151	−2.301
2624	LEU330	O	12.235	−6.909	−1.868
2625	HIS331	N	12.892	−4.907	−2.702
2626	HIS331	CA	11.552	−4.29	−2.803
2627	HIS331	CB	10.704	−5.064	−3.813
2628	HIS331	CG	11.369	−5.321	−5.157
2629	HIS331	ND1	11.608	−4.423	−6.134
2630	HIS331	CE1	12.232	−5.046	−7.155
2631	HIS331	NE2	12.351	−6.353	−6.837
2632	HIS331	CD2	11.81	−6.541	−5.613
2633	HIS331	C	10.828	−4.221	−1.458
2634	HIS331	O	10.363	−5.234	−0.927
2635	PHE332	N	10.643	−3.012	−0.96
2636	PHE332	CA	10.019	−2.869	0.358
2637	PHE332	CB	10.751	−1.814	1.199
2638	PHE332	CG	10.611	−0.328	0.854
2639	PHE332	CD1	10.345	0.567	1.884
2640	PHE332	CE1	10.21	1.924	1.612
2641	PHE332	CZ	10.349	2.387	0.311
2642	PHE332	CE2	10.631	1.5	−0.716
2643	PHE332	CD2	10.77	0.147	−0.444
2644	PHE332	C	8.521	−2.571	0.259
2645	PHE332	O	7.765	−2.965	1.155
2646	ASP333	N	8.08	−2.042	−0.873
2647	ASP333	CA	6.639	−1.806	−1.083
2648	ASP333	CB	6.371	−0.333	−1.385
2649	ASP333	CG	6.41	0.542	−0.133
2650	ASP333	OD1	5.398	1.177	0.128
2651	ASP333	OD2	7.496	0.729	0.393
2652	ASP333	C	6.127	−2.637	−2.248
2653	ASP333	O	6.868	−2.889	−3.201
2654	THR334	N	4.885	−3.076	−2.164
2655	THR334	CA	4.261	−3.823	−3.268
2656	THR334	CB	4.451	−5.315	−3.03
2657	THR334	OG1	5.842	−5.569	−3.102
2658	THR334	CG2	3.784	−6.155	−4.111
2659	THR334	C	2.776	−3.501	−3.398
2660	THR334	O	1.927	−4.082	−2.71
2661	VAL335	N	2.467	−2.604	−4.315
2662	VAL335	CA	1.071	−2.202	−4.518
2663	VAL335	CB	1.056	−0.764	−5.033
2664	VAL335	CG1	2.077	−0.54	−6.143
2665	VAL335	CG2	−0.339	−0.33	−5.468
2666	VAL335	C	0.35	−3.142	−5.485
2667	VAL335	O	0.794	−3.352	−6.619
2668	GLN336	N	−0.696	−3.783	−4.99
2669	GLN336	CA	−1.522	−4.653	−5.835
2670	GLN336	CB	−1.716	−5.997	−5.142
2671	GLN336	CG	−0.402	−6.545	−4.603
2672	GLN336	CD	−0.509	−8.044	−4.339
2673	GLN336	OE1	0.381	−8.807	−4.734
2674	GLN336	NE2	−1.581	−8.458	−3.687
2675	GLN336	C	−2.89	−4.022	−6.072
2676	GLN336	O	−3.719	−3.969	−5.156
2677	ILE337	N	−3.138	−3.56	−7.283
2678	ILE337	CA	−4.451	−2.964	−7.555
2679	ILE337	CB	−4.264	−1.664	−8.357
2680	ILE337	CG2	−4.066	−1.889	−9.851
2681	ILE337	CG1	−5.401	−0.667	−8.135
2682	ILE337	CD1	−6.725	−1.087	−8.757
2683	ILE337	C	−5.384	−4.008	−8.187
2684	ILE337	O	−5.196	−4.483	−9.315
2685	CYS338	N	−6.328	−4.44	−7.371

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2686	CYS338	CA	−7.307	−5.454	−7.775
2687	CYS338	CB	−7.882	−6.072	−6.508
2688	CYS338	SG	−6.674	−6.764	−5.355
2689	CYS338	C	−8.436	−4.846	−8.597
2690	CYS338	O	−8.987	−3.804	−8.222
2691	SER339	N	−8.796	−5.51	−9.68
2692	SER339	CA	−9.877	−5.015	−10.54
2693	SER339	CB	−9.806	−5.746	−11.871
2694	SER339	OG	−8.577	−5.402	−12.496
2695	SER339	C	−11.255	−5.198	−9.899
2696	SER339	O	−11.434	−6.011	−8.984
2697	LEU340	N	−12.193	−4.373	−10.337
2698	LEU340	CA	−13.576	−4.397	−9.821
2699	LEU340	CB	−14.293	−3.088	−10.177
2700	LEU340	CG	−13.922	−2.471	−11.533
2701	LEU340	CD1	−14.395	−3.295	−12.723
2702	LEU340	CD2	−14.511	−1.07	−11.654
2703	LEU340	C	−14.386	−5.623	−10.258
2704	LEU340	O	−13.829	−6.689	−10.558
2705	SER341	N	−15.703	−5.494	−10.181
2706	SER341	CA	−16.591	−6.638	−10.444
2707	SER341	CB	−17.981	−6.379	−9.856
2708	SER341	OG	−17.842	−6.062	−8.478
2709	SER341	C	−16.627	−7.018	−11.94
2710	SER341	O	−16.027	−8.051	−12.268
2711	PRO342	N	−17.217	−6.24	−12.848
2712	PRO342	CA	−16.988	−6.519	−14.27
2713	PRO342	CB	−18.101	−5.813	−14.98
2714	PRO342	CG	−18.689	−4.778	−14.032
2715	PRO342	CD	−17.992	−4.996	−12.698
2716	PRO342	C	−15.641	−5.944	−14.696
2717	PRO342	O	−15.574	−4.808	−15.175
2718	GLU343	N	−14.593	−6.744	−14.58
2719	GLU343	CA	−13.23	−6.237	−14.785
2720	GLU343	CB	−12.234	−7.022	−13.919
2721	GLU343	CG	−12.614	−8.468	−13.575
2722	GLU343	CD	−12.502	−9.414	−14.773
2723	GLU343	OE1	−13.501	−9.506	−15.473
2724	GLU343	OE2	−11.378	−9.747	−15.116
2725	GLU343	C	−12.771	−6.158	−16.244
2726	GLU343	O	−11.862	−5.377	−16.536
2727	VAL344	N	−13.446	−6.837	−17.155
2728	VAL344	CA	−13.109	−6.703	−18.576
2729	VAL344	CB	−11.832	−7.483	−18.905
2730	VAL344	CG1	−11.831	−8.896	−18.333
2731	VAL344	CG2	−11.546	−7.503	−20.404
2732	VAL344	C	−14.267	−7.174	−19.439
2733	VAL344	O	−14.695	−8.328	−19.321
2734	LEU345	N	−14.787	−6.242	−20.229
2735	LEU345	CA	−15.882	−6.475	−21.188
2736	LEU345	CB	−15.267	−6.97	−22.492
2737	LEU345	CG	−16.293	−6.985	−23.62
2738	LEU345	CD1	−16.959	−5.62	−23.766
2739	LEU345	CD2	−15.65	−7.417	−24.934
2740	LEU345	C	−16.92	−7.471	−20.671
2741	LEU345	O	−16.877	−8.668	−20.988
2742	GLY346	N	−17.777	−6.971	−19.794
2743	GLY346	CA	−18.819	−7.796	−19.171
2744	GLY346	C	−18.251	−9.025	−18.464
2745	GLY346	O	−17.351	−8.937	−17.614
2746	PRO347	N	−18.828	−10.168	−18.805
2747	PRO347	CA	−18.506	−11.452	−18.177
2748	PRO347	CB	−19.752	−12.261	−18.363
2749	PRO347	CG	−20.577	−11.632	−19.477
2750	PRO347	CD	−19.908	−10.302	−19.785
2751	PRO347	C	−17.338	−12.2	−18.822
2752	PRO347	O	−17.504	−13.385	−19.141
2753	SER348	N	−16.198	−11.559	−19.033
2754	SER348	CA	−15.058	−12.298	−19.597
2755	SER348	CB	−13.953	−11.325	−19.987
2756	SER348	OG	−14.463	−10.495	−21.022
2757	SER348	C	−14.549	−13.329	−18.59
2758	SER348	O	−14.243	−13.005	−17.439
2759	PRO349	N	−14.424	−14.565	−19.051

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2760	PRO349	CA	−14.432	−15.734	−18.147
2761	PRO349	CB	−14.813	−16.889	−19.022
2762	PRO349	CG	−14.821	−16.447	−20.476
2763	PRO349	CD	−14.576	−14.95	−20.457
2764	PRO349	C	−13.111	−16.055	−17.43
2765	PRO349	O	−13.063	−17.02	−16.656
2766	GLU350	N	−12.061	−15.283	−17.654
2767	GLU350	CA	−10.773	−15.645	−17.054
2768	GLU350	CB	−9.703	−15.689	−18.138
2769	GLU350	CG	−8.384	−16.207	−17.575
2770	GLU350	CD	−7.401	−16.49	−18.708
2771	GLU350	OE1	−7.869	−16.844	−19.782
2772	GLU350	OE2	−6.206	−16.421	−18.458
2773	GLU350	C	−10.383	−14.696	−15.925
2774	GLU350	O	−9.773	−13.641	−16.137
2775	GLY351	N	−10.695	−15.134	−14.718
2776	GLY351	CA	−10.363	−14.366	−13.516
2777	GLY351	C	−11.622	−13.991	−12.745
2778	GLY351	O	−12.663	−13.682	−13.338
2779	GLY352	N	−11.494	−13.953	−11.433
2780	GLY352	CA	−12.627	−13.621	−10.572
2781	GLY352	C	−12.764	−12.111	−10.438
2782	GLY352	O	−11.805	−11.356	−10.641
2783	GLY353	N	−13.989	−11.684	−10.194
2784	GLY353	CA	−14.269	−10.268	−9.962
2785	GLY353	C	−14.916	−10.093	−8.596
2786	GLY353	O	−15.571	−11.008	−8.081
2787	TRP354	N	−14.701	−8.935	−8.002
2788	TRP354	CA	−15.292	−8.653	−6.688
2789	TRP354	CB	−14.706	−7.367	−6.122
2790	TRP354	CG	−13.274	−7.455	−5.646
2791	TRP354	CD1	−12.233	−6.651	−6.047
2792	TRP354	NE1	−11.113	−7.037	−5.39
2793	TRP354	CE2	−11.365	−8.063	−4.553
2794	TRP354	CZ2	−10.567	−8.789	−3.681
2795	TRP354	CH2	−11.125	−9.813	−2.924
2796	TRP354	CZ3	−12.478	−10.114	−3.038
2797	TRP354	CE3	−13.284	−9.396	−3.914
2798	TRP354	CD2	−12.732	−8.374	−4.672
2799	TRP354	C	−16.805	−8.483	−6.744
2800	TRP354	O	−17.46	−8.689	−7.771
2801	HIS355	N	−17.357	−8.218	−5.575
2802	HIS355	CA	−18.748	−7.77	−5.485
2803	HIS355	CB	−19.539	−8.715	−4.587
2804	HIS355	CG	−21.042	−8.517	−4.658
2805	HIS355	ND1	−21.762	−7.588	−4
2806	HIS355	CE1	−23.064	−7.723	−4.322
2807	HIS355	NE2	−23.165	−8.749	−5.198
2808	HIS355	CD2	−21.926	−9.247	−5.416
2809	HIS355	C	−18.74	−6.356	−4.909
2810	HIS355	O	−19.012	−6.155	−3.72
2811	VAL356	N	−18.394	−5.398	−5.754
2812	VAL356	CA	−18.176	−4.01	−5.316
2813	VAL356	CB	−17.632	−3.215	−6.512
2814	VAL356	CG1	−18.64	−3.134	−7.654
2815	VAL356	CG2	−17.169	−1.812	−6.125
2816	VAL356	C	−19.449	−3.365	−4.756
2817	VAL356	O	−20.554	−3.566	−5.273
2818	HIS357	N	−19.295	−2.749	−3.595
2819	HIS357	CA	−20.379	−1.972	−2.989
2820	HIS357	CB	−20.777	−2.602	−1.66
2821	HIS357	CG	−21.865	−1.826	−0.945
2822	HIS357	ND1	−23.177	−1.843	−1.24
2823	HIS357	CE1	−23.835	−1.02	−0.398
2824	HIS357	NE2	−22.923	−0.476	0.44
2825	HIS357	CD2	−21.704	−0.964	0.116
2826	HIS357	C	−19.931	−0.53	−2.764
2827	HIS357	O	−19.011	−0.262	−1.983
2828	THR358	N	−20.556	0.381	−3.488
2829	THR358	CA	−20.212	1.803	−3.36
2830	THR358	CB	−20.25	2.457	−4.736
2831	THR358	OG1	−21.574	2.356	−5.243
2832	THR358	CG2	−19.305	1.767	−5.713
2833	THR358	C	−21.175	2.529	−2.427

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2834	THR358	O	−22.391	2.312	−2.472
2835	PHE359	N	−20.617	3.388	−1.591
2836	PHE359	CA	−21.431	4.174	−0.655
2837	PHE359	CB	−21.406	3.514	0.726
2838	PHE359	CG	−20.028	3.253	1.343
2839	PHE359	CD1	−19.343	4.272	1.993
2840	PHE359	CE1	−18.098	4.027	2.556
2841	PHE359	CZ	−17.538	2.758	2.478
2842	PHE359	CE2	−18.226	1.735	1.839
2843	PHE359	CD2	−19.471	1.981	1.275
2844	PHE359	C	−20.958	5.626	−0.563
2845	PHE359	O	−19.768	5.922	−0.724
2846	GLN360	N	−21.912	6.525	−0.388
2847	GLN360	CA	−21.597	7.946	−0.174
2848	GLN360	CB	−22.6	8.807	−0.94
2849	GLN360	CG	−22.566	8.509	−2.437
2850	GLN360	CD	−23.504	9.443	−3.204
2851	GLN360	OE1	−24.71	9.51	−2.933
2852	GLN360	NE2	−22.938	10.127	−4.185
2853	GLN360	C	−21.647	8.279	1.319
2854	GLN360	O	−22.381	7.633	2.076
2855	GLY361	N	−20.866	9.262	1.739
2856	GLY361	CA	−20.852	9.644	3.16
2857	GLY361	C	−20.314	11.055	3.399
2858	GLY361	O	−19.127	11.33	3.188
2859	ARG362	N	−21.193	11.937	3.847
2860	ARG362	CA	−20.795	13.321	4.139
2861	ARG362	CB	−21.846	14.283	3.596
2862	ARG362	CG	−21.887	14.327	2.073
2863	ARG362	CD	−22.923	15.344	1.606
2864	ARG362	NE	−22.916	15.506	0.143
2865	ARG362	CZ	−24.011	15.386	−0.611
2866	ARG362	NH1	−23.958	15.676	−1.914
2867	ARG362	NH2	−25.179	15.077	−0.045
2868	ARG362	C	−20.644	13.578	5.635
2869	ARG362	O	−21.596	13.427	6.409
2870	TRP363	N	−19.481	14.08	6.011
2871	TRP363	CA	−19.25	14.487	7.402
2872	TRP363	CB	−17.815	14.169	7.799
2873	TRP363	CG	−17.561	12.709	8.097
2874	TRP363	CD1	−17.845	12.061	9.276
2875	TRP363	NE1	−17.466	10.765	9.157
2876	TRP363	CE2	−16.942	10.52	7.941
2877	TRP363	CZ2	−16.424	9.377	7.351
2878	TRP363	CH2	−15.942	9.432	6.048
2879	TRP363	CZ3	−15.977	10.626	5.334
2880	TRP363	CE3	−16.495	11.776	5.917
2881	TRP363	CD2	−16.977	11.725	7.216
2882	TRP363	C	−19.501	15.976	7.61
2883	TRP363	O	−18.555	16.761	7.743
2884	VAL364	N	−20.77	16.346	7.677
2885	VAL364	CA	−21.141	17.739	7.96
2886	VAL364	CB	−22.651	17.857	7.784
2887	VAL364	CG1	−23.139	19.277	8.048
2888	VAL364	CG2	−23.071	17.394	6.394
2889	VAL364	C	−20.762	18.079	9.397
2890	VAL364	O	−21.284	17.443	10.32
2891	ARG365	N	−19.882	19.055	9.577
2892	ARG365	CA	−19.35	19.406	10.909
2893	ARG365	CB	−18.604	20.73	10.815
2894	ARG365	CG	−17.345	20.613	9.967
2895	ARG365	CD	−16.587	21.935	9.957
2896	ARG365	NE	−16.309	22.372	11.335
2897	ARG365	CZ	−15.228	23.074	11.681
2898	ARG365	NH1	−14.322	23.405	10.758
2899	ARG365	NH2	−15.047	23.435	12.954
2900	ARG365	C	−20.432	19.526	11.972
2901	ARG365	O	−21.51	20.075	11.713
2902	GLY366	N	−20.222	18.82	13.073
2903	GLY366	CA	−21.18	18.847	14.186
2904	GLY366	C	−22.293	17.799	14.077
2905	GLY366	O	−22.469	16.973	14.979
2906	PHE367	N	−23.04	17.848	12.986
2907	PHE367	CA	−24.222	16.998	12.834

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2908	PHE367	CB	−25.11	17.646	11.777
2909	PHE367	CG	−26.514	17.06	11.669
2910	PHE367	CD1	−27.5	17.474	12.555
2911	PHE367	CE1	−28.783	16.95	12.462
2912	PHE367	CZ	−29.08	16.013	11.481
2913	PHE367	CE2	−28.095	15.6	10.593
2914	PHE367	CD2	−26.813	16.125	10.686
2915	PHE367	C	−23.859	15.576	12.413
2916	PHE367	O	−24.44	14.61	12.922
2917	ASN368	N	−22.871	15.443	11.543
2918	ASN368	CA	−22.421	14.112	11.122
2919	ASN368	CB	−22.691	13.911	9.639
2920	ASN368	CG	−24.188	13.829	9.386
2921	ASN368	OD1	−24.746	14.638	8.634
2922	ASN368	ND2	−24.823	12.897	10.075
2923	ASN368	C	−20.939	13.914	11.374
2924	ASN368	O	−20.473	12.791	11.595
2925	SER369	N	−20.202	15.006	11.368
2926	SER369	CA	−18.762	14.911	11.598
2927	SER369	CB	−18.067	16.121	10.998
2928	SER369	OG	−16.685	16.01	11.298
2929	SER369	C	−18.445	14.845	13.08
2930	SER369	O	−18.365	15.88	13.756
2931	GLY370	N	−18.241	13.63	13.553
2932	GLY370	CA	−17.831	13.416	14.933
2933	GLY370	C	−16.435	12.82	14.971
2934	GLY370	O	−15.637	13.095	15.878
2935	GLY371	N	−16.171	11.971	13.996
2936	GLY371	CA	−14.829	11.416	13.82
2937	GLY371	C	−14.583	10.114	14.575
2938	GLY371	O	−15.305	9.122	14.417
2939	SER372	N	−13.589	10.172	15.443
2940	SER372	CA	−13.013	8.969	16.06
2941	SER372	CB	−11.585	9.286	16.484
2942	SER372	OG	−11.647	10.206	17.568
2943	SER372	C	−13.764	8.441	17.28
2944	SER372	O	−14.333	9.191	18.084
2945	GLN373	N	−13.508	7.163	17.514
2946	GLN373	CA	−14.041	6.396	18.66
2947	GLN373	CB	−13.476	4.971	18.635
2948	GLN373	CG	−13.783	4.198	17.354
2949	GLN373	CD	−12.562	4.155	16.435
2950	GLN373	OE1	−12.079	5.196	15.969
2951	GLN373	NE2	−12.057	2.954	16.218
2952	GLN373	C	−13.761	6.972	20.065
2953	GLN373	O	−14.7	6.916	20.867
2954	PRO374	N	−12.579	7.496	20.417
2955	PRO374	CA	−12.429	8.073	21.766
2956	PRO374	CB	−10.984	8.451	21.896
2957	PRO374	CG	−10.262	8.163	20.594
2958	PRO374	CD	−11.299	7.54	19.682
2959	PRO374	C	−13.322	9.291	22.044
2960	PRO374	O	−13.596	9.574	23.217
2961	ASN375	N	−13.846	9.947	21.021
2962	ASN375	CA	−14.905	10.918	21.271
2963	ASN375	CB	−14.728	12.173	20.422
2964	ASN375	CG	−15.705	13.266	20.872
2965	ASN375	OD1	−16.823	12.988	21.336
2966	ASN375	ND2	−15.293	14.502	20.666
2967	ASN375	C	−16.222	10.232	20.943
2968	ASN375	O	−16.894	10.585	19.966
2969	ALA376	N	−16.715	9.49	21.92
2970	ALA376	CA	−17.957	8.731	21.749
2971	ALA376	CB	−18.035	7.678	22.847
2972	ALA376	C	−19.216	9.6	21.779
2973	ALA376	O	−20.25	9.179	21.249
2974	GLU377	N	−19.06	10.869	22.128
2975	GLU377	CA	−20.192	11.794	22.151
2976	GLU377	CB	−19.845	12.924	23.109
2977	GLU377	CG	−19.541	12.39	24.502
2978	GLU377	CD	−18.972	13.508	25.368
2979	GLU377	OE1	−19.749	14.165	26.045
2980	GLU377	OE2	−17.756	13.648	25.372
2981	GLU377	C	−20.443	12.378	20.765

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
2982	GLU377	O	−21.499	12.975	20.525
2983	THR378	N	−19.476	12.221	19.875
2984	THR378	CA	−19.649	12.643	18.486
2985	THR378	CB	−18.62	13.72	18.176
2986	THR378	OG1	−17.327	13.139	18.292
2987	THR378	CG2	−18.716	14.912	19.122
2988	THR378	C	−19.472	11.471	17.519
2989	THR378	O	−19.879	11.556	16.353
2990	PHE379	N	−19.032	10.336	18.037
2991	PHE379	CA	−18.774	9.161	17.191
2992	PHE379	CB	−17.961	8.161	18.01
2993	PHE379	CG	−17.648	6.856	17.286
2994	PHE379	CD1	−16.893	6.872	16.121
2995	PHE379	CE1	−16.617	5.686	15.454
2996	PHE379	CZ	−17.095	4.482	15.954
2997	PHE379	CE2	−17.845	4.465	17.123
2998	PHE379	CD2	−18.121	5.652	17.789
2999	PHE379	C	−20.048	8.489	16.667
3000	PHE379	O	−20.016	7.862	15.599
3001	TRP380	N	−21.181	8.795	17.273
3002	TRP380	CA	−22.455	8.274	16.776
3003	TRP380	CB	−23.413	8.105	17.956
3004	TRP380	CG	−23.813	9.403	18.633
3005	TRP380	CD1	−23.122	10.069	19.621
3006	TRP380	NE1	−23.8	11.199	19.934
3007	TRP380	CE2	−24.933	11.306	19.212
3008	TRP380	CZ2	−25.934	12.26	19.169
3009	TRP380	CH2	−27.008	12.095	18.3
3010	TRP380	CZ3	−27.078	10.978	17.474
3011	TRP380	CE3	−26.076	10.017	17.51
3012	TRP380	CD2	−25.002	10.177	18.374
3013	TRP380	C	−23.094	9.18	15.715
3014	TRP380	O	−24.107	8.787	15.126
3015	THR381	N	−22.497	10.329	15.421
3016	THR381	CA	−23.104	11.233	14.436
3017	THR381	CB	−22.758	12.679	14.777
3018	THR381	OG1	−21.404	12.938	14.429
3019	THR381	CG2	−22.969	12.996	16.253
3020	THR381	C	−22.614	10.924	13.025
3021	THR381	O	−23.295	11.269	12.048
3022	ASN382	N	−21.535	10.159	12.938
3023	ASN382	CA	−20.977	9.75	11.644
3024	ASN382	CB	−19.738	8.895	11.892
3025	ASN382	CG	−18.627	9.652	12.617
3026	ASN382	OD1	−18.264	10.787	12.276
3027	ASN382	ND2	−18.142	9.025	13.669
3028	ASN382	C	−21.989	8.929	10.854
3029	ASN382	O	−22.889	8.315	11.443
3030	PRO383	N	−21.923	9.039	9.536
3031	PRO383	CA	−22.692	8.142	8.671
3032	PRO383	CB	−22.322	8.529	7.271
3033	PRO383	CG	−21.305	9.66	7.312
3034	PRO383	CD	−21.056	9.95	8.783
3035	PRO383	C	−22.35	6.684	8.97
3036	PRO383	O	−21.193	6.26	8.865
3037	GLN384	N	−23.352	5.96	9.438
3038	GLN384	CA	−23.168	4.559	9.824
3039	GLN384	CB	−24.024	4.287	11.053
3040	GLN384	CG	−23.614	5.197	12.208
3041	GLN384	CD	−24.538	4.986	13.401
3042	GLN384	OE1	−24.945	3.856	13.693
3043	GLN384	NE2	−24.859	6.075	14.076
3044	GLN384	C	−23.547	3.617	8.688
3045	GLN384	O	−24.568	3.799	8.015
3046	PHE385	N	−22.693	2.635	8.461
3047	PHE385	CA	−22.918	1.678	7.374
3048	PHE385	CB	−21.747	1.763	6.399
3049	PHE385	CG	−21.534	3.159	5.816
3050	PHE385	CD1	−20.403	3.892	6.156
3051	PHE385	CE1	−20.214	5.164	5.631
3052	PHE385	CZ	−21.153	5.703	4.761
3053	PHE385	CE2	−22.28	4.968	4.415
3054	PHE385	CD2	−22.47	3.697	4.941
3055	PHE385	C	−23.062	0.257	7.911

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
3056	PHE385	O	−22.137	−0.299	8.515
3057	ARG386	N	−24.236	−0.31	7.699
3058	ARG386	CA	−24.501	−1.677	8.155
3059	ARG386	CB	−26.005	−1.859	8.303
3060	ARG386	CG	−26.572	−0.891	9.334
3061	ARG386	CD	−28.086	−1.028	9.448
3062	ARG386	NE	−28.622	−0.09	10.447
3063	ARG386	CZ	−29.311	1.009	10.131
3064	ARG386	NH1	−29.755	1.817	11.096
3065	ARG386	NH2	−29.546	1.306	8.851
3066	ARG386	C	−23.949	−2.699	7.167
3067	ARG386	O	−24.447	−2.845	6.045
3068	LEU387	N	−22.908	−3.387	7.601

TABLE IV-continued

Atom No	Residue	Atom Name	X-Coord	Y-Coord	Z-Coord
3069	LEU387	CA	−22.281	−4.424	6.774
3070	LEU387	CB	−20.761	−4.245	6.758
3071	LEU387	CG	−20.232	−3.532	5.509
3072	LEU387	CD1	−20.71	−2.087	5.393
3073	LEU387	CD2	−18.708	−3.572	5.481
3074	LEU387	C	−22.626	−5.808	7.308
3075	LEU387	O	−23.527	−6.479	6.785

[1160]

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 102

<210> SEQ ID NO 1
<211> LENGTH: 2220
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens
<220> FEATURE:
<221> NAME/KEY: CDS
<222> LOCATION: (3)..(2207)

<400> SEQUENCE: 1

```
ag atg gca tcc agc agt ggg agg gtc acc atc cag ctc gtg gat gag        47
   Met Ala Ser Ser Ser Gly Arg Val Thr Ile Gln Leu Val Asp Glu
   1               5                  10                  15

gag gct ggg gtc gga gcc ggg cgc ctg cag ctt ttt cgg ggc cag agc       95
Glu Ala Gly Val Gly Ala Gly Arg Leu Gln Leu Phe Arg Gly Gln Ser
                20                  25                  30

tat gag gca att cgg gca gcc tgc ctg gat tcg ggg atc ctg ttc cgc      143
Tyr Glu Ala Ile Arg Ala Ala Cys Leu Asp Ser Gly Ile Leu Phe Arg
            35                  40                  45

gac cct tac ttc cct gct ggc cct gat gcc ctt ggc tat gac cag ctg      191
Asp Pro Tyr Phe Pro Ala Gly Pro Asp Ala Leu Gly Tyr Asp Gln Leu
        50                  55                  60

ggg ccg gac tcg gag aag gcc aaa ggc gtg aaa tgg atg agg ccc cat      239
Gly Pro Asp Ser Glu Lys Ala Lys Gly Val Lys Trp Met Arg Pro His
    65                  70                  75

gag ttc tgt gct gag ccg aag ttc atc tgt gaa gac atg agc cgc aca      287
Glu Phe Cys Ala Glu Pro Lys Phe Ile Cys Glu Asp Met Ser Arg Thr
80                  85                  90                  95

gac gtg tgt cag ggg agc ctg ggt aac tgc tgg ttc ctt gca gcc gcc      335
Asp Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Phe Leu Ala Ala Ala
                100                 105                 110

gcc tcc ctt act ctg tat ccc cgg ctc ctg cgc cgg gtg gtc cct cct      383
Ala Ser Leu Thr Leu Tyr Pro Arg Leu Leu Arg Arg Val Val Pro Pro
            115                 120                 125

gga cag gat ttc cag cat ggc tac gca ggc gtc ttc cac ttc cag ctc      431
Gly Gln Asp Phe Gln His Gly Tyr Ala Gly Val Phe His Phe Gln Leu
        130                 135                 140

tgg cag ttt ggc cgc tgg atg gac gtc gtg gtg gat gac agg ctg ccc      479
Trp Gln Phe Gly Arg Trp Met Asp Val Val Val Asp Asp Arg Leu Pro
    145                 150                 155
```

-continued

```
gtg cgt gag ggg aag ctg atg ttc gtg cgc tcg gaa cag cgg aat gag      527
Val Arg Glu Gly Lys Leu Met Phe Val Arg Ser Glu Gln Arg Asn Glu
160                 165                 170                 175

ttc tgg gcc cca ctc ctg gag aag gcc tac gcc aag ctc cac ggc tcc      575
Phe Trp Ala Pro Leu Leu Glu Lys Ala Tyr Ala Lys Leu His Gly Ser
                180                 185                 190

tat gag gtg atg cgg ggc ggc cac atg aat gag gct ttt gtg gat ttc      623
Tyr Glu Val Met Arg Gly Gly His Met Asn Glu Ala Phe Val Asp Phe
            195                 200                 205

aca ggc ggc gtg ggc gag gtg ctc tat ctg aga caa aac agc atg ggg      671
Thr Gly Gly Val Gly Glu Val Leu Tyr Leu Arg Gln Asn Ser Met Gly
        210                 215                 220

ctg ttc tct gcc ctg cgc cat gcc ctg gcc aag gag tcc ctc gtg ggc      719
Leu Phe Ser Ala Leu Arg His Ala Leu Ala Lys Glu Ser Leu Val Gly
    225                 230                 235

gcc act gcc ctg agt gat cgg ggt gag tac cgc aca gaa gag ggc ctg      767
Ala Thr Ala Leu Ser Asp Arg Gly Glu Tyr Arg Thr Glu Glu Gly Leu
240                 245                 250                 255

gta aag gga cac gcg tat tcc atc acg ggc aca cac aag gtg ttc ctg      815
Val Lys Gly His Ala Tyr Ser Ile Thr Gly Thr His Lys Val Phe Leu
                260                 265                 270

ggc ttc acc aag gtg cgg ctg ctg cgg ctg cgg aac cca tgg ggc tgc      863
Gly Phe Thr Lys Val Arg Leu Leu Arg Leu Arg Asn Pro Trp Gly Cys
            275                 280                 285

gtg gag tgg acg ggg gcc tgg agc gac agc tgc cca cgc tgg gac aca      911
Val Glu Trp Thr Gly Ala Trp Ser Asp Ser Cys Pro Arg Trp Asp Thr
        290                 295                 300

ctc ccc acc gag tgc cgc gat gcc ctg ctg gtg aaa aag gag gat ggc      959
Leu Pro Thr Glu Cys Arg Asp Ala Leu Leu Val Lys Lys Glu Asp Gly
    305                 310                 315

gag ttc tgg atg gag ctg cgg gac ttc ctc ctc cat ttc gac acc gtg     1007
Glu Phe Trp Met Glu Leu Arg Asp Phe Leu Leu His Phe Asp Thr Val
320                 325                 330                 335

cag atc tgc tcg ctg agc ccg gag gtg ctg ggc ccc agc ccg gag ggg     1055
Gln Ile Cys Ser Leu Ser Pro Glu Val Leu Gly Pro Ser Pro Glu Gly
                340                 345                 350

ggc ggc tgg cac gtc cac acc ttc caa ggc cgc tgg gtg cgt ggc ttc     1103
Gly Gly Trp His Val His Thr Phe Gln Gly Arg Trp Val Arg Gly Phe
            355                 360                 365

aac tcc ggc ggg agc cag cct aat gct gaa acc ttc tgg acc aat cct     1151
Asn Ser Gly Gly Ser Gln Pro Asn Ala Glu Thr Phe Trp Thr Asn Pro
        370                 375                 380

cag ttc cgt tta acg ctg ctg gag cct gat gag gag gat gac gag gat     1199
Gln Phe Arg Leu Thr Leu Leu Glu Pro Asp Glu Glu Asp Asp Glu Asp
    385                 390                 395

gag gaa ggg ccc tgg ggg ggc tgg ggg gct gca ggg gca cgg ggc cca     1247
Glu Glu Gly Pro Trp Gly Gly Trp Gly Ala Ala Gly Ala Arg Gly Pro
400                 405                 410                 415

gcg cgg ggg ggc cgc acg ccc aag tgc acg gtc ctt ctg tcc ctc atc     1295
Ala Arg Gly Gly Arg Thr Pro Lys Cys Thr Val Leu Leu Ser Leu Ile
                420                 425                 430

cag cgc aac cgg cgg cgc ctg aga gcc aag ggc ctc act tac ctc acc     1343
Gln Arg Asn Arg Arg Arg Leu Arg Ala Lys Gly Leu Thr Tyr Leu Thr
            435                 440                 445

gtt ggc ttc cac gtg ttc cag att cca gag gag ctg ctg ggc ctc tgg     1391
Val Gly Phe His Val Phe Gln Ile Pro Glu Glu Leu Leu Gly Leu Trp
        450                 455                 460
```

-continued

```
gat tcc ccg cgc agc cat gcg ctc ctg ccc cgg ctg ctg cgc gcc gac     1439
Asp Ser Pro Arg Ser His Ala Leu Leu Pro Arg Leu Leu Arg Ala Asp
    465                 470                 475

cgc tcg ccc ctc agc gcc cgc cgc gac gtg acc cgc cgc tgc tgc ctg     1487
Arg Ser Pro Leu Ser Ala Arg Arg Asp Val Thr Arg Arg Cys Cys Leu
480                 485                 490                 495

cgt cca ggc cac tac ctg gtg gtg ccg agc acc gcc cac gcc ggc gac     1535
Arg Pro Gly His Tyr Leu Val Val Pro Ser Thr Ala His Ala Gly Asp
                500                 505                 510

gag gct gac ttc act ctg cgt gtc ttc tcc gag cgc cgc cac acg gcc     1583
Glu Ala Asp Phe Thr Leu Arg Val Phe Ser Glu Arg Arg His Thr Ala
            515                 520                 525

gtg gag atc gac gac gtg atc agc gca gac ctg cag tct ctc cag gtg     1631
Val Glu Ile Asp Asp Val Ile Ser Ala Asp Leu Gln Ser Leu Gln Val
        530                 535                 540

ggg act gtt cct gga ggg gcg gca tgg ggc ggg gat ctt ggc cag ggc     1679
Gly Thr Val Pro Gly Gly Ala Ala Trp Gly Gly Asp Leu Gly Gln Gly
    545                 550                 555

ccc tac ctg ccc ctg gag ctg ggg ttg gag cag ctg ttt cag gag ctg     1727
Pro Tyr Leu Pro Leu Glu Leu Gly Leu Glu Gln Leu Phe Gln Glu Leu
560                 565                 570                 575

gct gga gag gag gaa gaa ctc aat gcc tct cag ctc cag gcc tta cta     1775
Ala Gly Glu Glu Glu Glu Leu Asn Ala Ser Gln Leu Gln Ala Leu Leu
                580                 585                 590

agc att gcc ctg gag cct gcc agg gcc cat acc tcc acc ccc aga gag     1823
Ser Ile Ala Leu Glu Pro Ala Arg Ala His Thr Ser Thr Pro Arg Glu
            595                 600                 605

atc ggg ctc agg acc tgt gag cag ctg ctg cag tgt ttc ggg cat ggg     1871
Ile Gly Leu Arg Thr Cys Glu Gln Leu Leu Gln Cys Phe Gly His Gly
        610                 615                 620

caa agc ctg gcc tta cac cac ttc cag cag ctc tgg ggc tac ctc ctg     1919
Gln Ser Leu Ala Leu His His Phe Gln Gln Leu Trp Gly Tyr Leu Leu
    625                 630                 635

gag tgg cag gcc ata ttc aac aag ttc gat gag gac acc tct gga acc     1967
Glu Trp Gln Ala Ile Phe Asn Lys Phe Asp Glu Asp Thr Ser Gly Thr
640                 645                 650                 655

atg aac tcc tac gag ctg agg ctg gca ctg aat gca gca ggc ttc cac     2015
Met Asn Ser Tyr Glu Leu Arg Leu Ala Leu Asn Ala Ala Gly Phe His
                660                 665                 670

ctg aac aac cag ctg acc cag acc ctc acc agc cgc tac cgg gat agc     2063
Leu Asn Asn Gln Leu Thr Gln Thr Leu Thr Ser Arg Tyr Arg Asp Ser
            675                 680                 685

cgt ctg cgt gtg gac ttc gag cgg ttc gtg tcc tgt gtg gcc cac ctc     2111
Arg Leu Arg Val Asp Phe Glu Arg Phe Val Ser Cys Val Ala His Leu
        690                 695                 700

acc tgc atc ttc tgc cac tgc agc cag cac ctg gat ggg ggt gag ggg     2159
Thr Cys Ile Phe Cys His Cys Ser Gln His Leu Asp Gly Gly Glu Gly
    705                 710                 715

gtc atc tgc ctg acc cac aga cag tgg atg gag gtg gcc acc ttc tcc     2207
Val Ile Cys Leu Thr His Arg Gln Trp Met Glu Val Ala Thr Phe Ser
720                 725                 730                 735

taggatctcc gga                                                     2220


<210> SEQ ID NO 2
<211> LENGTH: 735
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens
```

-continued

<400> SEQUENCE: 2

```
Met Ala Ser Ser Ser Gly Arg Val Thr Ile Gln Leu Val Asp Glu Glu
1               5                  10                  15

Ala Gly Val Gly Ala Gly Arg Leu Gln Leu Phe Arg Gly Gln Ser Tyr
            20                  25                  30

Glu Ala Ile Arg Ala Ala Cys Leu Asp Ser Gly Ile Leu Phe Arg Asp
        35                  40                  45

Pro Tyr Phe Pro Ala Gly Pro Asp Ala Leu Gly Tyr Asp Gln Leu Gly
    50                  55                  60

Pro Asp Ser Glu Lys Ala Lys Gly Val Lys Trp Met Arg Pro His Glu
65                  70                  75                  80

Phe Cys Ala Glu Pro Lys Phe Ile Cys Glu Asp Met Ser Arg Thr Asp
                85                  90                  95

Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Phe Leu Ala Ala Ala Ala
            100                 105                 110

Ser Leu Thr Leu Tyr Pro Arg Leu Leu Arg Arg Val Val Pro Pro Gly
        115                 120                 125

Gln Asp Phe Gln His Gly Tyr Ala Gly Val Phe His Phe Gln Leu Trp
    130                 135                 140

Gln Phe Gly Arg Trp Met Asp Val Val Val Asp Asp Arg Leu Pro Val
145                 150                 155                 160

Arg Glu Gly Lys Leu Met Phe Val Arg Ser Glu Gln Arg Asn Glu Phe
                165                 170                 175

Trp Ala Pro Leu Leu Glu Lys Ala Tyr Ala Lys Leu His Gly Ser Tyr
            180                 185                 190

Glu Val Met Arg Gly Gly His Met Asn Glu Ala Phe Val Asp Phe Thr
        195                 200                 205

Gly Gly Val Gly Glu Val Leu Tyr Leu Arg Gln Asn Ser Met Gly Leu
    210                 215                 220

Phe Ser Ala Leu Arg His Ala Leu Ala Lys Glu Ser Leu Val Gly Ala
225                 230                 235                 240

Thr Ala Leu Ser Asp Arg Gly Glu Tyr Arg Thr Glu Glu Gly Leu Val
                245                 250                 255

Lys Gly His Ala Tyr Ser Ile Thr Gly Thr His Lys Val Phe Leu Gly
            260                 265                 270

Phe Thr Lys Val Arg Leu Leu Arg Leu Arg Asn Pro Trp Gly Cys Val
        275                 280                 285

Glu Trp Thr Gly Ala Trp Ser Asp Ser Cys Pro Arg Trp Asp Thr Leu
    290                 295                 300

Pro Thr Glu Cys Arg Asp Ala Leu Leu Val Lys Lys Glu Asp Gly Glu
305                 310                 315                 320

Phe Trp Met Glu Leu Arg Asp Phe Leu Leu His Phe Asp Thr Val Gln
                325                 330                 335

Ile Cys Ser Leu Ser Pro Glu Val Leu Gly Pro Ser Pro Glu Gly Gly
            340                 345                 350

Gly Trp His Val His Thr Phe Gln Gly Arg Trp Val Arg Gly Phe Asn
        355                 360                 365

Ser Gly Gly Ser Gln Pro Asn Ala Glu Thr Phe Trp Thr Asn Pro Gln
    370                 375                 380

Phe Arg Leu Thr Leu Leu Glu Pro Asp Glu Glu Asp Asp Glu Asp Glu
```

-continued

```
385                 390                 395                 400

Glu Gly Pro Trp Gly Gly Trp Gly Ala Ala Gly Ala Arg Gly Pro Ala
                405                 410                 415

Arg Gly Gly Arg Thr Pro Lys Cys Thr Val Leu Leu Ser Leu Ile Gln
            420                 425                 430

Arg Asn Arg Arg Arg Leu Arg Ala Lys Gly Leu Thr Tyr Leu Thr Val
        435                 440                 445

Gly Phe His Val Phe Gln Ile Pro Glu Glu Leu Leu Gly Leu Trp Asp
    450                 455                 460

Ser Pro Arg Ser His Ala Leu Leu Pro Arg Leu Leu Arg Ala Asp Arg
465                 470                 475                 480

Ser Pro Leu Ser Ala Arg Arg Asp Val Thr Arg Arg Cys Cys Leu Arg
                485                 490                 495

Pro Gly His Tyr Leu Val Val Pro Ser Thr Ala His Ala Gly Asp Glu
            500                 505                 510

Ala Asp Phe Thr Leu Arg Val Phe Ser Glu Arg Arg His Thr Ala Val
        515                 520                 525

Glu Ile Asp Asp Val Ile Ser Ala Asp Leu Gln Ser Leu Gln Val Gly
    530                 535                 540

Thr Val Pro Gly Gly Ala Ala Trp Gly Gly Asp Leu Gly Gln Gly Pro
545                 550                 555                 560

Tyr Leu Pro Leu Glu Leu Gly Leu Glu Gln Leu Phe Gln Glu Leu Ala
                565                 570                 575

Gly Glu Glu Glu Glu Leu Asn Ala Ser Gln Leu Gln Ala Leu Leu Ser
            580                 585                 590

Ile Ala Leu Glu Pro Ala Arg Ala His Thr Ser Thr Pro Arg Glu Ile
        595                 600                 605

Gly Leu Arg Thr Cys Glu Gln Leu Leu Gln Cys Phe Gly His Gly Gln
    610                 615                 620

Ser Leu Ala Leu His His Phe Gln Gln Leu Trp Gly Tyr Leu Leu Glu
625                 630                 635                 640

Trp Gln Ala Ile Phe Asn Lys Phe Asp Glu Asp Thr Ser Gly Thr Met
                645                 650                 655

Asn Ser Tyr Glu Leu Arg Leu Ala Leu Asn Ala Ala Gly Phe His Leu
            660                 665                 670

Asn Asn Gln Leu Thr Gln Thr Leu Thr Ser Arg Tyr Arg Asp Ser Arg
        675                 680                 685

Leu Arg Val Asp Phe Glu Arg Phe Val Ser Cys Val Ala His Leu Thr
    690                 695                 700

Cys Ile Phe Cys His Cys Ser Gln His Leu Asp Gly Gly Glu Gly Val
705                 710                 715                 720

Ile Cys Leu Thr His Arg Gln Trp Met Glu Val Ala Thr Phe Ser
                725                 730                 735


<210> SEQ ID NO 3
<211> LENGTH: 714
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

Met Ser Glu Glu Ile Ile Thr Pro Val Tyr Cys Thr Gly Val Ser Ala
1               5                   10                  15
```

-continued

```
Gln Val Gln Lys Gln Arg Ala Arg Glu Leu Gly Leu Gly Arg His Glu
            20                  25                  30

Asn Ala Ile Lys Tyr Leu Gly Gln Asp Tyr Glu Gln Leu Arg Val Arg
        35                  40                  45

Cys Leu Gln Ser Gly Thr Leu Phe Arg Asp Glu Ala Phe Pro Pro Val
    50                  55                  60

Pro Gln Ser Leu Gly Tyr Lys Asp Leu Gly Pro Asn Ser Ser Lys Thr
65                  70                  75                  80

Tyr Gly Ile Lys Trp Lys Arg Pro Thr Glu Leu Leu Ser Asn Pro Gln
                85                  90                  95

Phe Ile Val Asp Gly Ala Thr Arg Thr Asp Ile Cys Gln Gly Ala Leu
            100                 105                 110

Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala Ser Leu Thr Leu Asn Asp
        115                 120                 125

Thr Leu Leu His Arg Val Val Pro His Gly Gln Ser Phe Gln Asn Gly
    130                 135                 140

Tyr Ala Gly Ile Phe His Phe Gln Leu Trp Gln Phe Gly Glu Trp Val
145                 150                 155                 160

Asp Val Val Val Asp Asp Leu Leu Pro Ile Lys Asp Gly Lys Leu Val
                165                 170                 175

Phe Val His Ser Ala Glu Gly Asn Glu Phe Trp Ser Ala Leu Leu Glu
            180                 185                 190

Lys Ala Tyr Ala Lys Val Asn Gly Ser Tyr Glu Ala Leu Ser Gly Gly
        195                 200                 205

Ser Thr Ser Glu Gly Phe Glu Asp Phe Thr Gly Gly Val Thr Glu Trp
    210                 215                 220

Tyr Glu Leu Arg Lys Ala Pro Ser Asp Leu Tyr Gln Ile Ile Leu Lys
225                 230                 235                 240

Ala Leu Glu Arg Gly Ser Leu Leu Gly Cys Ser Ile Asp Ile Ser Ser
                245                 250                 255

Val Leu Asp Met Glu Ala Ile Thr Phe Lys Lys Leu Val Lys Gly His
            260                 265                 270

Ala Tyr Ser Val Thr Gly Ala Lys Gln Val Asn Tyr Arg Gly Gln Val
        275                 280                 285

Val Ser Leu Ile Arg Met Arg Asn Pro Trp Gly Glu Val Glu Trp Thr
    290                 295                 300

Gly Ala Trp Ser Asp Ser Ser Ser Glu Trp Asn Asn Val Asp Pro Tyr
305                 310                 315                 320

Glu Arg Asp Gln Leu Arg Val Lys Met Glu Asp Gly Glu Phe Trp Met
                325                 330                 335

Ser Phe Arg Asp Phe Met Arg Glu Phe Thr Arg Leu Glu Ile Cys Asn
            340                 345                 350

Leu Thr Pro Asp Ala Leu Lys Ser Arg Thr Ile Arg Lys Trp Asn Thr
        355                 360                 365

Thr Leu Tyr Glu Gly Thr Trp Arg Arg Gly Ser Thr Ala Gly Gly Cys
    370                 375                 380

Arg Asn Tyr Pro Ala Thr Phe Trp Val Asn Pro Gln Phe Lys Ile Arg
385                 390                 395                 400

Leu Asp Glu Thr Asp Asp Pro Asp Asp Tyr Gly Asp Arg Glu Ser Gly
                405                 410                 415

Cys Ser Phe Val Leu Ala Leu Met Gln Lys His Arg Arg Arg Glu Arg
```

-continued

```
            420                 425                 430

Arg Phe Gly Arg Asp Met Glu Thr Ile Gly Phe Ala Val Tyr Glu Val
        435                 440                 445

Pro Pro Glu Leu Val Gly Gln Pro Ala Val His Leu Lys Arg Asp Phe
    450                 455                 460

Phe Leu Ala Asn Ala Ser Arg Ala Arg Ser Glu Gln Phe Ile Asn Leu
465                 470                 475                 480

Arg Glu Val Ser Thr Arg Phe Arg Leu Pro Pro Gly Glu Tyr Val Val
                485                 490                 495

Val Pro Ser Thr Phe Glu Pro Asn Lys Glu Gly Asp Phe Val Leu Arg
            500                 505                 510

Phe Phe Ser Glu Lys Ser Ala Gly Thr Val Glu Leu Asp Asp Gln Ile
        515                 520                 525

Gln Ala Asn Leu Pro Asp Glu Gln Val Leu Ser Glu Glu Glu Ile Asp
    530                 535                 540

Glu Asn Phe Lys Ala Leu Phe Arg Gln Leu Ala Gly Glu Asp Met Glu
545                 550                 555                 560

Ile Ser Val Lys Glu Leu Arg Thr Ile Leu Asn Arg Ile Ile Ser Lys
                565                 570                 575

His Lys Asp Leu Arg Thr Lys Gly Phe Ser Leu Glu Ser Cys Arg Ser
            580                 585                 590

Met Val Asn Leu Met Asp Arg Asp Gly Asn Gly Lys Leu Gly Leu Val
        595                 600                 605

Glu Phe Asn Ile Leu Trp Asn Arg Ile Arg Asn Tyr Leu Ser Ile Phe
    610                 615                 620

Arg Lys Phe Asp Leu Asp Lys Ser Gly Ser Met Ser Ala Tyr Glu Met
625                 630                 635                 640

Arg Met Ala Ile Glu Ser Ala Gly Phe Lys Leu Asn Lys Lys Leu Tyr
                645                 650                 655

Glu Leu Ile Ile Thr Arg Tyr Ser Glu Pro Asp Leu Ala Val Asp Phe
            660                 665                 670

Asp Asn Phe Val Cys Cys Leu Val Arg Leu Glu Thr Met Phe Arg Phe
        675                 680                 685

Phe Lys Thr Leu Asp Thr Asp Leu Asp Gly Val Val Thr Phe Asp Leu
    690                 695                 700

Phe Lys Trp Leu Gln Leu Thr Met Phe Ala
705                 710


<210> SEQ ID NO 4
<211> LENGTH: 700
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

Met Ala Gly Ile Ala Ala Lys Leu Ala Lys Asp Arg Glu Ala Ala Glu
1               5                   10                  15

Gly Leu Gly Ser His Glu Arg Ala Ile Lys Tyr Leu Asn Gln Asp Tyr
            20                  25                  30

Glu Ala Leu Arg Asn Glu Cys Leu Glu Ala Gly Thr Leu Phe Gln Asp
        35                  40                  45

Pro Ser Phe Pro Ala Ile Pro Ser Ala Leu Gly Phe Lys Glu Leu Gly
    50                  55                  60
```

-continued

```
Pro Tyr Ser Ser Lys Thr Arg Gly Met Arg Trp Lys Arg Pro Thr Glu
65                  70                  75                  80

Ile Cys Ala Asp Pro Gln Phe Ile Ile Gly Gly Ala Thr Arg Thr Asp
                85                  90                  95

Ile Cys Gln Gly Ala Leu Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala
            100                 105                 110

Ser Leu Thr Leu Asn Glu Glu Ile Leu Ala Arg Val Val Pro Leu Asn
        115                 120                 125

Gln Ser Phe Gln Glu Asn Tyr Ala Gly Ile Phe His Phe Gln Phe Trp
    130                 135                 140

Gln Tyr Gly Glu Trp Val Glu Val Val Val Asp Asp Arg Leu Pro Thr
145                 150                 155                 160

Lys Asp Gly Glu Leu Leu Phe Val His Ser Ala Glu Gly Ser Glu Phe
                165                 170                 175

Trp Ser Ala Leu Leu Glu Lys Ala Tyr Ala Lys Ile Asn Gly Cys Tyr
            180                 185                 190

Glu Ala Leu Ser Gly Gly Ala Thr Thr Glu Gly Phe Glu Asp Phe Thr
        195                 200                 205

Gly Gly Ile Ala Glu Trp Tyr Glu Leu Lys Lys Pro Pro Pro Asn Leu
    210                 215                 220

Phe Lys Ile Ile Gln Lys Ala Leu Gln Lys Gly Ser Leu Leu Gly Cys
225                 230                 235                 240

Ser Ile Asp Ile Thr Ser Ala Ala Asp Ser Glu Ala Ile Thr Phe Gln
                245                 250                 255

Lys Leu Val Lys Gly His Ala Tyr Ser Val Thr Gly Ala Glu Glu Val
            260                 265                 270

Glu Ser Asn Gly Ser Leu Gln Lys Leu Ile Arg Ile Arg Asn Pro Trp
        275                 280                 285

Gly Glu Val Glu Trp Thr Gly Arg Trp Asn Asp Asn Cys Pro Ser Trp
    290                 295                 300

Asn Thr Ile Asp Pro Glu Glu Arg Glu Arg Leu Thr Arg Arg His Glu
305                 310                 315                 320

Asp Gly Glu Phe Trp Met Ser Phe Ser Asp Phe Leu Arg His Tyr Ser
                325                 330                 335

Arg Leu Glu Ile Cys Asn Leu Thr Pro Asp Thr Leu Thr Ser Asp Thr
            340                 345                 350

Tyr Lys Lys Trp Lys Leu Thr Lys Met Asp Gly Asn Trp Arg Arg Gly
        355                 360                 365

Ser Thr Ala Gly Gly Cys Arg Asn Tyr Pro Asn Thr Phe Trp Met Asn
    370                 375                 380

Pro Gln Tyr Leu Ile Lys Leu Glu Glu Glu Asp Glu Asp Glu Glu Asp
385                 390                 395                 400

Gly Glu Ser Gly Cys Thr Phe Leu Val Gly Leu Ile Gln Lys His Arg
                405                 410                 415

Arg Arg Gln Arg Lys Met Gly Glu Asp Met His Thr Ile Gly Phe Gly
            420                 425                 430

Ile Tyr Glu Val Pro Glu Glu Leu Ser Gly Gln Thr Asn Ile His Leu
        435                 440                 445

Ser Lys Asn Phe Phe Leu Thr Asn Arg Ala Arg Glu Arg Ser Asp Thr
    450                 455                 460

Phe Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly
```

-continued

```
465                 470                 475                 480

Glu Tyr Ile Leu Val Pro Ser Thr Phe Glu Pro Asn Lys Asp Gly Asp
                485                 490                 495

Phe Cys Ile Arg Val Phe Ser Glu Lys Lys Ala Asp Tyr Gln Ala Val
            500                 505                 510

Asp Asp Glu Ile Glu Ala Asn Leu Glu Glu Phe Asp Ile Ser Glu Asp
        515                 520                 525

Asp Ile Asp Asp Gly Val Arg Arg Leu Phe Ala Gln Leu Ala Gly Glu
    530                 535                 540

Asp Ala Glu Ile Ser Ala Phe Glu Leu Gln Thr Ile Leu Arg Arg Val
545                 550                 555                 560

Leu Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr
                565                 570                 575

Cys Lys Ile Met Val Asp Met Leu Asp Ser Asp Gly Ser Gly Lys Leu
            580                 585                 590

Gly Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln
        595                 600                 605

Lys Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser
    610                 615                 620

Tyr Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Met Pro Cys
625                 630                 635                 640

Gln Leu His Gln Val Ile Val Ala Arg Phe Ala Asp Asp Gln Leu Ile
                645                 650                 655

Ile Asp Phe Asp Asn Phe Val Arg Cys Leu Val Arg Leu Glu Thr Leu
            660                 665                 670

Phe Lys Ile Phe Lys Gln Leu Asp Pro Glu Asn Thr Gly Thr Ile Glu
        675                 680                 685

Leu Asp Leu Ile Ser Trp Leu Cys Phe Ser Val Leu
    690                 695                 700


<210> SEQ ID NO 5
<211> LENGTH: 821
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

Met Pro Thr Val Ile Ser Ala Ser Val Ala Pro Arg Thr Ala Ala Glu
1               5                   10                  15

Pro Arg Ser Pro Gly Pro Val Pro His Pro Ala Gln Ser Lys Ala Thr
            20                  25                  30

Glu Ala Gly Gly Gly Asn Pro Ser Gly Ile Tyr Ser Ala Ile Ile Ser
        35                  40                  45

Arg Asn Phe Pro Ile Ile Gly Val Lys Glu Lys Thr Phe Glu Gln Leu
    50                  55                  60

His Lys Lys Cys Leu Glu Lys Lys Val Leu Tyr Val Asp Pro Glu Phe
65                  70                  75                  80

Pro Pro Asp Glu Thr Ser Leu Phe Tyr Ser Gln Lys Phe Pro Ile Gln
                85                  90                  95

Phe Val Trp Lys Arg Pro Pro Glu Ile Cys Glu Asn Pro Arg Phe Ile
            100                 105                 110

Ile Asp Gly Ala Asn Arg Thr Asp Ile Cys Gln Gly Glu Leu Gly Asp
        115                 120                 125
```

-continued

```
Cys Trp Phe Leu Ala Ala Ile Ala Cys Leu Thr Leu Asn Gln His Leu
    130                 135                 140

Leu Phe Arg Val Ile Pro His Asp Gln Ser Phe Ile Glu Asn Tyr Ala
145                 150                 155                 160

Gly Ile Phe His Phe Gln Phe Trp Arg Tyr Gly Glu Trp Val Asp Val
                165                 170                 175

Val Ile Asp Asp Cys Leu Pro Thr Tyr Asn Asn Gln Leu Val Phe Thr
            180                 185                 190

Lys Ser Asn His Arg Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala
        195                 200                 205

Tyr Ala Lys Leu His Gly Ser Tyr Glu Ala Leu Lys Gly Gly Asn Thr
    210                 215                 220

Thr Glu Ala Met Glu Asp Phe Thr Gly Gly Val Ala Glu Phe Phe Glu
225                 230                 235                 240

Ile Arg Asp Ala Pro Ser Asp Met Tyr Lys Ile Met Lys Lys Ala Ile
                245                 250                 255

Glu Arg Gly Ser Leu Met Gly Cys Ser Ile Asp Asp Gly Thr Asn Met
            260                 265                 270

Thr Tyr Gly Thr Ser Pro Ser Gly Leu Asn Met Gly Glu Leu Ile Ala
        275                 280                 285

Arg Met Val Arg Asn Met Asp Asn Ser Leu Leu Gln Asp Ser Asp Leu
    290                 295                 300

Asp Pro Arg Gly Ser Asp Glu Arg Pro Thr Arg Thr Ile Ile Pro Val
305                 310                 315                 320

Gln Tyr Glu Thr Arg Met Ala Cys Gly Leu Val Arg Gly His Ala Tyr
                325                 330                 335

Ser Val Thr Gly Leu Asp Glu Val Pro Phe Lys Gly Glu Lys Val Lys
            340                 345                 350

Leu Val Arg Leu Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser
        355                 360                 365

Trp Ser Asp Arg Trp Lys Asp Trp Ser Phe Val Asp Lys Asp Glu Lys
    370                 375                 380

Ala Arg Leu Gln His Gln Val Thr Glu Asp Gly Glu Phe Trp Met Ser
385                 390                 395                 400

Tyr Glu Asp Phe Ile Tyr His Phe Thr Lys Leu Glu Ile Cys Asn Leu
                405                 410                 415

Thr Ala Asp Ala Leu Gln Ser Asp Lys Leu Gln Thr Trp Thr Val Ser
            420                 425                 430

Val Asn Glu Gly Arg Trp Val Arg Gly Cys Ser Ala Gly Gly Cys Arg
        435                 440                 445

Asn Phe Pro Asp Thr Phe Trp Thr Asn Pro Gln Tyr Arg Leu Lys Leu
    450                 455                 460

Leu Glu Glu Asp Asp Asp Pro Asp Asp Ser Glu Val Ile Cys Ser Phe
465                 470                 475                 480

Leu Val Ala Leu Met Gln Lys Asn Arg Arg Lys Asp Arg Lys Leu Gly
                485                 490                 495

Ala Ser Leu Phe Thr Ile Gly Phe Ala Ile Tyr Glu Val Pro Lys Glu
            500                 505                 510

Met His Gly Asn Lys Gln His Leu Gln Lys Asp Phe Phe Leu Tyr Asn
        515                 520                 525

Ala Ser Lys Ala Arg Ser Lys Thr Tyr Ile Asn Met Arg Glu Val Ser
```

-continued

```
    530                 535                 540

Gln Arg Phe Arg Leu Pro Pro Ser Glu Tyr Val Ile Val Pro Ser Thr
545                 550                 555                 560

Tyr Glu Pro His Gln Glu Gly Glu Phe Ile Leu Arg Val Phe Ser Glu
                565                 570                 575

Lys Arg Asn Leu Ser Glu Glu Val Glu Asn Thr Ile Ser Val Asp Arg
            580                 585                 590

Pro Val Lys Lys Lys Lys Thr Lys Pro Ile Ile Phe Val Ser Asp Arg
        595                 600                 605

Ala Asn Ser Asn Lys Glu Leu Gly Val Asp Gln Glu Ser Glu Glu Gly
    610                 615                 620

Lys Gly Lys Thr Ser Pro Asp Lys Gln Lys Gln Ser Pro Gln Pro Gln
625                 630                 635                 640

Pro Gly Ser Ser Asp Gln Glu Ser Glu Glu Gln Gln Gln Phe Arg Asn
                645                 650                 655

Ile Phe Lys Gln Ile Ala Gly Asp Asp Met Glu Ile Cys Ala Asp Glu
            660                 665                 670

Leu Lys Lys Val Leu Asn Thr Val Val Asn Lys His Lys Asp Leu Lys
        675                 680                 685

Thr His Gly Phe Thr Leu Glu Ser Cys Arg Ser Met Ile Ala Leu Met
    690                 695                 700

Asp Thr Asp Gly Ser Gly Lys Leu Asn Leu Gln Glu Phe His His Leu
705                 710                 715                 720

Trp Asn Lys Ile Lys Ala Trp Gln Lys Ile Phe Lys His Tyr Asp Thr
                725                 730                 735

Asp Gln Ser Gly Thr Ile Asn Ser Tyr Glu Met Arg Asn Ala Val Asn
            740                 745                 750

Asp Ala Gly Phe His Leu Asn Asn Gln Leu Tyr Asp Ile Ile Thr Met
        755                 760                 765

Arg Tyr Ala Asp Lys His Met Asn Ile Asp Phe Asp Ser Phe Ile Cys
    770                 775                 780

Cys Phe Val Arg Leu Glu Gly Met Phe Arg Ala Phe His Ala Phe Asp
785                 790                 795                 800

Lys Asp Gly Asp Gly Ile Ile Lys Leu Asn Val Leu Glu Trp Leu Gln
                805                 810                 815

Leu Thr Met Tyr Ala
            820


<210> SEQ ID NO 6
<211> LENGTH: 639
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

Met Phe Ser Cys Val Lys Pro Tyr Glu Asp Gln Asn Tyr Ser Ala Leu
1               5                   10                  15

Arg Gln Asp Cys Arg Arg Arg Lys Val Leu Phe Glu Asp Pro Leu Phe
            20                  25                  30

Pro Ala Thr Asp Asp Ser Leu Tyr Tyr Lys Gly Thr Pro Gly Pro Ala
        35                  40                  45

Val Arg Trp Lys Arg Pro Lys Gly Ile Cys Glu Asp Pro Arg Leu Phe
    50                  55                  60
```

-continued

```
Val Asp Gly Ile Ser Ser His Asp Leu His Gln Gly Gln Val Gly Asn
65                  70                  75                  80

Cys Trp Phe Val Ala Ala Cys Ser Ser Leu Ala Ser Arg Glu Ser Leu
                85                  90                  95

Trp Gln Lys Val Ile Pro Asp Trp Lys Glu Gln Glu Trp Asp Pro Arg
            100                 105                 110

Lys Ala Gln Ala Tyr Ala Gly Ile Phe His Phe His Phe Trp Arg Leu
        115                 120                 125

Gly Met Val Asp Val Val Ile Asp Glu Arg Leu Pro Thr Val Asn Asn
    130                 135                 140

Gln Leu Ile Tyr Cys His Ser Asn Ser Arg Asn Glu Phe Trp Cys Ala
145                 150                 155                 160

Leu Val Glu Lys Ala Tyr Ala Lys Leu Ala Gly Cys Tyr Gln Ala Leu
                165                 170                 175

Asp Gly Gly Asn Thr Ala Asp Ala Leu Val Asp Phe Thr Gly Gly Val
            180                 185                 190

Ser Glu Pro Ile Asp Leu Thr Glu Gly Asp Phe Ala Asn Asp Glu Thr
        195                 200                 205

Lys Arg Asn Gln Leu Phe Glu Arg Met Leu Lys Val His Ser Arg Gly
    210                 215                 220

Gly Leu Ile Ser Ala Ser Ile Lys Ala Val Thr Ala Ala Asp Met Glu
225                 230                 235                 240

Ala Arg Leu Ala Cys Gly Leu Val Lys Gly His Ala Tyr Ala Val Thr
                245                 250                 255

Asp Val Arg Lys Val Arg Leu Gly His Gly Leu Leu Ala Phe Phe Lys
            260                 265                 270

Ser Glu Lys Leu Asp Met Ile Arg Leu Arg Asn Pro Trp Gly Glu Arg
        275                 280                 285

Glu Trp Asn Gly Pro Trp Ser Asp Thr Ser Glu Glu Trp Gln Lys Val
    290                 295                 300

Ser Lys Ser Glu Arg Glu Lys Met Gly Val Thr Val Gln Asp Asp Gly
305                 310                 315                 320

Glu Phe Trp Met Thr Phe Glu Asp Val Cys Arg Tyr Phe Thr Asp Ile
                325                 330                 335

Ile Lys Cys Arg Val Ile Asn Thr Ser His Leu Ser Ile His Lys Thr
            340                 345                 350

Trp Glu Glu Ala Arg Leu His Gly Ala Trp Thr Leu His Glu Asp Pro
        355                 360                 365

Arg Gln Asn Arg Gly Gly Gly Cys Ile Asn His Lys Asp Thr Phe Phe
    370                 375                 380

Gln Asn Pro Gln Tyr Ile Phe Glu Val Lys Lys Pro Glu Asp Glu Val
385                 390                 395                 400

Leu Ile Cys Ile Gln Gln Arg Pro Lys Arg Ser Thr Arg Arg Glu Gly
                405                 410                 415

Lys Gly Glu Asn Leu Ala Ile Gly Phe Asp Ile Tyr Lys Val Glu Glu
            420                 425                 430

Asn Arg Gln Tyr Arg Met His Ser Leu Gln His Lys Ala Ala Ser Ser
        435                 440                 445

Ile Tyr Ile Asn Ser Arg Ser Val Phe Leu Arg Thr Asp Gln Pro Glu
    450                 455                 460

Gly Arg Tyr Val Ile Ile Pro Thr Thr Phe Glu Pro Gly His Thr Gly
```

-continued

```
465                 470                 475                 480

Glu Phe Leu Leu Arg Val Phe Thr Asp Val Pro Ser Asn Cys Arg Glu
                485                 490                 495

Leu Arg Leu Asp Lys Pro Pro His Thr Cys Trp Ser Ser Leu Cys Gly
            500                 505                 510

Tyr Pro Gln Leu Val Thr Gln Val His Val Leu Gly Ala Ala Gly Leu
        515                 520                 525

Lys Asp Ser Pro Thr Gly Ala Asn Ser Tyr Val Ile Ile Lys Cys Glu
    530                 535                 540

Gly Asp Lys Val Arg Ser Ala Val Gln Lys Gly Thr Ser Thr Pro Glu
545                 550                 555                 560

Tyr Asn Val Lys Gly Ile Phe Tyr Arg Lys Lys Leu Ser Gln Pro Ile
                565                 570                 575

Thr Val Gln Val Trp Asn His Arg Val Leu Lys Asp Glu Phe Leu Gly
            580                 585                 590

Gln Val His Leu Lys Ala Asp Pro Asp Asn Leu Gln Ala Leu His Thr
        595                 600                 605

Leu His Leu Arg Asp Arg Asn Ser Arg Gln Pro Ser Asn Leu Pro Gly
    610                 615                 620

Thr Val Ala Val His Ile Leu Ser Ser Thr Ser Leu Met Ala Val
625                 630                 635


<210> SEQ ID NO 7
<211> LENGTH: 690
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

Met Pro Tyr Leu Tyr Arg Ala Pro Gly Pro Gln Ala His Pro Val Pro
1               5                   10                  15

Lys Asp Ala Arg Ile Thr His Ser Ser Gly Gln Ser Phe Glu Gln Met
            20                  25                  30

Arg Gln Glu Cys Leu Gln Arg Gly Thr Leu Phe Glu Asp Ala Asp Phe
        35                  40                  45

Pro Ala Ser Asn Ser Ser Leu Phe Tyr Ser Glu Arg Pro Gln Ile Pro
    50                  55                  60

Phe Val Trp Lys Arg Pro Gly Glu Ile Val Lys Asn Pro Glu Phe Ile
65                  70                  75                  80

Leu Gly Gly Ala Thr Arg Thr Asp Ile Cys Gln Gly Glu Leu Gly Asp
                85                  90                  95

Cys Trp Leu Leu Ala Ala Ile Ala Ser Leu Thr Leu Asn Gln Lys Ala
            100                 105                 110

Leu Ala Arg Val Ile Pro Gln Asp Gln Ser Phe Gly Pro Gly Tyr Ala
        115                 120                 125

Gly Ile Phe His Phe Gln Phe Trp Gln His Ser Glu Trp Leu Asp Val
    130                 135                 140

Val Ile Asp Asp Arg Leu Pro Thr Phe Arg Asp Arg Leu Val Phe Leu
145                 150                 155                 160

His Ser Ala Asp His Asn Glu Phe Trp Ser Ala Leu Leu Glu Lys Ala
                165                 170                 175

Tyr Ala Lys Leu Asn Gly Ser Tyr Glu Ala Leu Lys Gly Gly Ser Ala
            180                 185                 190
```

-continued

```
Ile Glu Ala Met Glu Asp Phe Thr Gly Gly Val Ala Glu Thr Phe Gln
        195                 200                 205

Thr Lys Glu Ala Pro Glu Asn Phe Tyr Glu Ile Leu Glu Lys Ala Leu
    210                 215                 220

Lys Arg Gly Ser Leu Leu Gly Cys Phe Ile Asp Thr Arg Ser Ala Ala
225                 230                 235                 240

Glu Ser Glu Ala Arg Thr Pro Phe Gly Leu Ile Lys Gly His Ala Tyr
                245                 250                 255

Ser Val Thr Gly Ile Asp Gln Val Ser Phe Arg Gly Gln Arg Ile Glu
            260                 265                 270

Leu Ile Arg Ile Arg Asn Pro Trp Gly Gln Val Glu Trp Asn Gly Ser
        275                 280                 285

Trp Ser Asp Ser Ser Pro Glu Trp Arg Ser Val Gly Pro Ala Glu Gln
    290                 295                 300

Lys Arg Leu Cys His Thr Ala Leu Asp Asp Gly Glu Phe Trp Met Ala
305                 310                 315                 320

Phe Lys Asp Phe Lys Ala His Phe Asp Lys Val Glu Ile Cys Asn Leu
                325                 330                 335

Thr Pro Asp Ala Leu Glu Glu Asp Ala Ile His Lys Trp Glu Val Thr
            340                 345                 350

Val His Gln Gly Ser Trp Val Arg Gly Ser Thr Ala Gly Gly Cys Arg
        355                 360                 365

Asn Phe Leu Asp Thr Phe Trp Thr Asn Pro Gln Ile Lys Leu Ser Leu
    370                 375                 380

Thr Glu Lys Asp Glu Gly Gln Glu Glu Cys Ser Phe Leu Val Ala Leu
385                 390                 395                 400

Met Gln Lys Asp Arg Arg Lys Leu Lys Arg Phe Gly Ala Asn Val Leu
                405                 410                 415

Thr Ile Gly Tyr Ala Ile Tyr Glu Cys Pro Asp Lys Asp Glu His Leu
            420                 425                 430

Asn Lys Asp Phe Phe Arg Tyr His Ala Ser Arg Ala Arg Ser Lys Thr
        435                 440                 445

Phe Ile Asn Leu Arg Glu Val Ser Asp Arg Phe Lys Leu Pro Pro Gly
    450                 455                 460

Glu Tyr Ile Leu Ile Pro Ser Thr Phe Glu Pro His Gln Glu Ala Asp
465                 470                 475                 480

Phe Cys Leu Arg Ile Phe Ser Glu Lys Lys Ala Ile Thr Arg Asp Met
                485                 490                 495

Asp Gly Asn Val Asp Ile Asp Leu Pro Glu Pro Pro Lys Pro Thr Pro
            500                 505                 510

Pro Asp Gln Glu Thr Glu Glu Glu Gln Arg Phe Arg Ala Leu Phe Glu
        515                 520                 525

Gln Val Ala Gly Glu Asp Met Glu Val Thr Ala Glu Glu Leu Glu Tyr
    530                 535                 540

Val Leu Asn Ala Val Leu Gln Lys Lys Lys Asp Ile Lys Phe Lys Lys
545                 550                 555                 560

Leu Ser Leu Ile Ser Cys Lys Asn Ile Ile Ser Leu Met Asp Thr Ser
                565                 570                 575

Gly Asn Gly Lys Leu Glu Phe Asp Glu Phe Lys Val Phe Trp Asp Lys
            580                 585                 590

Leu Lys Gln Trp Ile Asn Leu Phe Leu Arg Phe Asp Ala Asp Lys Ser
```

-continued

```
        595                 600                 605

Gly Thr Met Ser Thr Tyr Glu Leu Arg Thr Ala Leu Lys Ala Ala Gly
    610                 615                 620

Phe Gln Leu Ser Ser His Leu Leu Gln Leu Ile Val Leu Arg Tyr Ala
625                 630                 635                 640

Asp Glu Glu Leu Gln Leu Asp Phe Asp Asp Phe Leu Asn Cys Leu Val
                645                 650                 655

Arg Leu Glu Asn Ala Ser Arg Val Phe Gln Ala Leu Ser Thr Lys Asn
            660                 665                 670

Lys Glu Phe Ile His Leu Asn Ile Asn Glu Phe Ile His Leu Thr Met
        675                 680                 685

Asn Ile
    690


<210> SEQ ID NO 8
<211> LENGTH: 671
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met Arg Ala Gly Arg Gly Ala Thr Pro Ala Arg Glu Leu Phe Arg Asp
1               5                   10                  15

Ala Ala Phe Pro Ala Ala Asp Ser Ser Leu Phe Cys Asp Leu Ser Thr
            20                  25                  30

Pro Leu Ala Gln Phe Arg Glu Asp Ile Thr Trp Arg Arg Pro Gln Glu
        35                  40                  45

Ile Cys Ala Thr Pro Arg Leu Phe Pro Asp Asp Pro Arg Glu Gly Gln
    50                  55                  60

Val Lys Gln Gly Leu Leu Gly Asp Cys Trp Phe Leu Cys Ala Cys Ala
65                  70                  75                  80

Ala Leu Gln Lys Ser Arg His Leu Leu Asp Gln Val Ile Pro Pro Gly
                85                  90                  95

Gln Pro Ser Trp Ala Asp Gln Glu Tyr Arg Gly Ser Phe Thr Cys Arg
            100                 105                 110

Ile Trp Gln Phe Gly Arg Trp Val Glu Val Thr Thr Asp Asp Arg Leu
        115                 120                 125

Pro Cys Leu Ala Gly Arg Leu Cys Phe Ser Arg Cys Gln Arg Glu Asp
    130                 135                 140

Val Phe Trp Leu Pro Leu Leu Glu Lys Val Tyr Ala Lys Val His Gly
145                 150                 155                 160

Ser Tyr Glu His Leu Trp Ala Gly Gln Val Ala Asp Ala Leu Val Asp
                165                 170                 175

Leu Thr Gly Gly Leu Ala Glu Arg Trp Asn Leu Lys Gly Val Ala Gly
            180                 185                 190

Ser Gly Gly Gln Gln Asp Arg Pro Gly Arg Trp Glu His Arg Thr Cys
        195                 200                 205

Arg Gln Leu Leu His Leu Lys Asp Gln Cys Leu Ile Ser Cys Cys Val
    210                 215                 220

Leu Ser Pro Arg Ala Gly Ala Arg Glu Leu Gly Glu Phe His Ala Phe
225                 230                 235                 240

Ile Val Ser Asp Leu Arg Glu Leu Gln Gly Gln Ala Gly Gln Cys Ile
                245                 250                 255
```

-continued

```
Leu Leu Leu Arg Ile Gln Asn Pro Trp Gly Arg Arg Cys Trp Gln Gly
            260                 265                 270

Leu Trp Arg Glu Gly Gly Glu Gly Trp Ser Gln Val Asp Ala Ala Val
        275                 280                 285

Ala Ser Glu Leu Leu Ser Gln Leu Gln Glu Gly Glu Phe Trp Val Glu
    290                 295                 300

Glu Glu Glu Phe Leu Arg Glu Phe Asp Glu Leu Thr Val Gly Tyr Pro
305                 310                 315                 320

Val Thr Glu Ala Gly His Leu Gln Ser Leu Tyr Thr Glu Arg Leu Leu
                325                 330                 335

Cys His Thr Arg Ala Leu Pro Gly Ala Trp Val Lys Gly Gln Ser Ala
            340                 345                 350

Gly Gly Cys Arg Asn Asn Ser Gly Phe Pro Ser Asn Pro Lys Phe Trp
        355                 360                 365

Leu Arg Val Ser Glu Pro Ser Glu Val Tyr Ile Ala Val Leu Gln Arg
    370                 375                 380

Ser Arg Leu His Ala Ala Asp Trp Ala Gly Arg Ala Arg Ala Leu Val
385                 390                 395                 400

Gly Asp Ser His Thr Ser Trp Ser Pro Ala Ser Ile Pro Gly Lys His
                405                 410                 415

Tyr Gln Ala Val Gly Leu His Leu Trp Lys Val Glu Lys Arg Arg Val
            420                 425                 430

Asn Leu Pro Arg Val Leu Ser Met Pro Pro Val Ala Gly Thr Ala Cys
        435                 440                 445

His Ala Tyr Asp Arg Glu Val His Leu Arg Cys Glu Leu Ser Pro Gly
    450                 455                 460

Tyr Tyr Leu Ala Val Pro Ser Thr Phe Leu Lys Asp Ala Pro Gly Glu
465                 470                 475                 480

Phe Leu Leu Arg Val Phe Ser Thr Gly Arg Val Ser Leu Ser Ala Ile
                485                 490                 495

Arg Ala Val Ala Lys Asn Thr Thr Pro Gly Ala Ala Leu Pro Ala Gly
            500                 505                 510

Glu Trp Gly Thr Val Gln Leu Arg Gly Ser Trp Arg Val Gly Gln Thr
        515                 520                 525

Ala Gly Gly Ser Arg Asn Phe Ala Ser Tyr Pro Thr Asn Pro Cys Phe
    530                 535                 540

Pro Phe Ser Val Pro Glu Gly Pro Gly Pro Arg Cys Val Arg Ile Thr
545                 550                 555                 560

Leu His Gln His Cys Arg Pro Ser Asp Thr Glu Phe His Pro Ile Gly
                565                 570                 575

Phe His Ile Phe Gln Val Pro Glu Gly Gly Arg Ser Gln Asp Ala Pro
            580                 585                 590

Pro Leu Leu Leu Gln Glu Phe Leu Leu Ser Cys Val Pro His Arg Tyr
        595                 600                 605

Ala Gln Glu Val Ser Arg Leu Cys Leu Leu Pro Ala Gly Thr Tyr Lys
    610                 615                 620

Val Val Pro Ser Thr Tyr Leu Pro Asp Thr Glu Gly Ala Phe Thr Val
625                 630                 635                 640

Thr Ile Ala Thr Arg Ile Asp Pro Ser Ile His Ser Gln Glu Met Leu
                645                 650                 655

Gly Gln Phe Leu Gln Glu Val Ser Val Met Ala Val Met Lys Thr
```

-continued

660 665 670

<210> SEQ ID NO 9
<211> LENGTH: 702
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

Met Val Ala His Ile Asn Asn Ser Arg Leu Lys Ala Lys Gly Val Gly
1 5 10 15

Gln His Asp Asn Ala Gln Asn Phe Gly Asn Gln Ser Phe Glu Glu Leu
20 25 30

Arg Ala Ala Cys Leu Arg Lys Gly Glu Leu Phe Glu Asp Pro Leu Phe
35 40 45

Pro Ala Glu Pro Ser Ser Leu Gly Phe Lys Asp Leu Gly Pro Asn Ser
50 55 60

Lys Asn Val Gln Asn Ile Ser Trp Gln Arg Pro Lys Asp Ile Ile Asn
65 70 75 80

Asn Pro Leu Phe Ile Met Asp Gly Ile Ser Pro Thr Asp Ile Cys Gln
85 90 95

Gly Ile Leu Gly Asp Cys Trp Leu Leu Ala Ala Ile Gly Ser Leu Thr
100 105 110

Thr Cys Pro Lys Leu Leu Tyr Arg Val Val Pro Arg Gly Gln Ser Phe
115 120 125

Lys Lys Asn Tyr Ala Gly Ile Phe His Phe Gln Ile Trp Gln Phe Gly
130 135 140

Gln Trp Val Asn Val Val Val Asp Asp Arg Leu Pro Thr Lys Asn Asp
145 150 155 160

Lys Leu Val Phe Val His Ser Thr Glu Arg Ser Glu Phe Trp Ser Ala
165 170 175

Leu Leu Glu Lys Ala Tyr Ala Lys Leu Ser Gly Ser Tyr Glu Ala Leu
180 185 190

Ser Gly Gly Ser Thr Met Glu Gly Leu Glu Asp Phe Thr Gly Gly Val
195 200 205

Ala Gln Ser Phe Gln Leu Gln Arg Pro Pro Gln Asn Leu Leu Arg Leu
210 215 220

Leu Arg Lys Ala Val Glu Arg Ser Ser Leu Met Gly Cys Ser Ile Glu
225 230 235 240

Val Thr Ser Asp Ser Glu Leu Glu Ser Met Thr Asp Lys Met Leu Val
245 250 255

Arg Gly His Ala Tyr Ser Val Thr Gly Leu Gln Asp Val His Tyr Arg
260 265 270

Gly Lys Met Glu Thr Leu Ile Arg Val Arg Asn Pro Trp Gly Arg Ile
275 280 285

Glu Trp Asn Gly Ala Trp Ser Asp Ser Ala Arg Glu Trp Glu Glu Val
290 295 300

Ala Ser Asp Ile Gln Met Gln Leu Leu His Lys Thr Glu Asp Gly Glu
305 310 315 320

Phe Trp Met Ser Tyr Gln Asp Phe Leu Asn Asn Phe Thr Leu Leu Glu
325 330 335

Ile Cys Asn Leu Thr Pro Asp Thr Leu Ser Gly Asp Tyr Lys Ser Tyr
340 345 350

-continued

```
Trp His Thr Thr Phe Tyr Glu Gly Ser Trp Arg Arg Gly Ser Ser Ala
        355                 360                 365

Gly Gly Cys Arg Asn His Pro Gly Thr Phe Trp Thr Asn Pro Gln Phe
    370                 375                 380

Lys Ile Ser Leu Pro Glu Gly Asp Asp Pro Glu Asp Asp Ala Glu Gly
385                 390                 395                 400

Asn Val Val Val Cys Thr Cys Leu Val Ala Leu Met Gln Lys Asn Trp
                405                 410                 415

Arg His Ala Arg Gln Gln Gly Ala Gln Leu Gln Thr Ile Gly Phe Val
            420                 425                 430

Leu Tyr Ala Val Pro Lys Glu Phe Gln Asn Ile Gln Asp Val His Leu
        435                 440                 445

Lys Lys Glu Phe Phe Thr Lys Tyr Gln Asp His Gly Phe Ser Glu Ile
    450                 455                 460

Phe Thr Asn Ser Arg Glu Val Ser Ser Gln Leu Arg Leu Pro Pro Gly
465                 470                 475                 480

Glu Tyr Ile Ile Ile Pro Ser Thr Phe Glu Pro His Arg Asp Ala Asp
                485                 490                 495

Phe Leu Leu Arg Val Phe Thr Glu Lys His Ser Glu Ser Trp Glu Leu
            500                 505                 510

Asp Glu Val Asn Tyr Ala Glu Gln Leu Gln Glu Glu Lys Val Ser Glu
        515                 520                 525

Asp Asp Met Asp Gln Asp Phe Leu His Leu Phe Lys Ile Val Ala Gly
    530                 535                 540

Glu Gly Lys Glu Ile Gly Val Tyr Glu Leu Gln Arg Leu Leu Asn Arg
545                 550                 555                 560

Met Ala Ile Lys Phe Lys Ser Phe Lys Thr Lys Gly Phe Gly Leu Asp
                565                 570                 575

Ala Cys Arg Cys Met Ile Asn Leu Met Asp Lys Asp Gly Ser Gly Lys
            580                 585                 590

Leu Gly Leu Leu Glu Phe Lys Ile Leu Trp Lys Lys Leu Lys Lys Trp
        595                 600                 605

Met Asp Ile Phe Arg Glu Cys Asp Gln Asp His Ser Gly Thr Leu Asn
    610                 615                 620

Ser Tyr Glu Met Arg Leu Val Ile Glu Lys Ala Gly Ile Lys Leu Asn
625                 630                 635                 640

Asn Lys Val Met Gln Val Leu Val Ala Arg Tyr Ala Asp Asp Asp Leu
                645                 650                 655

Ile Ile Asp Phe Asp Ser Phe Ile Ser Cys Phe Leu Arg Leu Lys Thr
            660                 665                 670

Met Phe Thr Phe Phe Leu Thr Met Asp Pro Lys Asn Thr Gly His Ile
        675                 680                 685

Cys Leu Ser Leu Glu Gln Trp Leu Gln Met Thr Met Trp Gly
    690                 695                 700


<210> SEQ ID NO 10
<211> LENGTH: 694
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

Met Ser Leu Trp Pro Pro Phe Arg Cys Arg Trp Lys Leu Ala Pro Arg
1               5                   10                  15
```

-continued

```
Tyr Ser Arg Arg Ala Ser Pro Gln Gln Pro Gln Gln Asp Phe Glu Ala
            20                  25                  30

Leu Leu Ala Glu Cys Leu Arg Asn Gly Cys Leu Phe Glu Asp Thr Ser
        35                  40                  45

Phe Pro Ala Thr Leu Ser Ser Ile Gly Ser Gly Ser Leu Leu Gln Lys
    50                  55                  60

Leu Pro Pro Arg Leu Gln Trp Lys Arg Pro Pro Glu Leu His Ser Asn
65                  70                  75                  80

Pro Gln Phe Tyr Phe Ala Lys Ala Lys Arg Leu Asp Leu Cys Gln Gly
                85                  90                  95

Ile Val Gly Asp Cys Trp Phe Leu Ala Ala Leu Gln Ala Leu Ala Leu
            100                 105                 110

His Gln Asp Ile Leu Ser Arg Val Val Pro Leu Asn Gln Ser Phe Thr
        115                 120                 125

Glu Lys Tyr Ala Gly Ile Phe Arg Phe Trp Phe Trp His Tyr Gly Asn
    130                 135                 140

Trp Val Pro Val Val Ile Asp Asp Arg Leu Pro Val Asn Glu Ala Gly
145                 150                 155                 160

Gln Leu Val Phe Val Ser Ser Thr Tyr Lys Asn Leu Phe Trp Gly Ala
                165                 170                 175

Leu Leu Glu Lys Ala Tyr Ala Lys Leu Ser Gly Ser Tyr Glu Asp Leu
            180                 185                 190

Gln Ser Gly Gln Val Ser Glu Ala Leu Val Asp Phe Thr Gly Gly Val
        195                 200                 205

Thr Met Thr Ile Asn Leu Ala Glu Ala His Gly Asn Leu Trp Asp Ile
    210                 215                 220

Leu Ile Glu Ala Thr Tyr Asn Arg Thr Leu Ile Gly Cys Gln Thr His
225                 230                 235                 240

Ser Gly Glu Lys Ile Leu Glu Asn Gly Leu Val Glu Gly His Ala Tyr
                245                 250                 255

Thr Leu Thr Gly Ile Arg Lys Val Thr Cys Lys His Arg Pro Glu Tyr
            260                 265                 270

Leu Val Lys Leu Arg Asn Pro Trp Gly Lys Val Glu Trp Lys Gly Asp
        275                 280                 285

Trp Ser Asp Ser Ser Ser Lys Trp Glu Leu Leu Ser Pro Lys Glu Lys
    290                 295                 300

Ile Leu Leu Leu Arg Lys Asp Asn Asp Gly Glu Phe Trp Met Thr Leu
305                 310                 315                 320

Gln Asp Phe Lys Thr His Phe Val Leu Leu Val Ile Cys Lys Leu Thr
                325                 330                 335

Pro Gly Leu Leu Ser Gln Glu Ala Ala Gln Lys Trp Thr Tyr Thr Met
            340                 345                 350

Arg Glu Gly Arg Trp Glu Lys Arg Ser Thr Ala Gly Gly Gln Arg Gln
        355                 360                 365

Leu Leu Gln Asp Thr Phe Trp Lys Asn Pro Gln Phe Leu Leu Ser Val
    370                 375                 380

Trp Arg Pro Glu Glu Gly Arg Arg Ser Leu Arg Pro Cys Ser Val Leu
385                 390                 395                 400

Val Ser Leu Leu Gln Lys Pro Arg His Arg Cys Arg Lys Arg Lys Pro
                405                 410                 415
```

-continued

```
Leu Leu Ala Ile Gly Phe Tyr Leu Tyr Arg Tyr His Asp Asp Gln Arg
            420                 425                 430

Arg Leu Pro Pro Glu Phe Phe Gln Arg Asn Thr Pro Leu Ser Gln Pro
        435                 440                 445

Asp Arg Phe Leu Lys Glu Lys Glu Val Ser Gln Glu Leu Cys Leu Glu
    450                 455                 460

Pro Gly Thr Tyr Leu Ile Val Pro Cys Ile Leu Glu Ala His Gln Lys
465                 470                 475                 480

Ser Glu Phe Val Leu Arg Val Phe Ser Arg Lys His Ile Phe Tyr Glu
                485                 490                 495

Ile Gly Ser Asn Ser Gly Val Val Phe Ser Lys Glu Ile Glu Asp Gln
            500                 505                 510

Asn Glu Arg Gln Asp Glu Phe Phe Thr Lys Phe Phe Glu Lys His Pro
        515                 520                 525

Glu Ile Asn Ala Val Gln Leu Gln Asn Leu Leu Asn Gln Met Thr Trp
    530                 535                 540

Ser Ser Leu Gly Ser Arg Gln Pro Phe Phe Ser Leu Glu Ala Cys Gln
545                 550                 555                 560

Gly Ile Leu Ala Leu Leu Asp Leu Asn Ala Ser Gly Thr Met Ser Ile
                565                 570                 575

Gln Glu Phe Arg Asp Leu Trp Lys Gln Leu Lys Leu Ser Gln Lys Val
            580                 585                 590

Phe His Lys Gln Asp Arg Gly Ser Gly Tyr Leu Asn Trp Glu Gln Leu
        595                 600                 605

His Ala Ala Met Arg Glu Ala Gly Arg His Arg Lys Ser Trp Ser Cys
    610                 615                 620

Gly His Thr Arg Ala Gly Cys Thr Leu Ile Arg Gln Arg Arg Gly Asp
625                 630                 635                 640

Val Trp His Ala Glu Val Thr Leu Ile Arg Ser Val Thr Leu Lys Asp
                645                 650                 655

Val Asp Leu Gln Ser Thr Pro Thr Phe Phe Met Ile Val Pro Val Ile
            660                 665                 670

Leu Ala Asn Ile Asp Gly Gly Val Ala His Ser Thr Ser Tyr Leu Ile
        675                 680                 685

Phe Asn Thr Thr Leu Leu
    690


<210> SEQ ID NO 11
<211> LENGTH: 713
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

Met Thr Glu Glu Leu Ile Thr Pro Val Tyr Cys Thr Gly Val Ser Ala
1               5                   10                  15

Gln Val Gln Lys Lys Arg Asp Lys Glu Leu Gly Leu Gly Arg His Glu
            20                  25                  30

Asn Ala Ile Lys Tyr Leu Gly Gln Asp Tyr Glu Thr Leu Arg Ala Arg
        35                  40                  45

Cys Leu Gln Ser Gly Val Leu Phe Gln Asp Glu Ala Phe Pro Pro Val
    50                  55                  60

Ser His Ser Leu Gly Phe Lys Glu Leu Gly Pro His Ser Ser Lys Thr
65                  70                  75                  80
```

-continued

```
Tyr Gly Ile Lys Trp Lys Arg Pro Thr Glu Leu Met Ser Asn Pro Gln
                85                  90                  95

Phe Ile Val Asp Gly Ala Thr Arg Thr Asp Ile Cys Gln Gly Ala Leu
            100                 105                 110

Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala Ser Leu Thr Leu Asn Glu
        115                 120                 125

Thr Ile Leu His Arg Val Val Pro Tyr Gly Gln Ser Phe Gln Asp Gly
    130                 135                 140

Tyr Ala Gly Ile Phe His Phe Gln Leu Trp Gln Phe Gly Glu Trp Val
145                 150                 155                 160

Asp Val Val Ile Asp Asp Leu Leu Pro Thr Lys Asp Gly Lys Leu Val
                165                 170                 175

Phe Val His Ser Ala Gln Gly Asn Glu Phe Trp Ser Ala Leu Leu Glu
            180                 185                 190

Lys Ala Tyr Ala Lys Val Asn Gly Ser Tyr Glu Ala Leu Ser Gly Gly
        195                 200                 205

Cys Thr Ser Glu Ala Phe Glu Asp Phe Thr Gly Gly Val Thr Glu Trp
    210                 215                 220

Tyr Asp Leu Gln Lys Ala Pro Ser Asp Leu Tyr Gln Ile Ile Leu Lys
225                 230                 235                 240

Ala Leu Glu Arg Gly Ser Leu Leu Gly Cys Ser Ile Asn Ile Ser Asp
                245                 250                 255

Ile Arg Asp Leu Glu Ala Ile Thr Phe Lys Asn Leu Val Arg Gly His
            260                 265                 270

Ala Tyr Ser Val Thr Gly Ala Lys Gln Val Thr Tyr Gln Gly Gln Arg
        275                 280                 285

Val Asn Leu Ile Arg Met Arg Asn Pro Trp Gly Glu Val Glu Trp Lys
    290                 295                 300

Gly Pro Trp Ser Asp Ser Ser Tyr Glu Trp Asn Lys Val Asp Pro Tyr
305                 310                 315                 320

Glu Arg Glu Gln Leu Arg Val Lys Met Glu Asp Gly Glu Phe Trp Met
                325                 330                 335

Ser Phe Arg Asp Phe Ile Arg Glu Phe Thr Lys Leu Glu Ile Cys Asn
            340                 345                 350

Leu Thr Pro Asp Ala Leu Lys Ser Arg Thr Leu Arg Asn Trp Asn Thr
        355                 360                 365

Thr Phe Tyr Glu Gly Thr Trp Arg Arg Gly Ser Thr Ala Gly Gly Cys
    370                 375                 380

Arg Asn Tyr Pro Ala Thr Phe Trp Val Asn Pro Gln Phe Lys Ile Arg
385                 390                 395                 400

Leu Glu Glu Val Asp Asp Ala Asp Asp Tyr Asp Asn Arg Glu Ser Gly
                405                 410                 415

Cys Ser Phe Leu Leu Ala Leu Met Gln Lys His Arg Arg Arg Glu Arg
            420                 425                 430

Arg Phe Gly Arg Asp Met Glu Thr Ile Gly Phe Ala Val Tyr Gln Val
        435                 440                 445

Pro Arg Glu Leu Ala Gly Gln Pro Val His Leu Lys Arg Asp Phe Phe
    450                 455                 460

Leu Ala Asn Ala Ser Arg Ala Gln Ser Glu His Phe Ile Asn Leu Arg
465                 470                 475                 480
```

-continued

```
Glu Val Ser Asn Arg Ile Arg Pro Pro Pro Gly Glu Tyr Ile Val Val
                485                 490                 495

Pro Ser Thr Phe Glu Pro Asn Lys Glu Gly Asp Phe Leu Leu Arg Phe
            500                 505                 510

Phe Ser Glu Lys Lys Ala Gly Thr Gln Glu Leu Asp Asp Gln Ile Gln
        515                 520                 525

Ala Asn Leu Pro Asp Glu Lys Val Leu Ser Glu Glu Glu Ile Asp Asp
    530                 535                 540

Asn Phe Lys Thr Leu Phe Ser Lys Leu Ala Gly Asp Asp Met Glu Ile
545                 550                 555                 560

Ser Val Lys Glu Leu Gln Thr Ile Leu Asn Arg Ile Ile Ser Lys His
                565                 570                 575

Lys Asp Leu Arg Thr Asn Gly Phe Ser Leu Glu Ser Cys Arg Ser Met
            580                 585                 590

Val Asn Leu Met Asp Arg Asp Gly Asn Gly Lys Leu Gly Leu Val Glu
        595                 600                 605

Phe Asn Ile Leu Trp Asn Arg Ile Arg Asn Tyr Leu Thr Ile Phe Arg
    610                 615                 620

Lys Phe Asp Leu Asp Lys Ser Gly Ser Met Ser Ala Tyr Glu Met Arg
625                 630                 635                 640

Met Ala Ile Glu Ala Ala Gly Phe Lys Leu Asn Lys Lys Leu His Glu
                645                 650                 655

Leu Ile Ile Thr Arg Tyr Ser Glu Pro Asp Leu Ala Val Asp Phe Asp
            660                 665                 670

Asn Phe Val Cys Cys Leu Val Arg Leu Glu Thr Met Phe Arg Phe Phe
        675                 680                 685

Lys Leu Leu Asp Thr Asp Leu Asn Gly Val Val Thr Phe Asp Leu Phe
    690                 695                 700

Lys Trp Leu Gln Leu Thr Met Phe Ala
705                 710


<210> SEQ ID NO 12
<211> LENGTH: 700
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

Met Ala Gly Ile Ala Ile Lys Leu Ala Lys Asp Arg Glu Ala Ala Glu
1               5                   10                  15

Gly Leu Gly Ser His Glu Arg Ala Ile Lys Tyr Leu Asn Gln Asp Tyr
            20                  25                  30

Glu Thr Leu Arg Asn Glu Cys Leu Glu Ala Gly Ala Leu Phe Gln Asp
        35                  40                  45

Pro Ser Phe Pro Ala Leu Pro Ser Ser Leu Gly Tyr Lys Glu Leu Gly
    50                  55                  60

Pro Tyr Ser Ser Lys Thr Arg Gly Ile Glu Trp Lys Arg Pro Thr Glu
65                  70                  75                  80

Ile Cys Ala Asp Pro Gln Phe Ile Ile Gly Gly Ala Thr Arg Thr Asp
                85                  90                  95

Ile Cys Gln Gly Ala Leu Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala
            100                 105                 110

Ser Leu Thr Leu Asn Glu Glu Ile Leu Ala Arg Val Val Pro Pro Asp
        115                 120                 125
```

-continued

```
Gln Ser Phe Gln Glu Asn Tyr Ala Gly Ile Phe His Phe Gln Phe Trp
    130                 135                 140

Gln Tyr Gly Glu Trp Val Glu Val Val Val Asp Asp Arg Leu Pro Thr
145                 150                 155                 160

Lys Asp Gly Glu Leu Leu Phe Val His Ser Ala Glu Gly Ser Glu Phe
                165                 170                 175

Trp Ser Ala Leu Leu Glu Lys Ala Tyr Ala Lys Ile Asn Gly Cys Tyr
            180                 185                 190

Glu Ala Leu Ser Gly Gly Ala Thr Thr Glu Gly Phe Glu Asp Phe Thr
        195                 200                 205

Gly Gly Ile Gly Glu Trp Tyr Glu Leu Arg Lys Pro Pro Pro Asn Leu
    210                 215                 220

Phe Lys Ile Ile Gln Lys Ala Leu Glu Lys Gly Ser Leu Leu Gly Cys
225                 230                 235                 240

Ser Ile Asp Ile Thr Ser Ala Ala Asp Ser Glu Ala Val Thr Tyr Gln
                245                 250                 255

Lys Leu Val Lys Gly His Ala Tyr Ser Val Thr Gly Ala Glu Glu Val
            260                 265                 270

Glu Ser Ser Gly Ser Leu Gln Lys Leu Ile Arg Ile Arg Asn Pro Trp
        275                 280                 285

Gly Gln Val Glu Trp Thr Gly Lys Trp Asn Asp Asn Cys Pro Ser Trp
    290                 295                 300

Asn Thr Val Asp Pro Glu Val Arg Ala Asn Leu Thr Glu Arg Gln Glu
305                 310                 315                 320

Asp Gly Glu Phe Trp Met Ser Phe Ser Asp Phe Leu Arg His Tyr Ser
                325                 330                 335

Arg Leu Glu Ile Cys Asn Leu Thr Pro Asp Thr Leu Thr Cys Asp Ser
            340                 345                 350

Tyr Lys Lys Trp Lys Leu Thr Lys Met Asp Gly Asn Trp Arg Arg Gly
        355                 360                 365

Ser Thr Ala Gly Gly Cys Arg Asn Tyr Pro Asn Thr Phe Trp Met Asn
    370                 375                 380

Pro Gln Tyr Leu Ile Lys Leu Glu Glu Glu Asp Glu Asp Glu Glu Asp
385                 390                 395                 400

Gly Glu Arg Gly Cys Thr Phe Leu Val Gly Leu Ile Gln Lys His Arg
                405                 410                 415

Arg Arg Gln Arg Lys Met Gly Glu Asp Met His Thr Ile Gly Phe Gly
            420                 425                 430

Ile Tyr Glu Val Pro Glu Glu Leu Thr Gly Gln Thr Asn Ile His Leu
        435                 440                 445

Gly Lys Asn Phe Phe Leu Thr Thr Arg Ala Arg Glu Arg Ser Asp Thr
    450                 455                 460

Phe Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly
465                 470                 475                 480

Glu Tyr Val Leu Val Pro Ser Thr Phe Glu Pro His Lys Asp Gly Asp
                485                 490                 495

Phe Cys Ile Arg Val Phe Ser Glu Lys Lys Ala Asp Tyr Gln Ala Val
            500                 505                 510

Asp Asp Glu Ile Glu Ala Asn Ile Glu Glu Ile Asp Ala Asn Glu Glu
        515                 520                 525
```

-continued

```
Asp Ile Asp Asp Gly Phe Arg Arg Leu Phe Val Gln Leu Ala Gly Glu
    530                 535                 540

Asp Ala Glu Ile Ser Ala Phe Glu Leu Gln Thr Ile Leu Arg Arg Val
545                 550                 555                 560

Leu Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr
                565                 570                 575

Cys Lys Ile Met Val Asp Met Leu Asp Glu Asp Gly Ser Gly Lys Leu
            580                 585                 590

Gly Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln
        595                 600                 605

Lys Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser
    610                 615                 620

Tyr Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Leu Pro Cys
625                 630                 635                 640

Gln Leu His Gln Val Ile Val Ala Arg Phe Ala Asp Asp Glu Leu Ile
                645                 650                 655

Ile Asp Phe Asp Asn Phe Val Arg Cys Leu Val Arg Leu Glu Thr Leu
            660                 665                 670

Phe Lys Ile Phe Lys Gln Leu Asp Pro Glu Asn Thr Gly Thr Ile Gln
        675                 680                 685

Leu Asn Leu Ala Ser Trp Leu Ser Phe Ser Val Leu
    690                 695                 700


<210> SEQ ID NO 13
<211> LENGTH: 641
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 13

Met Gly Pro Pro Leu Lys Leu Phe Lys Asn Gln Lys Tyr Gln Glu Leu
1               5                   10                  15

Lys Gln Glu Cys Met Lys Asp Gly Arg Leu Phe Cys Asp Pro Thr Phe
            20                  25                  30

Leu Pro Glu Asn Asp Ser Leu Phe Phe Asn Arg Leu Leu Pro Gly Lys
        35                  40                  45

Val Val Trp Lys Arg Pro Gln Asp Ile Ser Asp Asp Pro His Leu Ile
    50                  55                  60

Val Gly Asn Ile Ser Asn His Gln Leu Ile Gln Gly Arg Leu Gly Asn
65                  70                  75                  80

Lys Ala Met Ile Ser Ala Phe Ser Cys Leu Ala Val Gln Glu Ser His
                85                  90                  95

Trp Thr Lys Ala Ile Pro Asn His Lys Asp Gln Glu Trp Asp Pro Arg
            100                 105                 110

Lys Pro Glu Lys Tyr Ala Gly Ile Phe His Phe Arg Phe Trp His Phe
        115                 120                 125

Gly Glu Trp Thr Glu Val Val Ile Asp Asp Leu Leu Pro Thr Ile Asn
    130                 135                 140

Gly Asp Leu Val Phe Ser Phe Ser Thr Ser Met Asn Glu Phe Trp Asn
145                 150                 155                 160

Ala Leu Leu Glu Lys Ala Tyr Ala Lys Leu Leu Gly Cys Tyr Glu Ala
                165                 170                 175

Leu Asp Gly Leu Thr Ile Thr Asp Ile Ile Met Asp Phe Thr Gly Thr
            180                 185                 190
```

-continued

```
Leu Ala Glu Ile Ile Asp Met Gln Lys Gly Arg Tyr Thr Asp Leu Val
        195                 200                 205

Glu Glu Lys Tyr Lys Leu Phe Gly Glu Leu Tyr Lys Thr Phe Thr Lys
    210                 215                 220

Gly Gly Leu Ile Cys Cys Ser Ile Glu Ser Pro Ser Gln Glu Glu Gln
225                 230                 235                 240

Glu Val Glu Thr Asp Trp Gly Leu Leu Lys Gly Tyr Thr Tyr Thr Met
                245                 250                 255

Thr Asp Ile Arg Lys Leu Arg Leu Gly Glu Arg Leu Val Glu Val Phe
            260                 265                 270

Ser Thr Glu Lys Leu Tyr Met Val Arg Leu Arg Asn Pro Leu Gly Arg
        275                 280                 285

Gln Glu Trp Ser Gly Pro Trp Ser Glu Ile Ser Glu Glu Trp Gln Gln
    290                 295                 300

Leu Thr Val Thr Asp Arg Lys Asn Leu Gly Leu Val Met Ser Asp Asp
305                 310                 315                 320

Gly Glu Phe Trp Met Ser Leu Glu Asp Phe Cys His Asn Phe His Lys
                325                 330                 335

Leu Asn Val Cys Arg Asn Val Asn Asn Pro Val Phe Gly Arg Lys Glu
            340                 345                 350

Leu Glu Ser Val Val Gly Cys Trp Thr Val Asp Asp Asp Pro Leu Met
        355                 360                 365

Asn Arg Ser Gly Gly Cys Tyr Asn Asn Arg Asp Thr Phe Leu Gln Asn
    370                 375                 380

Pro Gln Tyr Ile Phe Thr Val Pro Glu Asp Gly His Lys Val Ile Met
385                 390                 395                 400

Ser Leu Gln Gln Lys Asp Leu Arg Thr Tyr Arg Arg Met Gly Arg Pro
                405                 410                 415

Asp Asn Tyr Ile Ile Gly Phe Glu Leu Phe Lys Val Glu Met Asn Arg
            420                 425                 430

Arg Phe Arg Leu His His Leu Tyr Ile Gln Glu Arg Ala Gly Thr Ser
        435                 440                 445

Thr Tyr Ile Asp Thr Arg Thr Val Phe Leu Ser Lys Tyr Leu Lys Lys
    450                 455                 460

Gly Ser Tyr Val Leu Val Pro Thr Met Phe Gln His Gly Arg Thr Ser
465                 470                 475                 480

Glu Phe Leu Leu Arg Ile Phe Ser Glu Val Pro Val Gln Leu Arg Glu
                485                 490                 495

Leu Thr Leu Asp Met Pro Lys Met Ser Cys Trp Asn Leu Ala Arg Gly
            500                 505                 510

Tyr Pro Lys Val Val Thr Gln Ile Thr Val His Ser Ala Glu Gly Leu
        515                 520                 525

Glu Lys Lys Tyr Ala Asn Glu Thr Val Asn Pro Tyr Leu Ile Ile Lys
    530                 535                 540

Cys Gly Lys Glu Glu Val Arg Ser Pro Val Gln Lys Asn Thr Val His
545                 550                 555                 560

Ala Ile Phe Asp Thr Gln Ala Val Phe Tyr Arg Arg Thr Thr Asp Ile
                565                 570                 575

Pro Ile Ile Ile Gln Val Trp Asn Ser Arg Lys Phe Cys Asp Gln Phe
            580                 585                 590
```

-continued

```
Leu Gly Gln Val Thr Leu Asp Ala Asp Pro Ser Asp Cys Arg Asp Leu
        595                 600                 605

Lys Ser Leu Tyr Leu Arg Lys Lys Gly Gly Pro Thr Ala Lys Val Lys
    610                 615                 620

Gln Gly His Ile Ser Phe Lys Val Ile Ser Ser Asp Asp Leu Thr Glu
625                 630                 635                 640

Leu
```

<210> SEQ ID NO 14
<211> LENGTH: 813
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 14

```
Met Asp Ala Ser Ala Leu Glu Arg Asp Ala Val Gln Phe Ala Arg Leu
1               5                   10                  15

Ala Val Gln Arg Asp His Glu Gly Arg Tyr Ser Glu Ala Val Phe Tyr
            20                  25                  30

Tyr Lys Glu Ala Ala Gln Ala Leu Ile Tyr Ala Glu Met Ala Gly Ser
        35                  40                  45

Ser Leu Glu Arg Ile Gln Glu Lys Ile Asn Glu Tyr Leu Glu Arg Val
    50                  55                  60

Gln Ala Leu His Ser Ala Val Gln Ser Lys Ser Thr Asp Pro Leu Lys
65                  70                  75                  80

Ser Lys His Gln Leu Asp Leu Glu Arg Ala His Phe Leu Val Thr Gln
                85                  90                  95

Ala Phe Asp Glu Asp Glu Lys Gly Asn Val Glu Asp Ala Ile Glu Leu
            100                 105                 110

Tyr Thr Glu Ala Val Glu Leu Cys Leu Lys Thr Ser Ser Glu Thr Ala
        115                 120                 125

Asp Lys Thr Leu Gln Asn Lys Leu Lys Gln Leu Ala Arg Gln Ala Leu
    130                 135                 140

Asp Arg Ala Glu Ala Leu Ser Glu Pro Leu Thr Lys Pro Phe Cys Lys
145                 150                 155                 160

Leu Lys Ser Ala Asn Met Lys Thr Lys Thr Pro Pro Val Arg Thr His
                165                 170                 175

Phe Pro Leu Gly Pro Asn Pro Phe Val Glu Lys Pro Gln Ala Phe Ile
            180                 185                 190

Ser Pro Gln Ser Cys Asp Ala Gln Gly Gln Lys Tyr Thr Ala Glu Glu
        195                 200                 205

Ile Glu Val Leu Arg Thr Thr Ser Lys Ile Asn Gly Val Glu Tyr Val
    210                 215                 220

Pro Phe Met Ser Val Asp Leu Arg Glu Arg Phe Ala Tyr Pro Met Pro
225                 230                 235                 240

Phe Cys Asp Arg Leu Gly Lys Leu Pro Leu Ser Pro Lys Gln Lys Thr
                245                 250                 255

Thr Phe Ser Lys Trp Val Arg Pro Glu Asp Leu Thr Asn Asn Pro Thr
            260                 265                 270

Met Ile Tyr Thr Val Ser Ser Phe Ser Ile Lys Gln Thr Ile Val Ser
        275                 280                 285

Asp Cys Ser Phe Val Ala Ser Leu Ala Ile Ser Ala Ala Tyr Glu Arg
    290                 295                 300
```

-continued

```
Arg Phe Asn Lys Lys Leu Ile Thr Ser Ile Ile Tyr Pro Gln Asn Lys
305                 310                 315                 320

Asp Gly Glu Pro Glu Tyr Asn Pro Cys Gly Lys Tyr Met Val Lys Leu
                325                 330                 335

His Leu Asn Gly Val Pro Arg Lys Val Ile Ile Asp Asp Gln Leu Pro
            340                 345                 350

Val Asp His Lys Gly Glu Leu Leu Cys Ser Tyr Ser Asn Asn Lys Ser
        355                 360                 365

Glu Leu Trp Val Ser Leu Ile Glu Lys Ala Tyr Met Lys Val Met Gly
    370                 375                 380

Gly Tyr Asp Phe Pro Gly Ser Asn Ser Asn Ile Asp Leu His Ala Leu
385                 390                 395                 400

Thr Gly Trp Ile Pro Glu Arg Ile Ala Met His Ser Asp Ser Gln Thr
                405                 410                 415

Phe Ser Lys Asp Asn Ser Phe Arg Met Leu Tyr Gln Arg Phe His Lys
            420                 425                 430

Gly Asp Val Leu Ile Thr Ala Ser Thr Gly Val Met Thr Glu Ala Glu
        435                 440                 445

Gly Glu Lys Trp Gly Leu Val Pro Thr His Ala Tyr Ala Val Leu Asp
    450                 455                 460

Ile Arg Glu Phe Lys Gly Leu Arg Phe Ile Gln Leu Lys Asn Pro Trp
465                 470                 475                 480

Ser His Leu Arg Trp Lys Gly Arg Tyr Ser Glu Asn Asp Val Lys Asn
                485                 490                 495

Trp Thr Pro Glu Leu Gln Lys Tyr Leu Asn Phe Asp Pro Arg Thr Ala
            500                 505                 510

Gln Lys Ile Asp Asn Gly Ile Phe Trp Ile Ser Trp Asp Asp Leu Cys
        515                 520                 525

Gln Tyr Tyr Asp Val Val Tyr Leu Ser Trp Asn Pro Ala Leu Phe Lys
    530                 535                 540

Glu Ser Thr Cys Ile His Ser Thr Trp Asp Ala Lys Gln Gly Pro Val
545                 550                 555                 560

Lys Asp Ala Tyr Ser Leu Ala Asn Asn Pro Gln Tyr Lys Leu Glu Val
                565                 570                 575

Gln Cys Pro Gln Gly Gly Ala Ala Val Trp Val Leu Leu Ser Arg His
            580                 585                 590

Ile Thr Asp Lys Asp Asp Phe Ala Asn Asn Arg Glu Phe Ile Thr Met
        595                 600                 605

Val Val Tyr Lys Thr Asp Gly Lys Lys Val Tyr Tyr Pro Ala Asp Pro
    610                 615                 620

Pro Pro Tyr Ile Asp Gly Ile Arg Ile Asn Ser Pro His Tyr Leu Thr
625                 630                 635                 640

Lys Ile Lys Leu Thr Thr Pro Gly Thr His Thr Phe Thr Leu Val Val
                645                 650                 655

Ser Gln Tyr Glu Lys Gln Asn Thr Ile His Tyr Thr Val Arg Val Tyr
            660                 665                 670

Ser Ala Cys Ser Phe Thr Phe Ser Lys Ile Pro Ser Pro Tyr Thr Leu
        675                 680                 685

Ser Lys Arg Ile Asn Gly Lys Trp Ser Gly Gln Ser Ala Gly Gly Cys
    690                 695                 700

Gly Asn Phe Gln Glu Thr His Lys Asn Asn Pro Ile Tyr Gln Phe His
```

-continued

```
705                 710                 715                 720

Ile Asp Lys Thr Gly Pro Leu Leu Ile Glu Leu Arg Gly Pro Arg Gln
                725                 730                 735

Tyr Ser Val Gly Phe Glu Val Val Ala Val Ser Ile Met Gly Asp Pro
            740                 745                 750

Gly Pro His Gly Phe Gln Arg Lys Ser Ser Gly Asp Tyr Arg Cys Gly
        755                 760                 765

Phe Cys Tyr Leu Glu Leu Glu Asn Ile Pro Ala Gly Ile Phe Asn Ile
    770                 775                 780

Ile Pro Ser Thr Phe Leu Pro Lys Gln Glu Gly Pro Phe Phe Leu Asp
785                 790                 795                 800

Phe Asn Ser Thr Val Pro Ile Lys Thr Thr Gln Leu Gln
                805                 810


<210> SEQ ID NO 15
<211> LENGTH: 605
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 15

Met Arg Ala Val Arg Ala Glu Thr Pro Ala Arg Glu Leu Phe Arg Asp
1               5                   10                  15

Ala Ala Phe Pro Ala Ser Asp Ser Ser Leu Phe Tyr Asn Leu Ser Thr
            20                  25                  30

Pro Leu Ala Gln Phe Arg Glu Asp Ile Thr Trp Arg Arg Pro Gln Glu
        35                  40                  45

Ile Cys Ala Thr Pro Gln Leu Phe Pro Asp Asn Pro Trp Glu Gly Gln
    50                  55                  60

Val Lys Gln Gly Leu Leu Gly Asp Cys Trp Phe Leu Cys Ala Cys Ala
65                  70                  75                  80

Ala Leu Gln Lys Ser Gln His Leu Leu Asp Gln Val Phe Pro Pro Gly
                85                  90                  95

Gln Pro Gly Trp Ser Asp Gln Lys Tyr Gln Gly Phe Phe Thr Cys Arg
            100                 105                 110

Ile Trp Gln Phe Gly His Trp Glu Glu Val Thr Ile Asp Asp Arg Leu
        115                 120                 125

Pro Cys Leu Ala Gly Arg Leu Cys Phe Ser Arg Cys Gln Arg Glu Asp
    130                 135                 140

Val Phe Trp Leu Pro Leu Leu Glu Lys Ala Tyr Ala Lys Val His Gly
145                 150                 155                 160

Ser Tyr Glu His Leu Trp Ala Gly Gln Val Ala Asp Ala Leu Val Asp
                165                 170                 175

Leu Thr Gly Ser Leu Ala Glu Arg Trp Ser Leu Lys Asp Val Thr Lys
            180                 185                 190

Ala Ser Gly Gln Gln Asp Arg Pro Ser Gly Gly Glu His Arg Thr Cys
        195                 200                 205

Arg Gln Leu Leu His Leu Lys Asp Arg Cys Leu Ile Ser Cys Ser Val
    210                 215                 220

Leu Ser Pro Arg Ala Gly Ala Arg Glu Leu Gly Glu Phe His Ala Phe
225                 230                 235                 240

Ile Ile Ser Asp Leu Gln Glu Leu Arg Ser Gln Thr Gly Gln Gly Ile
                245                 250                 255
```

-continued

```
Leu Leu Leu Arg Ile His Asn Pro Trp Gly Arg Arg Cys Trp Gln Gly
            260                 265                 270

Leu Trp Arg Glu Gly Gly Glu Gly Trp Asn Gln Val Glu Pro Ala Lys
        275                 280                 285

Glu Ser Glu Leu Leu Ala Gln Leu Gln Glu Gly Glu Phe Trp Val Glu
    290                 295                 300

Glu Glu Glu Phe Leu Arg Glu Phe Asp Glu Val Thr Ile Gly Tyr Pro
305                 310                 315                 320

Val Thr Glu Ala Gly His Leu Gln Ser Leu His Thr Glu Arg Val Leu
                325                 330                 335

Cys His Thr Arg Thr Leu Pro Gly Ala Trp Val Thr Gly Gln Ser Ala
            340                 345                 350

Gly Gly Cys Arg Asn Asn Ser Cys Phe Pro Cys Asn Pro Lys Phe Trp
        355                 360                 365

Leu Arg Leu Leu Glu Pro Ser Glu Val Cys Val Ala Val Leu Gln Arg
    370                 375                 380

Pro Arg Arg Arg Leu Val Gly Gln Thr Arg Ala Leu Ala Gly Ala Ser
385                 390                 395                 400

Pro Ala Pro Val Asn Leu Pro Gly Lys Asp Tyr Gln Ala Val Gly Leu
                405                 410                 415

His Ile Trp Lys Val Glu Lys Arg Lys Ile Ser Leu Pro Arg Val Leu
            420                 425                 430

Ser Ala Pro Pro Val Ala Gly Thr Ala Cys His Ala Tyr Asp Arg Glu
        435                 440                 445

Ile His Leu Arg Cys Glu Leu Ser Pro Gly Tyr Tyr Leu Ala Val Pro
    450                 455                 460

Ser Thr Phe Leu Lys Asp Val Pro Gly Gln Phe Leu Leu Arg Val Phe
465                 470                 475                 480

Ser Thr Gly Lys Ile Ser Leu Ser Ala Val Arg Leu Ala Thr Lys Gly
                485                 490                 495

Ala Ser Pro Gly Thr Ala Leu Pro Ala Gly Glu Trp Glu Thr Val Gln
            500                 505                 510

Leu Gln Gly Cys Trp Arg Ala Gly Gln Thr Ala Gly Gly Ser Arg Asn
        515                 520                 525

Phe Ala Ser Tyr Pro Cys Asn Pro Cys Leu Pro Phe Ser Val Pro Glu
    530                 535                 540

Gly Ala Gly Pro Arg Tyr Ile Arg Ile Thr Leu Gln Gln His Cys Arg
545                 550                 555                 560

Leu Ser Asp Ser Gln Leu His Pro Ile Gly Phe His Val Phe Gln Val
                565                 570                 575

Pro Ala Asp Gly Glu Asn Gln Asp Ala Cys Ser Leu Leu Leu Gln Glu
            580                 585                 590

Pro Leu Leu Ser Cys Val Pro His Arg Thr Pro Arg Lys
        595                 600                 605


<210> SEQ ID NO 16
<211> LENGTH: 720
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 16

Met Ala Ser Gly Asn Arg Lys Val Thr Ile Gln Leu Val Asp Asp Gly
1               5                   10                  15
```

-continued

```
Ala Gly Thr Gly Ala Gly Gly Pro Gln Leu Phe Lys Gly Gln Asn Tyr
            20                  25                  30

Glu Ala Ile Arg Arg Ala Cys Leu Asp Ser Gly Ile Leu Phe Arg Asp
        35                  40                  45

Pro Cys Phe Pro Ala Gly Pro Asp Ala Leu Gly Tyr Asp Lys Leu Gly
    50                  55                  60

Pro Asp Ser Glu Lys Ala Lys Gly Val Glu Trp Lys Arg Pro His Glu
65                  70                  75                  80

Phe Cys Ala Glu Pro Gln Phe Ile Cys Glu Asp Met Ser Arg Thr Asp
                85                  90                  95

Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Leu Leu Ala Ala Ala Ala
            100                 105                 110

Ser Leu Thr Leu Tyr Pro Arg Leu Leu Tyr Arg Val Val Pro Pro Gly
        115                 120                 125

Gln Gly Phe Gln Asp Gly Tyr Ala Gly Val Phe His Phe Gln Leu Trp
    130                 135                 140

Gln Phe Gly Arg Trp Val Asp Val Val Val Asp Asp Lys Leu Pro Val
145                 150                 155                 160

Arg Glu Gly Lys Leu Met Phe Val Arg Ser Glu Gln Arg Asn Glu Phe
                165                 170                 175

Trp Ala Pro Leu Leu Glu Lys Ala Tyr Ala Lys Leu His Gly Ser Tyr
            180                 185                 190

Glu Val Met Arg Gly Gly His Met Asn Glu Ala Phe Val Asp Phe Thr
        195                 200                 205

Gly Gly Val Gly Glu Val Leu Tyr Leu Arg Gln Asn Thr Pro Gly Val
    210                 215                 220

Phe Ala Ala Leu Arg His Ala Leu Ala Lys Glu Ser Leu Val Gly Ala
225                 230                 235                 240

Thr Ala Leu Ser Asp Arg Gly Glu Ile Arg Thr Asp Glu Gly Leu Val
                245                 250                 255

Lys Gly His Ala Tyr Ser Val Thr Gly Thr His Lys Met Ser Leu Gly
            260                 265                 270

Phe Thr Lys Val Arg Leu Leu Arg Leu Arg Asn Pro Trp Gly Arg Val
        275                 280                 285

Glu Trp Ser Gly Pro Trp Ser Asp Ser Cys Pro Arg Trp Asp Met Leu
    290                 295                 300

Pro Ser Glu Trp Arg Asp Ala Leu Leu Val Lys Lys Glu Asp Gly Glu
305                 310                 315                 320

Phe Trp Met Glu Leu Gln Asp Phe Leu Thr His Phe Asn Thr Val Gln
                325                 330                 335

Ile Cys Ser Leu Ser Pro Glu Val Leu Gly Pro Ser Pro Ala Gly Gly
            340                 345                 350

Gly Trp His Ile His Ile Phe Gln Gly Arg Trp Val Arg Gly Phe Asn
        355                 360                 365

Ser Gly Gly Ser Gln Pro Ser Ala Glu Asn Phe Trp Thr Asn Pro Gln
    370                 375                 380

Phe Arg Leu Thr Leu Leu Glu Pro Asp Glu Glu Glu Asp Asp Asp Asp
385                 390                 395                 400

Glu Glu Gly Pro Trp Gly Gly Trp Gly Ala Ala Gly Ala Arg Gly Pro
                405                 410                 415
```

-continued

```
Ala Arg Gly Gly Arg Val Pro Lys Cys Thr Val Leu Leu Ser Leu Ile
            420                 425                 430

Gln Arg Asn Arg Arg Cys Leu Arg Ala Lys Gly Leu Thr Tyr Leu Thr
        435                 440                 445

Val Gly Phe His Val Phe Gln Ile Pro Glu Glu Leu Leu Asp Leu Trp
    450                 455                 460

Asp Ser Pro Arg Ser Arg Ala Leu Leu Pro Gly Leu Leu Arg Ala Asp
465                 470                 475                 480

Arg Ser Val Phe Cys Ala Arg Arg Asp Val Ser Arg Arg Cys Arg Leu
                485                 490                 495

Pro Pro Gly His Tyr Leu Val Val Pro Ser Ala Ser Arg Val Gly Asp
            500                 505                 510

Glu Ala Asp Phe Thr Leu Arg Ile Phe Ser Glu Arg Ser His Thr Ala
        515                 520                 525

Val Glu Ile Asp Asp Val Ile Ser Ala Asp Leu Asp Ala Leu Gln Ala
    530                 535                 540

Pro Tyr Lys Pro Leu Glu Leu Glu Leu Ala Gln Leu Phe Leu Glu Leu
545                 550                 555                 560

Ala Gly Glu Glu Glu Glu Leu Asn Ala Leu Gln Leu Gln Thr Leu Ile
                565                 570                 575

Ser Ile Ala Leu Glu Pro Ala Arg Ala Asn Thr Arg Thr Pro Gly Glu
            580                 585                 590

Ile Gly Leu Arg Thr Cys Glu Gln Leu Val Gln Cys Phe Gly Arg Gly
        595                 600                 605

Gln Arg Leu Ser Leu His His Phe Gln Glu Leu Trp Gly His Leu Met
    610                 615                 620

Ser Trp Gln Ala Thr Phe Asp Lys Phe Asp Glu Asp Ala Ser Gly Thr
625                 630                 635                 640

Met Asn Ser Cys Glu Leu Arg Leu Ala Leu Thr Ala Ala Gly Phe His
                645                 650                 655

Leu Asn Asn Gln Leu Thr Gln Ser Leu Thr Ser Arg Tyr Arg Asp Ser
            660                 665                 670

Arg Leu Arg Val Asp Phe Glu Arg Phe Val Gly Cys Ala Ala Arg Leu
        675                 680                 685

Thr Cys Ile Phe Arg His Cys Cys Gln His Leu Asp Gly Gly Glu Gly
    690                 695                 700

Val Val Cys Leu Thr His Lys Gln Trp Ser Glu Val Ala Thr Phe Ser
705                 710                 715                 720


<210> SEQ ID NO 17
<211> LENGTH: 20
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 17

cagagctatg aggcaattcg                                              20


<210> SEQ ID NO 18
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 18

tcatccattt cacgccttt                                               19
```

-continued

<210> SEQ ID NO 19
<211> LENGTH: 699
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 19

Ala Gly Ile Ala Ala Lys Leu Ala Lys Asp Arg Glu Ala Ala Glu Gly
1 5 10 15

Leu Gly Ser His Glu Arg Ala Ile Lys Tyr Leu Asn Gln Asp Tyr Glu
20 25 30

Ala Leu Arg Asn Glu Cys Leu Glu Ala Gly Thr Leu Phe Gln Asp Pro
35 40 45

Ser Phe Pro Ala Ile Pro Ser Ala Leu Gly Phe Lys Glu Leu Gly Pro
50 55 60

Tyr Ser Ser Lys Thr Arg Gly Met Arg Trp Lys Arg Pro Thr Glu Ile
65 70 75 80

Cys Ala Asp Pro Gln Phe Ile Ile Gly Gly Ala Thr Arg Thr Asp Ile
85 90 95

Cys Gln Gly Ala Leu Gly Asp Cys Trp Leu Leu Ala Ala Ile Ala Ser
100 105 110

Leu Thr Leu Asn Glu Glu Ile Leu Ala Arg Val Val Pro Leu Asn Gln
115 120 125

Ser Phe Gln Glu Asn Tyr Ala Gly Ile Phe His Phe Gln Phe Trp Gln
130 135 140

Tyr Gly Glu Trp Val Glu Val Val Val Asp Asp Arg Leu Pro Thr Lys
145 150 155 160

Asp Gly Glu Leu Leu Phe Val His Ser Ala Glu Gly Ser Glu Phe Trp
165 170 175

Ser Ala Leu Leu Glu Lys Ala Tyr Ala Lys Ile Asn Gly Cys Tyr Glu
180 185 190

Ala Leu Ser Gly Gly Ala Thr Thr Glu Gly Phe Glu Asp Phe Thr Gly
195 200 205

Gly Leu Ala Glu Trp Tyr Glu Leu Lys Lys Pro Pro Pro Asn Leu Phe
210 215 220

Lys Ile Ile Gln Lys Ala Leu Gln Lys Gly Ser Leu Leu Gly Cys Ser
225 230 235 240

Ile Asp Ile Thr Ser Ala Ala Asp Ser Glu Ala Ile Thr Phe Gln Lys
245 250 255

Leu Val Lys Gly His Ala Tyr Ser Val Thr Gly Ala Glu Glu Val Glu
260 265 270

Ser Asn Gly Ser Leu Gln Lys Leu Ile Arg Ile Arg Asn Pro Trp Gly
275 280 285

Glu Val Glu Trp Thr Gly Ala Trp Asn Asp Asn Cys Pro Ser Trp Asn
290 295 300

Thr Ile Asp Pro Glu Glu Arg Glu Arg Leu Thr Arg Arg His Glu Asp
305 310 315 320

Gly Glu Phe Trp Met Ser Phe Ser Asp Phe Leu Arg His Tyr Ser Arg
325 330 335

Leu Glu Ile Cys Asn Leu Thr Pro Asp Thr Leu Thr Ser Asp Thr Tyr
340 345 350

Lys Lys Trp Lys Leu Thr Lys Met Asp Gly Asn Trp Arg Arg Gly Ser

-continued

```
        355                 360                 365

Thr Ala Gly Gly Cys Arg Asn Tyr Pro Asn Thr Phe Trp Met Asn Pro
    370                 375                 380

Gln Tyr Leu Ile Lys Leu Glu Glu Glu Asp Glu Asp Glu Glu Asp Gly
385                 390                 395                 400

Glu Ser Gly Cys Thr Phe Leu Val Gly Leu Ile Gln Lys His Arg Arg
                405                 410                 415

Arg Gln Arg Lys Met Gly Glu Asp Met His Thr Ile Gly Phe Gly Ile
            420                 425                 430

Tyr Glu Val Pro Glu Glu Leu Ser Gly Gln Thr Asn Leu His Leu Ser
        435                 440                 445

Lys Asn Phe Phe Leu Thr Asn Arg Ala Arg Glu Arg Ser Asp Thr Phe
    450                 455                 460

Ile Asn Leu Arg Glu Val Leu Asn Arg Phe Lys Leu Pro Pro Gly Glu
465                 470                 475                 480

Tyr Ile Leu Val Pro Ser Thr Phe Glu Pro Asn Lys Asp Gly Asp Phe
                485                 490                 495

Cys Ile Arg Val Phe Ser Glu Lys Lys Ala Asp Tyr Gln Ala Val Asp
            500                 505                 510

Asp Glu Ile Glu Ala Asn Leu Glu Glu Phe Asp Ile Ser Glu Asp Asp
        515                 520                 525

Ile Asp Asp Gly Val Arg Arg Leu Phe Ala Gln Leu Ala Gly Glu Asp
    530                 535                 540

Ala Glu Ile Ser Ala Phe Glu Leu Gln Thr Ile Leu Arg Arg Val Leu
545                 550                 555                 560

Ala Lys Arg Gln Asp Ile Lys Ser Asp Gly Phe Ser Ile Glu Thr Cys
                565                 570                 575

Lys Ile Met Val Asp Met Leu Asp Ser Asp Gly Ser Gly Lys Leu Gly
            580                 585                 590

Leu Lys Glu Phe Tyr Ile Leu Trp Thr Lys Ile Gln Lys Tyr Gln Lys
        595                 600                 605

Ile Tyr Arg Glu Ile Asp Val Asp Arg Ser Gly Thr Met Asn Ser Tyr
    610                 615                 620

Glu Met Arg Lys Ala Leu Glu Glu Ala Gly Phe Lys Met Pro Cys Gln
625                 630                 635                 640

Leu His Gln Val Ile Val Ala Arg Phe Ala Asp Asp Gln Leu Ile Ile
                645                 650                 655

Asp Phe Asp Asn Phe Val Arg Cys Leu Val Arg Leu Glu Thr Leu Phe
            660                 665                 670

Lys Ile Phe Lys Gln Leu Asp Pro Glu Asn Thr Gly Thr Ile Glu Leu
        675                 680                 685

Asp Leu Ile Ser Trp Leu Cys Phe Ser Val Leu
    690                 695


<210> SEQ ID NO 20
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

Met Ala Ser Ser Ser Gly Arg Val Thr Ile Gln Leu
1               5                   10
```

-continued

<210> SEQ ID NO 21
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

Gln Leu Gly Pro Asp Ser Glu Lys Ala Lys Gly Val Lys
1 5 10

<210> SEQ ID NO 22
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 22

Gly Ala Thr Ala Leu Ser Asp Arg Gly Glu Tyr Arg Thr
1 5 10

<210> SEQ ID NO 23
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 23

Tyr Ser Ile Thr Gly Thr His Lys Val Phe Leu Gly Phe
1 5 10

<210> SEQ ID NO 24
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 24

Ala Arg Gly Gly Arg Thr Pro Lys Cys Thr Val Leu Leu
1 5 10

<210> SEQ ID NO 25
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 25

Leu Gly Leu Trp Asp Ser Pro Arg Ser His Ala Leu Leu
1 5 10

<210> SEQ ID NO 26
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 26

Asp Arg Ser Pro Leu Ser Ala Arg Arg Asp Val Thr Arg
1 5 10

<210> SEQ ID NO 27
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 27

Ala Arg Arg Asp Val Thr Arg Arg Cys Cys Leu Arg Pro
1 5 10

-continued

<210> SEQ ID NO 28
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 28

Asp Glu Ala Asp Phe Thr Leu Arg Val Phe Ser Glu Arg
1 5 10

<210> SEQ ID NO 29
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 29

Thr Leu Arg Val Phe Ser Glu Arg Arg His Thr Ala Val
1 5 10

<210> SEQ ID NO 30
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 30

Arg Ala His Thr Ser Thr Pro Arg Glu Ile Gly Leu Arg
1 5 10

<210> SEQ ID NO 31
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 31

Leu Thr Gln Thr Leu Thr Ser Arg Tyr Arg Asp Ser Arg
1 5 10

<210> SEQ ID NO 32
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 32

Gly Val Ile Cys Leu Thr His Arg Gln Trp Met Glu Val
1 5 10

<210> SEQ ID NO 33
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 33

Ile Cys Glu Asp Met Ser Arg Thr Asp Val Cys Gln Gly Ser
1 5 10

<210> SEQ ID NO 34
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 34

Pro Gln Phe Arg Leu Thr Leu Leu Glu Pro Asp Glu Glu Asp

-continued

1 5 10

<210> SEQ ID NO 35
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 35

Ser Glu Arg Arg His Thr Ala Val Glu Ile Asp Asp Val Ile
1 5 10

<210> SEQ ID NO 36
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 36

Arg Ala His Thr Ser Thr Pro Arg Glu Ile Gly Leu Arg Thr
1 5 10

<210> SEQ ID NO 37
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 37

Arg Val Phe Ser Glu Arg Arg His Thr Ala Val Glu Ile Asp
1 5 10

<210> SEQ ID NO 38
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 38

Glu Glu Glu Glu Leu Asn Ala Ser Gln Leu Gln Ala Leu Leu
1 5 10

<210> SEQ ID NO 39
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 39

Val Asp Glu Glu Ala Gly Val Gly Ala Gly Arg Leu Gln Leu Phe Arg
1 5 10 15

<210> SEQ ID NO 40
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 40

Thr Asp Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Phe Leu Ala Ala
1 5 10 15

<210> SEQ ID NO 41
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 41

-continued

Tyr Glu Val Met Arg Gly Gly His Met Asn Glu Ala Phe Val Asp Phe
1 5 10 15

<210> SEQ ID NO 42
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

Arg Gln Asn Ser Met Gly Leu Phe Ser Ala Leu Arg His Ala Leu Ala
1 5 10 15

<210> SEQ ID NO 43
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

Tyr Arg Thr Glu Glu Gly Leu Val Lys Gly His Ala Tyr Ser Ile Thr
1 5 10 15

<210> SEQ ID NO 44
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

Gly Phe Asn Ser Gly Gly Ser Gln Pro Asn Ala Glu Thr Phe Trp Thr
1 5 10 15

<210> SEQ ID NO 45
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

Glu Glu Gly Pro Trp Gly Gly Trp Gly Ala Ala Gly Ala Arg Gly Pro
1 5 10 15

<210> SEQ ID NO 46
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46

Pro Trp Gly Gly Trp Gly Ala Ala Gly Ala Arg Gly Pro Ala Arg Gly
1 5 10 15

<210> SEQ ID NO 47
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 47

Gln Ser Leu Gln Val Gly Thr Val Pro Gly Gly Ala Ala Trp Gly Gly
1 5 10 15

<210> SEQ ID NO 48
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

-continued

Gly Thr Val Pro Gly Gly Ala Ala Trp Gly Gly Asp Leu Gly Gln Gly
1 5 10 15

<210> SEQ ID NO 49
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

Gly Gly Ala Ala Trp Gly Gly Asp Leu Gly Gln Gly Pro Tyr Leu Pro
1 5 10 15

<210> SEQ ID NO 50
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

Thr Pro Arg Glu Ile Gly Leu Arg Thr Cys Glu Gln Leu Leu Gln Cys
1 5 10 15

<210> SEQ ID NO 51
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 51

Gln Cys Phe Gly His Gly Gln Ser Leu Ala Leu His His Phe Gln Gln
1 5 10 15

<210> SEQ ID NO 52
<211> LENGTH: 16
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 52

Asp Glu Asp Thr Ser Gly Thr Met Asn Ser Tyr Glu Leu Arg Leu Ala
1 5 10 15

<210> SEQ ID NO 53
<211> LENGTH: 23
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 53

Ile Phe Asn Lys Phe Asp Glu Asp Thr Ser Gly Thr Met Asn Ser Tyr
1 5 10 15

Glu Leu Arg Leu Ala Leu Asn
20

<210> SEQ ID NO 54
<211> LENGTH: 22
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54

Arg Thr Asp Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Phe Leu Ala
1 5 10 15

Ala Ala Ala Ser Leu Thr
20

-continued

<210> SEQ ID NO 55
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 55

gcacgtccac accttccaa 19

<210> SEQ ID NO 56
<211> LENGTH: 22
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 56

ttggtccaga aggtttcagc at 22

<210> SEQ ID NO 57
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57

tccggcggga gccagcct 18

<210> SEQ ID NO 58
<211> LENGTH: 10
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 58

Glu Pro Asp Glu Glu Asp Asp Glu Asp Glu
1 5 10

<210> SEQ ID NO 59
<211> LENGTH: 18
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 59

agatggcatc cagcagtg 18

<210> SEQ ID NO 60
<211> LENGTH: 19
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 60

tccggagatc ctaggagaa 19

<210> SEQ ID NO 61
<211> LENGTH: 39
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61

gcagcagcgg ccgcatgagc cgcacagacg tgtgtcagg 39

<210> SEQ ID NO 62
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 62

gcagcagtcg acggagaagg tggccacctc catccac 37

<210> SEQ ID NO 63
<211> LENGTH: 38
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 63

gcagcagcgg ccgcatggca tccagcagtg ggagggtc 38

<210> SEQ ID NO 64
<211> LENGTH: 37
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 64

gcagcagtcg acggcctgcc actccaggag gtagccc 37

<210> SEQ ID NO 65
<211> LENGTH: 245
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 65

Ser Arg Thr Asp Val Cys Gln Gly Ser Leu Gly Asn Cys Trp Phe Leu
1 5 10 15

Ala Ala Ala Ala Ser Leu Thr Leu Tyr Pro Arg Leu Leu Arg Arg Val
20 25 30

Val Pro Pro Gly Gln Asp Phe Gln His Gly Tyr Ala Gly Val Phe His
35 40 45

Phe Gln Leu Trp Gln Phe Gly Arg Trp Met Asp Val Val Val Asp Asp
50 55 60

Arg Leu Pro Val Arg Glu Gly Lys Leu Met Phe Val Arg Ser Glu Gln
65 70 75 80

Arg Asn Glu Phe Trp Ala Pro Leu Leu Glu Lys Ala Tyr Ala Lys Leu
85 90 95

His Gly Ser Tyr Glu Val Met Arg Gly Gly His Met Asn Glu Ala Phe
100 105 110

Val Asp Phe Thr Gly Gly Val Gly Glu Val Leu Tyr Leu Arg Gln Asn
115 120 125

Ser Met Gly Leu Phe Ser Ala Leu Arg His Ala Leu Ala Lys Glu Ser
130 135 140

Leu Val Gly Ala Thr Ala Leu Ser Asp Arg Gly Glu Tyr Arg Thr Glu
145 150 155 160

Glu Gly Leu Val Lys Gly His Ala Tyr Ser Ile Thr Gly Thr His Lys
165 170 175

Val Phe Leu Gly Phe Thr Lys Val Arg Leu Leu Arg Leu Arg Asn Pro
180 185 190

Trp Gly Cys Val Glu Trp Thr Gly Ala Trp Ser Asp Ser Cys Pro Arg
195 200 205

Trp Asp Thr Leu Pro Thr Glu Cys Arg Asp Ala Leu Leu Val Lys Lys
210 215 220

Glu Asp Gly Glu Phe Trp Met Glu Leu Arg Asp Phe Leu Leu His Phe

-continued

225 230 235 240

Asp Thr Val Gln Ile
245

<210> SEQ ID NO 66
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: bacteriophage T7

<400> SEQUENCE: 66

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> SEQ ID NO 67
<211> LENGTH: 733
<212> TYPE: DNA
<213> ORGANISM: homo sapiens

<400> SEQUENCE: 67

gggatccgga gcccaaatct tctgacaaaa ctcacacatg cccaccgtgc ccagcacctg 60

aattcgaggg tgcaccgtca gtcttcctct tccccccaaa acccaaggac accctcatga 120

tctcccggac tcctgaggtc acatgcgtgg tggtggacgt aagccacgaa gaccctgagg 180

tcaagttcaa ctggtacgtg gacggcgtgg aggtgcataa tgccaagaca aagccgcggg 240

aggagcagta caacagcacg taccgtgtgg tcagcgtcct caccgtcctg caccaggact 300

ggctgaatgg caaggagtac aagtgcaagg tctccaacaa agccctccca acccccatcg 360

agaaaaccat ctccaaagcc aaagggcagc cccgagaacc acaggtgtac accctgcccc 420

catcccggga tgagctgacc aagaaccagg tcagcctgac ctgcctggtc aaaggcttct 480

atccaagcga catcgccgtg gagtgggaga gcaatgggca gccggagaac aactacaaga 540

ccacgcctcc cgtgctggac tccgacggct ccttcttcct ctacagcaag ctcaccgtgg 600

acaagagcag gtggcagcag gggaacgtct tctcatgctc cgtgatgcat gaggctctgc 660

acaaccacta cacgcagaag agcctctccc tgtctccggg taaatgagtg cgacggccgc 720

gactctagag gat 733

<210> SEQ ID NO 68
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 68

caggtgcagc tggtgcagtc tgg 23

<210> SEQ ID NO 69
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69

caggtcaact taagggagtc tgg 23

<210> SEQ ID NO 70
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

-continued

<400> SEQUENCE: 70

gaggtgcagc tggtggagtc tgg 23

<210> SEQ ID NO 71
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 71

caggtgcagc tgcaggagtc ggg 23

<210> SEQ ID NO 72
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 72

gaggtgcagc tgttgcagtc tgc 23

<210> SEQ ID NO 73
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 73

caggtacagc tgcagcagtc agg 23

<210> SEQ ID NO 74
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 74

tgaggagacg gtgaccaggg tgcc 24

<210> SEQ ID NO 75
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 75

tgaagagacg gtgaccattg tccc 24

<210> SEQ ID NO 76
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 76

tgaggagacg gtgaccaggg ttcc 24

<210> SEQ ID NO 77
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 77

tgaggagacg gtgaccgtgg tccc 24

<210> SEQ ID NO 78

-continued

<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 78

gacatccaga tgacccagtc tcc 23

<210> SEQ ID NO 79
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 79

gatgttgtga tgactcagtc tcc 23

<210> SEQ ID NO 80
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 80

gatattgtga tgactcagtc tcc 23

<210> SEQ ID NO 81
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 81

gaaattgtgt tgacgcagtc tcc 23

<210> SEQ ID NO 82
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 82

gacatcgtga tgacccagtc tcc 23

<210> SEQ ID NO 83
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 83

gaaacgacac tcacgcagtc tcc 23

<210> SEQ ID NO 84
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 84

gaaattgtgc tgactcagtc tcc 23

<210> SEQ ID NO 85
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 85

-continued

cagtctgtgt tgacgcagcc gcc 23

<210> SEQ ID NO 86
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 86

cagtctgccc tgactcagcc tgc 23

<210> SEQ ID NO 87
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 87

tcctatgtgc tgactcagcc acc 23

<210> SEQ ID NO 88
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 88

tcttctgagc tgactcagga ccc 23

<210> SEQ ID NO 89
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 89

cacgttatac tgactcaacc gcc 23

<210> SEQ ID NO 90
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 90

caggctgtgc tcactcagcc gtc 23

<210> SEQ ID NO 91
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 91

aattttatgc tgactcagcc cca 23

<210> SEQ ID NO 92
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 92

acgtttgatt tccaccttgg tccc 24

<210> SEQ ID NO 93
<211> LENGTH: 24
<212> TYPE: DNA

-continued

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 93

acgtttgatc tccagcttgg tccc 24

<210> SEQ ID NO 94
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 94

acgtttgata tccactttgg tccc 24

<210> SEQ ID NO 95
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 95

acgtttgatc tccaccttgg tccc 24

<210> SEQ ID NO 96
<211> LENGTH: 24
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 96

acgtttaatc tccagtcgtg tccc 24

<210> SEQ ID NO 97
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 97

cagtctgtgt tgacgcagcc gcc 23

<210> SEQ ID NO 98
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 98

cagtctgccc tgactcagcc tgc 23

<210> SEQ ID NO 99
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 99

tcctatgtgc tgactcagcc acc 23

<210> SEQ ID NO 100
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 100

tcttctgagc tgactcagga ccc 23

-continued

<210> SEQ ID NO 101
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 101

cacgttatac tgactcaacc gcc 23

<210> SEQ ID NO 102
<211> LENGTH: 23
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 102

caggctgtgc tcactcagcc gtc 23

What is claimed is:

1. An isolated nucleic acid molecule comprising a polynucleotide having a nucleotide sequence selected from the group consisting of:

(a) a polynucleotide fragment of SEQ ID NO:1 or a polynucleotide fragment of the cDNA sequence included in ATCC Deposit No: PTA-3745, which is hybridizable to SEQ ID NO:1;

(b) a polynucleotide encoding a polypeptide fragment of SEQ ID NO:2 or a polypeptide fragment encoded by the cDNA sequence included in ATCC Deposit No: PTA-3745, which is hybridizable to SEQ ID NO:1;

(c) a polynucleotide encoding a polypeptide domain of SEQ ID NO:2 or a polypeptide domain encoded by the cDNA sequence included in ATCC Deposit No: PTA-3745, which is hybridizable to SEQ ID NO:1;

(d) a polynucleotide encoding a polypeptide epitope of SEQ ID NO:2 or a polypeptide epitope encoded by the cDNA sequence included in ATCC Deposit No: PTA-3745, which is hybridizable to SEQ ID NO:1;

(e) a polynucleotide encoding a polypeptide of SEQ ID NO:2 or the cDNA sequence included in ATCC Deposit No: PTA-3745, which is hybridizable to SEQ ID NO:1, having biological activity;

(f) an isolated polynucleotide comprising nucleotides 4 to 2207 of SEQ ID NO:1, wherein said nucleotides encode a polypeptide corresponding to amino acids 2 to 735 of SEQ ID NO:2 of SEQ ID NO:2 minus the start methionine;

(g) an isolated polynucleotide comprising nucleotides 1 to 2207 of SEQ ID NO:1, wherein said nucleotides encode a polypeptide corresponding to amino acids 1 to 735 of SEQ ID NO:2 including the start methionine;

(h) a polynucleotide which represents the complimentary sequence (antisense) of SEQ ID NO:1; and

(i) a polynucleotide capable of hybridizing under stringent conditions to any one of the polynucleotides specified in (a)-(h), wherein said polynucleotide does not hybridize under stringent conditions to a nucleic acid molecule having a nucleotide sequence of only A residues or of only T residues.

2. The isolated nucleic acid molecule of claim 1, wherein the polynucleotide fragment consists of a nucleotide sequence encoding a human cysteine protease.

3. A recombinant vector comprising the isolated nucleic acid molecule of claim 1.

4. A recombinant host cell comprising the vector sequences of claim 3.

5. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: PTA-3745;

(b) a polypeptide fragment of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: PTA-3745, having protease activity;

(c) a polypeptide domain of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: PTA-3745;

(d) a polypeptide epitope of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: PTA-3745;

(e) a full length protein of SEQ ID NO:2 or the encoded sequence included in ATCC Deposit No: PTA-3745;

(f) a polypeptide comprising amino acids 2 to 735 of SEQ ID NO:2, wherein said amino acids 2 to 735 comprising a polypeptide of SEQ ID NO:2 minus the start methionine; and

(g) a polypeptide comprising amino acids 1 to 735 of SEQ ID NO:2.

6. The isolated polypeptide of claim 5, wherein the full length protein comprises sequential amino acid deletions from either the C-terminus or the N-terminus.

7. An isolated antibody that binds specifically to the isolated polypeptide of claim 5.

8. A recombinant host cell that expresses the isolated polypeptide of claim 5.

9. A method of making an isolated polypeptide comprising:

(a) culturing the recombinant host cell of claim 8 under conditions such that said polypeptide is expressed; and

(b) recovering said polypeptide.

10. The polypeptide produced by claim 9.

11. A method for preventing, treating, or ameliorating a medical condition, comprising the step of administering to a mammalian subject a therapeutically effective amount of the polypeptide of claim 5, or a modulator thereof.

12. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or absence of a mutation in the polynucleotide of claim 1; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or absence of said mutation.

13. A method of diagnosing a pathological condition or a susceptibility to a pathological condition in a subject comprising:

(a) determining the presence or amount of expression of the polypeptide of claim 5 in a biological sample; and

(b) diagnosing a pathological condition or a susceptibility to a pathological condition based on the presence or amount of expression of the polypeptide.

14. An isolated nucleic acid molecule consisting of a polynucleotide having a nucleotide sequence selected from the group consisting of:

(a) a polynucleotide encoding a polypeptide of SEQ ID NO:2;

(b) an isolated polynucleotide consisting of nucleotides 4 to 2207 of SEQ ID NO:1, wherein said nucleotides encode a polypeptide corresponding to amino acids 2 to 735 of SEQ ID NO:2 minus the start methionine;

(c) an isolated polynucleotide consisting of nucleotides 1 to 2207 of SEQ ID NO:1, wherein said nucleotides encode a polypeptide corresponding to amino acids 1 to 735 of SEQ ID NO:2 including the start methionine;

(d) a polynucleotide encoding the Protease-42 polypeptide encoded by the cDNA clone contained in ATCC Deposit No. PTA-3745; and

(e) a polynucleotide which represents the complimentary sequence (antisense) of SEQ ID NO:1.

15. The isolated nucleic acid molecule of claim 14, wherein the polynucleotide comprises a nucleotide sequence encoding a human cysteine protease.

16. A recombinant vector comprising the isolated nucleic acid molecule of claim 15.

17. A recombinant host cell comprising the recombinant vector of claim 16.

18. An isolated polypeptide consisting of an amino acid sequence selected from the group consisting of:

(a) a polypeptide fragment of SEQ ID NO:2 having protease activity;

(b) a polypeptide domain of SEQ ID NO:2 having protease activity;

(c) a full length protein of SEQ ID NO:2;

(d) a polypeptide corresponding to amino acids 2 to 735 of SEQ ID NO:2, wherein said amino acids 2 to 735 consisting of a polypeptide of SEQ ID NO:2 minus the start methionine;

(e) a polypeptide corresponding to amino acids 1 to 735 of SEQ ID NO:2; and

(f) a polypeptide encoded by the cDNA contained in ATCC Deposit No. PTA-3745.

19. The method of diagnosing a pathological condition of claim 15 wherein the condition is a member of the group consisting of: a disorder related to aberrant calpain activity; a disorder associated with deficiencies in calpain activity; a disorder associated with hypercalpain activity; a disorder related to aberrant protease regulation; a disorder related to aberrant calcium regulation; a disorder related to aberrant cell cycle regulation; female reproductive tract disorders; infertility; carcinomas of the female reproductive tract; dysfunctional uterine bleeding, amenorrhea, primary dysmenorrhea, sexual dysfunction, infertility, pelvic inflammatory disease, endometriosis, placental aromatase deficiency, premature menopause, placental dysfunction, pelvic inflammatory disease, tubal pregnancy, and Chlamydial infection; neural disorders; hepatic disorders; immune disorders; hematopoietic disorders; renal disorders; pulmonary disorders; an inflammatory condition; an inflammatory disease wherein calpains, either directly or indirectly, are involved in disease progression; gastrointestinal disorders; colon disorders; proliferative disorder of the colon or gastrointestinal tissue; colon cancer; colon adenocarcinoma; ischemia-reperfusion injury; hearing disorders; hearing loss; multiple sclerosis; cataracts; and myocarditis.

20. The method for preventing, treating, or ameliorating a medical condition of claim 11, wherein the medical condition is selected from the group consisting of: a disorder related to aberrant calpain activity; a disorder associated with deficiencies in calpain activity; a disorder associated with hypercalpain activity; a disorder related to aberrant protease regulation; a disorder related to aberrant calcium regulation; a disorder related to aberrant cell cycle regulation; female reproductive tract disorders; infertility; carcinomas of the female reproductive tract; dysfunctional uterine bleeding, amenorrhea, primary dysmenorrhea, sexual dysfunction, infertility, pelvic inflammatory disease, endometriosis, placental aromatase deficiency, premature menopause, placental dysfunction, pelvic inflammatory disease, tubal pregnancy, and Chlamydial infection; neural disorders; hepatic disorders; immune disorders; hematopoietic disorders; renal disorders; pulmonary disorders; an inflammatory condition; an inflammatory disease wherein calpains, either directly or indirectly, are involved in disease progression; gastrointestinal disorders; colon disorders; proliferative disorder of the colon or gastrointestinal tissue; colon cancer; colon adenocarcinoma; ischemia-reperfusion injury; hearing disorders; hearing loss; multiple sclerosis; cataracts; and myocarditis.

21. A computer for producing a three-dimensional representation of a molecule or molecular complex, wherein said molecule or molecular complex comprises the structural coordinates of Protease-42 provided in **FIG. 6** in accordance with Table IV, wherein said computer comprises:

(a) A machine-readable data storage medium, comprising a data storage material encoded with machine readable data, wherein the data is defined by the set of structure coordinates of the model;

(b) a working memory for storing instructions for processing said machine-readable data;

(c) a central-processing unit coupled to said working memory and to said machine-readable data storage medium for processing said machine readable data into said three-dimensional representation; and

(d) a display coupled to said central-processing unit for displaying said three-dimensional representation.

22. A method for identifying a mutant with altered biological properties, function, or activity of Protease-42, wherein said method comprises the steps of:

(a) using a model of said polypeptide according to the structural coordinates of said model to identify amino acids to mutate; and

(b) mutating said amino acids to create a mutant protein with altered biological function or properties.

23. The method according to claim 22 wherein the mutant is a member of the group consisting of:

(a) a mutant with mutations in the active site domain of Protease-42 comprised of amino acids from about S93 to about I337 of SEQ ID NO:2 according to Table IV with altered calpain function or properties; and

(b) a mutant with mutations in catalytic amino acid residues within the Protease-42 active site comprised of amino acids from about C105, H259, and N283 of SEQ ID NO:2 according to Table IV with altered calpain function or properties.

24. A method for designing or selecting compounds as potential modulators of Protease-42, wherein said method comprises the steps of:

(a) identifying a structural or chemical feature of said member using the structural coordinates of said member; and

(b) rationally designing compounds that bind to said feature.

* * * * *