



(12) **Patent Application Publication**
Eaton et al.

(43) **Pub. Date:** **Sep. 25, 2003**

FIGURE 1

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATG
AGCAGGTCTGAAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGT
TTCAGCAAGGCCTCAGTTTTCCTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTTCATATTT
TCATACATTACTGCAGTAACACTCCACCATATAGACCCGGCTTTACCTTATATCAGTGACACTGG
TACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTAAATATTGCGGCAGTTTTATGCATTG
CTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAGAACGTTATCATCAA
TTAAACAAGGCTGGCCTTGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTCGTAAGTTC
CCAGAAAAACAACCCTTTTTGCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCAT
TATATATGTTTGTTCAGACCATCCTTTTCCTACCAAATGCAGCCCCAAAATCCATGGCAAACAAGTC
TTCTGGATCAGACTGTTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTC
ATCAGTTTTGCACAGTGGCAATTTTGGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGG
ACAAAGGTTATGTGCTTCACATGATCACTACTGCAGCAGAATGGTCTATGTCATTTTCCTTCTTT
GGTTTTTTCCTGACTTACATTTCGTGATTTTCAGAAAATTTCTTTACGGGTGGAAGCCAATTTACA
TGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAAACAATGAACGAACACGGCTACTTTCCA
GAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG
TTCACAGAAGTTGCTTATTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACT
GATGAATGCTGATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCAT
CAAGAAGACTATTAAAAACACCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACT
ATG

FIGURE 2

<subunit 1 of 1, 266 aa, 1 stop

<MW: 29766, pI: 8.39, NX(S/T): 0

MWWFQQGLSFLPSALVIWTSAAFIIFS YITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAV
LCIATIIYVRYKQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFG
MGSLYMFVQTIILSYQMOPKIHGKQVFWIRLLLVICGVSALSMLTCSSVLHSGNFGTDLEQKLHW
NPEDKGYVLHMITTAAEWSMSFSFFGFFLT YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTR
LLSRDI

Important features:

Type II transmembrane domain:

amino acids 13-33

Other Transmembrane domains:

amino acids 54-73, 94-113, 160-180, 122-141

N-myristoylation sites.

amino acids 57-63, 95-101, 99-105, 124-130, 183-189

FIGURE 3

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGACGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACC
 GTGTGAGGGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTC
 CTTACACTTCGCCATGAGTTTCCTCATCGACTCCAGCATCATGATTACCTCCAGATACTATTTTTTG
 GATTTGGGTGGCTTTTCTTCATGCGCCAATTGTTTAAAGACTATGAGATACGTCAGTATGTTGTACAG
 GTGATCTTCTCCGTGACGTTTGCATTTTCTTGACCATGTTTGAGCTCATCATCTTTGAAATCTTAGG
 AGTATTGAATAGCAGCTCCCGTTATTTTCACTGGAAAATGAACCTGTGTGTAATTCTGCTGATCCTGG
 TTTTCATGGTGCCTTTTTTACATTGGCTATTTTATTGTGAGCAATATCCGACTACTGCATAAACAACGA
 CTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGAACTAGGAGATCCCTTTCC
 CATTCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTGGTGTGATTGGAG
 TGA CTCTCATGGCTCTTCTTTCTGGATTTGGTGTGTCAACTGCCCATACACTTACATGTCTTACTTC
 CTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGAT
 CATAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACA
 AACCATCAGGTTTCTGGGAATGATAAAAAAGTGTTACCCTTCAGCATCAGGAAGTGAAAATCTTACT
 CTTATTCAACAGGAAGTGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCT
 ATATGCTACCAAGGAGAGAATAGAATACTCCAAAACCTTCAAGGGGAAATATTTTAATTTTCTTGGTT
 ACTTTTTCTCTATTTACTGTGTTTGGAAAATTTTCATGGCTACCATCAATATTTGTTTTGATCGAGTT
 GGGAAAACGGATCCTGTCACAAGAGGCATTGAGATCACTGTGAATTATCTGGGAATCCAATTTGATGT
 GAAGTTTTGGTCCCAACACATTTCTTCATTCTGTTGGAATAATCATCGTCACATCCATCAGAGGAT
 TGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCCTCCAATGTCATTGTCCTG
 CTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGAGTATGCCTTT
 AGAATACCGCACCATATACTCACTGAAGTCCCTTGAGAACTGCAGTTCAACTTCTATCACCGTTGGTTTG
 ATGTGATCTTCCTGGTCAGCGCTCTCTCTAGCATACTCTTCCTCTATTTGGCTCACAAACAGGCACCA
 GAGAAGCAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTTCAAATTTA
 GATATAAGAGGGGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGC
 ATTTTTTACCTTCATAGCATACTCCTTCCCCGTCAGGTGATACTATGACCATGAGTAGCATCAGCCAG
 AACATGAGAGGGGAGAACTAACTCAAGACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGG
 TGTAGAGGCGGAGAGGAGCCAAGAACTAAAGGTGAAAAATACACTGGAACCTCTGGGGCAAGACATGT
 CTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTAAGGTTACATGGAAAAGGTTATAGCTTTG
 CCTTGAGATTGACTCATTAAAATCAGAGACTGTAACAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCG
 ACTCTAGAGTCGACCTGCAGAAGCTTGCCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

FIGURE 4

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGV
LNSSSRYPFHWMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLEFSCLLWLTFMYFFWKLGBP
FPILSPKHGILSIEQLISRVGVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQ
TMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWGMIKSVTTSASGSENLTLIQQEVDALAEELSRQ
LFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVFDRVGKTDPVTRGIEI
TVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKSSNVIVLLLAQIMGMY
FVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSAALSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 5

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGTTCC
 AGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATCTGAGGT
 GTTTCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTGCTTCTCACT
 TCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCTTTCCATCCAGGT
 GTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGCAGCTGAATTTACAG
 AAGCTAAGGAGGCTGTAGGCTGCTGGGACTAAGTTTGGCCGGCAAGGACCAAGTTGAAACAGCC
 TTGAAAGCTAGCTTTGAAACTTGCAGCTATGGCTGGGTGGAGATGGATTCTGTTGCTCATCTCTAG
 GATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTGTCTGATTTGGAAGGTTCCAGTGA
 GCCGACAGTTTGCAGCCTATTGTTACAACCTCATCTGATACTTGGACTAACTCGTGCATTCCAGAA
 ATTATCACCACCAAAGATCCCATATTCAACACTCAAACGCAACACAAACAAGAAATTTATTGT
 CAGTGACAGTACCTACTCGGTGGCATCCCCCTTACTCTACAATACCTGCCCCCTACTACTCTCTC
 CTGCTCCAGCTTCCACTTCTATTCCACGGAGAAAAAATGATTTGTGTACAGAAAGTTTTTATG
 GAACTAGCACCATGTCTACAGAACTGAACCATTGTTGAAAATAAAGCAGCATTCAAGAATGA
 AGCTGCTGGGTTTGGAGGTGTCCCCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTG
 CAGCTGGTCTTGGATTTTGCTATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAAT
 CAGCAGAAGGAAATGATCGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAA
 TGAGGAATCAAAGAAAACCTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGC
 GATGCCTGGAAGCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTTCTTT
 CATGCTCCTTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCA
 CCCTTGGTTCTTAAGTGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAAT
 GCCCTTCTCCTTATTGTAACCCCTGTCTGGATCCTATCCTCCTACCTCCAAAGCTTCCCACGGCCT
 TTCTAGCCTGGCTATGTCTAATAATATCCCACTGGGAGAAAGGAGTTTTGCAAAGTGCAAGGAC
 CTAAACATCTCATCAGTATCCAGTGGTAAAAGGCCTCCTGGCTGTCTGAGGCTAGGTGGGTG
 AAAGCCAAGGAGTCACTGAGACCAAGGCTTTCTCTACTGATTCCGCAGCTCAGACCCTTTCTTCA
 GCTCTGAAAGAGAAACACGTATCCACCTGACATGTCTTCTGAGCCCGGTAAGAGCAAAAGAAT
 GGCAGAAAAGTTTAGCCCCCTGAAAGCCATGGAGATTCTCATAACTTGAGACCTAATCTCTGTAAA
 GCTAAAATAAAGAAATAGAACAAGGCTGAGGATACGACAGTACACTGTCAGCAGGGACTGTAAAC
 ACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGAATCACTGTTTAGAACACACACA
 CTTACTTTTTCTGGTCTCTACCACTGCTGATATTTTCTCTAGGAAATATACTTTTACAAGTAACA
 AAAATAAAAACTCTTATAAATTTCTATTTTTATCTGAGTTACAGAAATGATTACTAAGGAAGATT
 ACTCAGTAATTTGTTTTAAAAAGTAATAAAATTCAACAAACATTTGCTGAATAGCTACTATATGTC
 AAGTGCTGTGCAAGGTATTACACTCTGTAATTGAATATTATTCCTCAAAAAATTGCACATAGTAG
 AACGCTATCTGGGAAGCTATTTTTTTCAGTTTTGATATTTCTAGCTTATCTACTTCCAACTAAT
 TTTTATTTTTGCTGAGACTAATCTTATTCATTTTCTCTAATATGGCAACCATTATAACCTTAATT
 TATTATTAACATACCTAAGAAGTACATTGTTACCTCTATATACCAAAGCACATTTTAAAGTGCC
 ATTAACAAATGTATCACTAGCCCTCCTTTTTTCCAACAAGAAGGGACTGAGAGATGCAGAAATATT
 TGTGACAAAAAATTAAAGCATTTAGAAAACCTT

FIGURE 6

MARCFSLVLLLTISIWTRLLVQGSRLAEELSIQVSCRIMGITLVSKKANQQLNFTEAKEACRLLG
LSLAGKDQVETALKASFETCSYGWVGDFVVISRISPNPKCGKNGVGVLIWKVPVSRQFAAYCYN
SSDTWTNSCIPEIITTKDPIFNTQTATQTTEFIVSDSTYSVASPYSTIPAPTTTPAPASTSIPR
RKKLICVTEVFMETSTMSTETEPFVENKAAFKNEAAGFGGVPTALLVLALLFFGAAAGLGFCYVK
RYVKAFPFITKNKQOKEMIETKVVKEEKANDSNPNEESKKTDKNPEESKSPSKTTVRCLEAEV

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 235-254

N-glycosylation site.

amino acids 53-57, 130-134, 289-293

Casein kinase II phosphorylation site.

amino acids 145-149, 214-218

Tyrosine kinase phosphorylation site.

amino acids 79-88

N-myristoylation site.

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

FIGURE 7

CGCCGCGCTCCCGCACCCGCGGCCCCGCCACCGCGCCGCTCCCGCATCTGCACCCGAGCCCGGC
GGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGCCCCGAGCGCAACTCGGTCCAGTCGGGGCGG
CGGCTGCGGGCGCAGAGCGGAGATGCGAGCGGCTTGGGGCCACCCTGCTGTGCCTGCTGCTGGCGG
CGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTCAAGCCCGGCCCG
GCTCTCAGCTACCCGAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGAGGTTGAGGAACTGAT
GGAGGACACGCAGCACAAATTGCGCAGCGCGGTGGAAGAGATGGAGGCAGAAGAAGCTGCTGCTA
AAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCCAGCTATCACAATGAGACCAACACAGAC
ACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAATTCACAAGATAACCAACAACCAGAC
TGGACAAATGGTCTTTTCAGAGACAGTTATCACATCTGTGGGAGACGAAGAAGGCAGAAGGAGCC
ACGAGTGCATCATCGACGAGGACTGTGGGCCCAGCATGTACTGCCAGTTTGCCAGCTTCCAGTAC
ACCTGCCAGCCATGCCGGGGCCAGAGGATGCTCTGCACCCGGGACAGTGAAGTGTGTGGAGACCA
GCTGTGTGTCTGGGGTCACTGCACCAAATGGCCACCAGGGGCAGCAATGGGACCATCTGTGACA
ACCAGAGGGACTGCCAGCCGGGGCTGTGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGC
ACACCCCTGCCCGTGGAGGGCGAGCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCAC
CTGGGAGCTAGAGCCTGATGGAGCCTTGGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGC
CCACAGCCACAGCCTGGTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGG
GAGATCCTGCTGCCAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCG
CCAGGAGCTGGAGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTG
CCGCCGCTGCACTGCTGGGAGGGGAAGAGATTAGATCTGGACCAGGCTGTGGGTAGATGTGCAA
TAGAAATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTA
CATCTTCTTCCAGTAAGTTTCCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTAGCT
CCCCCAGGCTGTTCTCCAGGCTTCACAGTCTGGTGCTTGGGAGAGTCAGGCAGGGTTAACTGCA
GGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAGACAGCCG
TTTGTCTTACATGGCTTTGATAATTGTTTGAGGGGAGGAGATGGAAACAATGTGGAGTCTCCCTC
TGATTGGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAACATCAACCTGGCAAAAATG
CAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTGTGCCTTCAGCTGTTGC
AGATGAAATGTTCTGTTCACCCTGCATTACATGTGTTTATTCATCCAGCAGTGTGTGCTCAGCTCC
TACCTCTGTGCCAGGGCAGCATTTTCATATCCAAGATCAATTCCCTCTCTCAGCACAGCCTGGGG
AGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAGGCTCAGAGACTGCAAGCTGCTTGCC
CAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCATCTGGTTGTGACTCTAAGCTCAGTGCTCT
CTCCACTACCCACACACAGCCTTGGTGCCACCAAAAAGTGCTCCCCAAAAGGAAGGAGAATGGGAT
TTTTCTTGAGGCATGCACATCTGGAATTAAGGTCAAACCTAATTCACATCCCTCTAAAAGTAA
CTACTGTTAGGAACAGCAGTGTCTCACAGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGAT
ATTGACACTGTCCCTCTTTGGCAGTTGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCA
TACAGGTTAACCTGCAGAAACAGTACTTAGGTAATTGTAGGGCGAGGATTATAAATGAAATTTGC
AAAATCACTTAGCAGCAACTGAAGACAATTATCAACCACGTGGAGAAAATCAAACCGAGCAGGGC
TGTGTGAAACATGGTTGTAATATGCGACTGCGAACACTGAACCTCTACGCCACTCCACAAATGATG
TTTTCAGGTGTATGAGCTGTTGCCACCATGTATTATCCAGAGTTCTTAAAGTTTAAAGTTGCA
CATGATTGTATAAGCATGCTTTCTTTGAGTTTTAAATATGTATAAACATAAGTTGCATTTAGAA
ATCAAGCATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 8

MQRLGATLLCLLLAAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQHKL
RSAVEEMEAEAAAAKASSEVNLANLP SYHNETNTDTKVGNN TIHVHREIHKITNNQTGQMVFSE
TVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQQRMLCTRDSECCGDQLCVWGHC
TKMATRGSNGTICDNQRDCQPLCCAFQRGLLFPVCTPLPVEGELCHDPASRLDLITWELEPDG
ALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEVGSFMEEVRQEELEDLE
RSLTEEMALGEPAAAAAALLGGEI

Signal sequence:

amino acids 1-19

N-glycosylation site.

amino acids 96-100, 106-110, 121-125, 204-208

Casein kinase II phosphorylation site.

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,
327-331

N-myristoylation site.

amino acids 202-208, 217-223

Amidation site.

amino acids 140-144

FIGURE 9

CGGACGCGTGGGCGGACGCGTGGGGGCTGTGAGAAAGTGCCAATAAATACATCATGCAACCCAC
GGCCACCTTGTGAACTCCTCGTGCCAGGGCTGATGTGCGTCTTCCAGGGCTACTCATCCAAAG
GCCTAATCCAACGTTCTGTCTTCAATCTGCAAATCTATGGGGTCCCTGGGGCTCTTCTGGACCCTT
AACTGGGTACTGGCCCTGGGCCAATGCGTCCTCGCTGGAGCCTTTGCCTCCTTCTACTGGGCCTT
CCACAAGCCCCAGGACATCCCTACCTTCCCCCTTAATCTCTGCCTTCATCCGCACACTCCGTTACC
ACACTGGGTCATTGGCATTGGAGCCCTCATCCTGACCCTTGTGCAGATAGCCCGGGTCATCTTG
GAGTATATTGACCACAAGCTCAGAGGAGTGCAGAACCCTGTAGCCCGCTGCATCATGTGCTGTTT
CAAGTGCTGCCTCTGGTGTCTGGAAAAATTTATCAAGTTCCTAAACCGCAATGCATACATCATGA
TCGCCATCTACGGGAAGAATTTCTGTGTCTCAGCCAAAAATGCGTTCATGCTACTCATGCGAAAC
ATTGTCAGGGTGGTTCGTCCTGGACAAAGTCACAGACCTGCTGCTGTTCTTTGGGAAGCTGCTGGT
GGTCGGAGGCGTGGGGGTCCCTGTCCTTCTTTTTTTCTCCGGTCGCATCCCGGGGCTGGGTAAAG
ACTTTAAGAGCCCCCACCTCAACTATTACTGGCTGCCCATCATGACCTCCATCCTGGGGGCCTAT
GTCATCGCCAGCGGCTTCTTCAGCGTTTTTCGGCATGTGTGTGGACACGCTCTTCCTCTGCTTCCT
GGAAGACCTGGAGCGGAACAACGGCTCCCTGGACCGGCCCTACTACATGTCCAAGAGCCTTCTAA
AGATTCTGGGCAAGAAGAACGAGGCGCCCCCGGACAACAAGAAGAGGAAGAAGTGACAGCTCCGG
CCCTGATCCAGGACTGCACCCACCCCCACCGTCCAGCCATCCAACCTCACTTCGCCTTACAGGT
CTCCATTTTGTGGTAAAAAAGGTTTTAGGCCAGGCGCCGTGGCTCACGCCTGTAATCCAACACT
TTGAGAGGCTGAGGCGGGCGGATCACCTGAGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTG
AAACCTCCGTCTCTATTAAAAATACAAAAATTAGCCGAGAGTGGTGGCATGCACCTGTCATCCCA
GCTACTCGGGAGGCTGAGGCAGGAGAATCGCTTGAACCCGGGAGGCAGAGGTTGCAGTGAGCCGA
GATCGCGCCACTGCACTCCAACCTGGGTGACAGACTCTGTCTCCAAAACAAAACAAACAAACAA
AAGATTTTATTAAAGATATTTTGTTAACTC

FIGURE 10

RTRGRTRGGCEKVPINTSCNPHTAHLVNSSCPGLMCVFQGYSSKGLIQRSVFNLQIYGVLGLFWTL
NWVLALGQCVLAGAFASFYWAFHKPQDIPTFPLISAFIRTLRYHTGSLAFGALILTLVQIARVIL
EYIDHKLRGVQNPVARCIMCCFKCCLWCLEKFIKFLNRNAYIMIAIYGKNFCVSAKNAFMLLMRN
IVRVVVLDKVTDLLFFGKLLVVGGVGLSFFFFSGRIPGLGKDFKSPHLNYYWLPIMTSILGAY
VIASGFFSVFGMCVDTLFLCFLEDLERNNGSLDRPYMSKSLLKILGKKNEAPPDNKKRKK

Important features:

Transmembrane domains:

amino acids 57-80 (type II), 110-126, 215-231, 254-274

N-glycosylation sites.

amino acids 16-20, 27-31, 289-293

Hypothetical YBR002c family proteins.

amino acids 276-288

Ammonium transporters proteins.

amino acids 204-231

N-myristoylation sites.

amino acids 60-66, 78-84

Amidation site.

amino acids 306-310

FIGURE 11

GCCCCGCGCCCGGCGCCGGGCGCCCGAAGCCGGGAGCCACCGCCATGGGGGCCTGCCTGGGAGCCTGC
TCCCTGCTCAGCTGCGCGTCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCTGCTGCCCCGC
CAGCCGCAACTCCACCGTGAGCCGCCTCATCTTCACGTTCTTCCTCTTCCTGGGGGTGCTGGTGTCCA
TCATTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCTGGGTGTGTGAGGAGGGGGCC
GGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCCTGCTTGGCTACCGCGCTGTCTACCG
CATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTTTACCCTGCTCATGCTCTGCGTGAGCAGCA
GCCGGGACCCCCGGGCTGCCATCCAGAATGGGTTTTGGTTCTTTAAGTTCCTGATCCTGGTGGGCCTC
ACCGTGGGTGCCTTCTACATCCCTGACGGCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGG
CTCCTTCTCTTCATCCTCATCCAGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGT
GGCTGGGCAAGGCCGAGGAGTGCGATTCCCGTGCCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTC
TTCTACTTGCTGTGATCGCGGCCGTGGCGCTGATGTTTCATGTACTACACTGAGCCCAGCGGCTGCCA
CGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTCGC
CCAAGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTTCATCACCTCTACACCATG
TTTGTCACTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTGCCAACCCAGCT
GGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATGCCCCGAGCATTG
TGGGCCTCATCATCTTCCTCCTGTGCACCCTCTTCATCAGTCTGCGCTCCTCAGACCACCGGCAGGTG
AACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGT
GGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACGGCGTCACCTACAGCTACTCCTTCTTCC
ACTTCTGCCTGGTGTGCTGGCCTCACTGCACGTGATGATGACGCTCACCAACTGGTACAAGCCCCGGTGAG
ACCCGGAAGATGATCAGCACGTGGACCGCCGTGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCT
CCTCTACCTGTGGACCCTGGTAGCCCCACTCCTCCTGCGCAACCGCGACTTCAGCTTGAGGCAGCCTCA
CAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTC
CCCACACCAATCAGCCAGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCCTGAGCCGGGC
CTTCTAGTCGTAGTGCCCTTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCAC
ACCCACACGGTGGAGCTGCCTCTTCCCTTCCCCCTCCTCCCTGTTGCCATACTCAGCATCTCGGATGAA
AGGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACCTCCCACCACAG
TGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGACCTGCCCCCTTCTCTG
GACTTCGTGCCTTACTGAGTCTCTAAGACTTTTTCTAATAAACAAGCCAGTGCGTGTAACAAAAA

FIGURE 12

MGACLGACSLSCASCLCGSAPCILCSCCPASRNSTVSRIFTFFFLGVLVSIIMLSPGVESQL
YKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFTLLMLCVSSSRDPRAAIQ
NGFWFFKFLILVGLTVGAFYIPDGSTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQRWLKAE
ECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVSIAAVLPKV
QDAQPN SGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGYETQWW DAPSI
VGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRAFDNEQDGV TYSY
SFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVVWKICASWAGLLLLYLWTLVAPLLLRNRD
FS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 13

CGGGCCAGCCTGGGGCGGGCCGGCCAGGAACCACCCGTTAAGGTGTCTTCTCTTTAGGGATGGTGA
GGTTGGAAAAAGACTCCTGTAAACCCTCCTCCAGGATGAACCACCTGCCAGAAGACATGGAGAACG
CTCTCACCGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCACACAA
CTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTCAGGAGGAC
TTTCTGTTTGTGTGTACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATAGAGTTAAATG
TGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTACTATTCTTCATAT
TTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCATATGCTGTGTGCAG
ACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCCTTTTTACTAGCAAAAG
TGATCCTTTCGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTGCCCATCATTTCAATTCATC
CTTGCCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCTCAAGAAGCAGAAGAAGAAAA
CAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGCAGCACTTATACCTGGTGGTCTTTCTGATG
GTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAAGAAGCTGAAGAAAAACAGGACAGT
GAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTAAATGTGAAAAACCTCACAGAAAGTC
ATCGAGGCCAAAAGAGGCAGGCAGTGGAGTCTCCCTGTGACAGTAAAGTTGAAATGGTGACGTC
CACTGCTGGCTTTATTGAACAGCTAATAAAGATTTATTTATTGTAATACCTCACAAACGTTGTAC
CATATCCATGCACATTTAGTTGCCTGCCTGTGGCTGGTAAGGTAATGTCATGATTCATCCTCTCT
TCAGTGAGACTGAGCCTGATGTGTTAACAAATAGGTGAAGAAAGTCTTGTGCTGTATTCCTAATC
AAAAGACTTAATATATTGAAGTAACACTTTTTTAGTAAGCAAGATACCTTTTTATTTCAATTCAC
AGAATGGAATTTTTTTGTTCATGTCTCAGATTTATTTGTATTTCTTTTTTAACACTCTACATT
TCCCTTGTTTTTTAACTCATGCACATGTGCTCTTTGTACAGTTTTAAAAAGTGTAATAAAATCTG
ACATGTCAATGTGGCTAGTTTTATTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGA
CTTGCAAAAGGGGAAGAAAGGAATTGCGAATACATGTAAAATGTCACCAGACATTTGTATTATTT
TTATCATGAAATCATGTTTTTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGC
ACAAAATGACTTAAACCATTCATATCATGTTTCCTTTGCGTTCAGCCAATTTCAATTAAATGAA
CTAAATTAAAAA

FIGURE 14

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDLLE
VTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAIALTT
AVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLPOEAEENRLLIVQDASER
AALIPGGLSDGQFYSPPESEAGSEEAEEKQDSEKPLLEL

Important features of the protein:

Signal peptide:

amino acids 1-20

Transmembrane domains:

amino acids 54-72, 100-118, 130-144, 146-166

N-myristoylation sites.

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

FIGURE 15

ACTCGAACGCAGTTGCTTCGGGACCCAGGACCCCCCTCGGGCCCCGACCCGCCAGGAAAGACTGAGG
CCGCGGCCTGCCCCGCCCGGCTCCCTGCGCCGCCCGCCGCCTCCCGGGACAGAAGATGTGCTCCAG
GGTCCCTCTGCTGCTGCCGCTGCTCCTGCTACTGGCCCTGGGGCCTGGGGTGCAGGGCTGCCCCAT
CCGGCTGCCAGTGACGCCAGCCACAGACAGTCTTCTGCACTGCCCCGCCAGGGGACCACGGTGCCC
CGAGACGTGCCACCCGACACGGTGGGGCTGTACGTCTTTGAGAACGGCATCACCATGCTCGACGC
AGGCAGCTTTGCCGGCCTGCCGGGCTGCAGCTCCTGGACCTGTACAGAACCCAGATCGCCAGCC
TGCCCAGCGGGGTCTTCCAGCCACTCGCCAACCTCAGCAACCTGGACCTGACGGCCAACAGGCTG
CATGAAATCACCAATGAGACCTTCCGTGGCCTGCGGCGCCTCGAGCGCCTCTACCTGGGCAAGAA
CCGCATCCGCCACATCCAGCCTGGTGCCCTCGACACGCTCGACCGCCTCCTGGAGCTCAAGCTGC
AGGACAACGAGCTGCGGGCACTGCCCCCGCTGCGCCTGCCCCGCTGCTGCTGCTGGACCTCAGC
CACAACAGCCTCCTGGCCCTGGAGCCCGGCATCCTGGACACTGCCAACGTGGAGGCGCTGCGGCT
GGCTGGTCTGGGGCTGCAGCAGCTGGACGAGGGGCTCTTCAGCCGCTTGCGCAACCTCCACGACC
TGGATGTGTCCGACAACCAGCTGGAGCGAGTGCCACCTGTGATCCGAGGCCTCCGGGGCCTGACG
CGCCTGCGGCTGGCCGGCAACACCCGCATTGCCAGCTGCGGCCCGAGGACCTGGCCGGCCTGGC
TGCCCTGCAGGAGCTGGATGTGAGCAACCTAAGCCTGCAGGCCCTGCCTGGCGACCTCTCGGGCC
TCTTCCCCCGCCTGCGGCTGCTGGCAGCTGCCCGCAACCCCTTCAACTGCGTGTGCCCCCTGAGC
TGGTTTGGCCCCCTGGGTGCGCGAGAGCCACGTCACTGGCCAGCCCTGAGGAGACGCGCTGCCA
CTTCCCGCCCAAGAACGCTGGCCGGCTGCTCCTGGAGCTTGACTACGCCGACTTTGGCTGCCAG
CCACCACCACACAGCCACAGTGCCACACAGAGGCCCGTGGTGCGGGAGCCACAGCCTGTCT
TCTAGCTTGGCTCCTACCTGGCTTAGCCCCACAGCGCCGGCCACTGAGGCCCCCAGCCCGCCCTC
CACTGCCCCACCGACTGTAGGGCTGTCCCCCAGCCCCAGGACTGCCACCGTCCACCTGCCTCA
ATGGGGGCACATGCCACCTGGGGACACGGCACCACCTGGCGTGCTTGTGCCCCGAAGGCTTCAG
GGCCTGTACTGTGAGAGCCAGATGGGGCAGGGGACACGGCCAGCCCTACACCAGTCACGCCGAG
GCCACCACGGTCCCTGACCCTGGGCATCGAGCCGGTGAGCCCCACCTCCCTGCGCGTGGGGCTGC
AGCGCTACCTCCAGGGGAGCTCCGTGCAGCTCAGGAGCCTCCGTCTCACCTATCGCAACCTATCG
GGCCCTGATAAGCGGCTGGTGACGCTGCGACTGCCTGCCTCGCTCGCTGAGTACAGGTCACCCA
GCTGCGGGCCCAACGCCACTTACTCCGTCTGTGTATGCCTTTGGGGCCCCGGGCGGGTGCCGGAGG
GCGAGGAGGCTGCGGGGAGGCCATACACCCCCAGCCGTCCACTCCAACCACGCCCCAGTCACC
CAGGCCCGCGAGGGCAACCTGCCGCTCCTCATTTGCGCCCGCCCTGGCCGCGGTGCTCCTGGCCGC
GCTGGCTGCGGTGGGGGCAGCCTACTGTGTGCGGCGGGGGCGGGCCATGGCAGCAGCGGCTCAGG
ACAAAGGGCAGGTGGGGCCAGGGGCTGGGCCCCTGGAAGTGGAGGGAGTGAAGGTCCCCTTGGAG
CCAGGCCCGAAGGCAACAGAGGGCGGTGGAGAGGCCCTGCCAGCGGGTCTGAGTGTGAGGTGCC
ACTCATGGGCTTCCCAGGGCCTGGCCTCCAGTCACCCCTCCACGCAAAGCCCTACATCTAAGCCA
GAGAGAGACAGGGCAGCTGGGGCCGGGCTCTCAGCCAGTGAGATGGCCAGCCCCCTCCTGCTGCC
ACACCACGTAAGTTCTCAGTCCCAACCTCGGGGATGTGTGCAGACAGGGCTGTGTGACCACAGCT
GGGCCCTGTTCCCTCTGGACCTCGGTCTCCTCATCTGTGAGATGCTGTGGCCAGCTGACGAGCC
CTAACGTCCCCAGAACCAGTGCTATGAGGACAGTGTCGCCCCCTGCCCTCCGCAACGTGCAGTC
CCTGGGCACGGCGGGCCCTGCCATGTGCTGGTAACGCATGCCTGGGTCTGCTGGGCTCTCCAC
TCCAGGCGGACCTGGGGGCCAGTGAAGGAAGCTCCCGGAAAGAGCAGAGGGAGAGCGGGTAGGC
GGCTGTGTGACTCTAGTCTTGGCCCCAGGAAGCGAAGGAACAAAAGAACTGGAAAGGAAGATGC
TTTAGGAACATGTTTTGCTTTTTTAAATATATATATTTATAAGAGATCCTTTCCCATTTATTCT
GGGAAGATGTTTTTCAAACCTCAGAGACAAGGACTTTGGTTTTTTGTAAGACAAACGATGATATGAA
GGCCTTTTGTAAAGAAAAATAAAAGATGAAGTGTGAAA

FIGURE 16

MCSRVP L L L L L L L L L L A L G P G V Q G C P S G C Q C S Q P Q T V F C T A R Q G T T V P R D V P P D T V G L Y V F E N G I T
M L D A G S F A G L P G L Q L L D L S Q N Q I A S L P S G V F Q P L A N L S N L D L T A N R L H E I T N E T F R G L R R L E R L Y
L G K N R I R H I Q P G A F D T L D R L L E L K L Q D N E L R A L P P L R L P R L L L L D L S H N S L L A L E P G I L D T A N V E
A L R L A G L G L Q L D E G L F S R L R N L H D L D V S D N Q L E R V P P V I R G L R G L T R L R L A G N T R I A Q L R P E D L
A G L A A L Q E L D V S N L S L Q A L P G D L S G L F P R L R L L A A A R N P F N C V C P L S W F G P W V R E S H V T L A S P E E
T R C H F P P K N A G R L L L E L D Y A D F G C P A T T T T A T V P T T R P V V R E P T A L S S S L A P T W L S P T A P A T E A P
S P P S T A P P T V G P V P Q P Q D C P P S T C L N G G T C H L G T R H H L A C L C P E G F T G L Y C E S Q M G Q G T R P S P T P
V T P R P P R S L T L G I E P V S P T S L R V G L Q R Y L Q G S S V Q L R S L R L T Y R N L S G P D K R L V T L R L P A S L A E Y
T V T Q L R P N A T Y S V C V M P L G P G R V P E G E E A C G E A H T P P A V H S N H A P V T Q A R E G N L P L L I A P A L A A V
L L A A L A A V G A A Y C V R R G R A M A A A A Q D K G Q V G P G A G P L E L E G V K V P L E P G P K A T E G G G E A L P S G S E
C E V P L M G F P G P G L Q S P L H A K P Y I

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 579-599

EGF-like domain cysteine pattern signature.

amino acids 430-442

Leucine zipper pattern.

amino acids 197-219, 269-291

N-glycosylation sites.

amino acids 101-105, 117-121, 273-277, 500-504, 528-532

Tyrosine kinase phosphorylation sites.

amino acids 124-131, 337-345

N-myristoylation sites.

amino acids 23-29, 27-33, 70-76, 142-148, 187-193, 348-354,
594-600, 640-646

FIGURE 17

[illegible]

FIGURE 18

MRVRIGLTLLLCVLLSLASASSDEEGSQDESLDSTTLTSDSVKDHTTAGRVVAGQIFLDSESESEL
ESSIQEEEDSLKSQEGESVTEDISFLESPNPENKDYEEPKKVRKPALTAIEGTAHGEPCHFPFLFLDK
EYDECTSDGREDGRLWCATTYDYKADEKWGFCETEEEEAAKRRQMQEAEEMMYQTGMKILNGSNKKSQKR
EAYRYLQKAASMNHTKALERSYALLFGDYLPQNIQAAREMFEKLTEEGSPKGQTALGFLYASGLGVN
SSQAKALVYYTTFGALGGNLIAMVLVSRL

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 195-199, 217-221, 272-276

Tyrosine kinase phosphorylation site.

amino acids 220-228

N-myristoylation sites.

amino acids 120-126, 253-259, 268-274, 270-274, 285-291, 289-295

Glycosaminoglycan attachment site.

amino acids 267-271

Microbodies C-terminal targeting signal.

amino acids 299-303

Type II fibronectin collagen-binding domain protein.

amino acids 127-169

Fructose-bisphosphate aldolase class-II protein.

amino acids 101-119

FIGURE 19

AATTCAGATTTTAAAGCCCATTCTGCAGTGGAATTTTCATGAACTAGCAAGAGGACACCATCTTCTT
GTATTATACAAGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGG
CCTCCTAATCCTCTGTGGTTTTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTG
ATAAGTACATTTTTATCACTGGATGTGACTCGGGCTTTGGAACTTGGCAGCCAGAACCTTTTGAT
AAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAATCAGGATCAACAGCTTTAAAGGCAGA
AACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGAATGTCAAGAGGACTG
CCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTGGTGT
CCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCCTATTGAAGTGAA
CCTGTTTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAG
TTATTAATGTCTCCAGTGTGGAGGTGCGCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAA
TATGCAGTGAAGGTTTCAATGACAGCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTC
ATGCATTGAACCAGGATTGTTCAAAACAACTTGGCAGATCCAGTAAAGGTAATTGAAAAAAAC
TCGCCATTTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATGGAGAAGGTTACATTGAAAAA
AGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGGTGGTAGAGTG
CATGGACCAGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAGATGCCAAAA
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAA
GCAGAGCTGGCTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGA
AATTGGCCGATTTCAAGAACACATCTCCTTTTCAACCCCATTCCTTATCTGCTCCAACCTGGACT
CATTTAGATCGTGCTTATTTGGATTGCAAAAGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGT
CCCTGCTCAAGTTTTCTTTGAAAAGGAGGGCTGGAATGGTACATCACATAGGCAAGTCTGCCCT
GTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTTGCCCATTCAAAATGA
TCTTTACCGTGGCTGCCCATGCTTATGGTCCCCAGCATTTACAGTAACTGTGAATGTTAAGT
ATCATCTCTTATCTAAATATTAAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAA

FIGURE 20

MLFWVLGLLILCGFLWTRKGKLIKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESG
STALKAETSERLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDY
REPIEVNLFGLISVTLNMLPLVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDSLRRDMK
AFGVHVSCIEPGLFKTNLADPVKVKIEKKLAIWEQLSPDIKQQYGEGYIEKSLDKLKGNKSYVNMD
LSPVVECMDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAALQDFLLLKQKAELANPKAV

Important features of the protein:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 136-152

N-glycosylation sites.

amino acids 161-163, 187-190 and 253-256

Glycosaminoglycan attachment site.

amino acids 39-42

N-myristoylation sites.

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

FIGURE 21

CTGAGGCGGCGGTAGCATGGAGGGGGAGAGTACGTCGGCGGTGCTCTCGGGCTTTGTGCTCGGCG
 CACTCGCTTTCCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTTGGGGAAGTAAAA
 GGTGAAGCCAAGAACAGCATTACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGA
 CATTTCAGAAATATATTCCATGCTATCAGCTTTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATG
 AGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAATGTGGTAGGTTGGTACAAATTCGGT
 CGTCATTTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAACCTTGCAGGAGCATTT
 TTCAAACCAAGACCTTGTTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTCTACTC
 ATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTT
 GCCAATCTGGGCATGTCTGAACAACTGGGTTATAAACTGTATFCAGGTTCTGTATGTCCACTGG
 TTTTAGCCGAGCAGTACAAACACACAGCTCTAAATTTTTTGAAGAAGATGGATCCTTAAAGGAGG
 TACATAAGATAAAATGAAATGTATGCTTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTG
 GAAGACAGTGAACAAGCAGTAGATAAACTAGTAAAGGATGTAAACAGATTAAAACGAGAAATTGA
 GAAAAGGAGAGGAGCACAGATTCAGGCAGCAAGAGAGAAGAACATCCAAAAGACCCTCAGGAGA
 ACATTTTTCTTTGTTCAGGCATTACGGACCTTTTTTCCAAATCTGAATTTCTTCATTTCATGTGTT
 ATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAAC TACAACCACCATCTCGATGTAGT
 AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCAC
 AAATCATTAAAGCATAAAGCCTTAGACTTAGATGACAGATGGCAATTCAAGAGATCTCGGTTGTTA
 GATACACAAGACAAACGATCTAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAAT
 GAGCAGCCCAGAAACAGATGAAGAAATTGAAAAGATGAAGGGTTTTGGTGAATATTCACGGTCTC
 CTACATTTTTGATTCCTTTTAACTTACAAGGAGATTTTTTTATTTGGCTGATGGGTAAAGCCAAAC
 ATTTCTATTGTTTTTACTATGTTGAGCTACTTGCAGTAAGTTCATTTGTTTTTACTATGTTTACC
 TGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAAACATCA
 GATGCTTTTATTTCCAAACCTTTTTTTTACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAG
 ACACATTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACT
 TAGGGAAGACAAGTCAGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATT
 GAGACCATGTCTATTAAAAAATAAAATGGAAAAGCAAGAATAGCCTTATTTTCAAATATGGAAA
 GAAATTTATATGAAAATTTATCTGAGTCATTAAAATTCCTTAAGTGATACTTTTTTAGAAGTA
 CATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAATAAATTTGCAAACATCATCT
 AAAATTTAAAAAAAAAAAAAAAAAAAAA

FIGURE 22

MEGESTSAVLSGFVLGALAFQHLNTSDTEGFLLGGEVKGEAKNSITDSQMDDVEVVYTIDIQYI
PCYQLFSFYNSSGEVNEQALKKILSNVKKNVVGWYKFRRHSDQIMTFRERLLHKNLQEHFSNQDL
VLLLLTPSIITESCSTHRLEHSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSRAV
QTHSSKFFEEEDGSLKEVHKINEMYASLQEELKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGA
QIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMSLKNRHVSKSSCNYNHHLDVVDNLTL
MVEHTDIPEASPASTPQIIKHKALDLDLRWQFKRSRLDLDQDKRSKANTGSSNQDKASKMSSPET
DEEIEKMKGFGEYSRSPTE

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 75-79, 322-326

N-myristoylation site.

amino acids 184-154

Growth factor and cytokines receptors family.

amino acids 134-150

FIGURE 23

GGCACAGCCGCGCGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCCCAA
 GCAGCGCGCAGCGAACGCCCGCCGCCGCCACACCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCCTTCCCC
 GCGTCCCCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCCGAGGTCACCAGCCCGCGCCTCT
 GCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCGGCGCCTGGCACCGGGGACCCTTGCCTGA
 CGCGAGGCCAGCTCTACTTTTCCGCCCGCGTCTCCTCCGCTGCTGCGCTCTTCCACCAACTCCAACCTCTTCCCC
 TCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCCGCTTCCCGTCCGGTCCCAA
 GGTGGGAACCGCTCCGCCCGCGCCCGCACCATGGCACGGTTGCGCTTGCCCGCGCTTCTCTGCACCCTGGCAGTGCTC
 AGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCACGCTCTTTACGTGTCCAAAGGCTTC
 AACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTTGAAGATCTGTCCCAGGGTTCTACCTGCTGCTCT
 CAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAAAGTGTGGTCAGCGAACAGTGCAATCATTTG
 CAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTTCTCAAAGAACTACTTGAAAATGCAGAGAAATCCCTG
 AATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAAATTTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTG
 AAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAAATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATG
 TTCGCCCTGGTGAACCTCCAGTACCCTTTACAGATGAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAG
 CCCTTCGGAGATGTCCCTCGCAAATGAAGCTCCAGGTTACTCGTGCTTTTGTAGCAGCCCGTACTTTGCTCAAGGC
 TTAGCGGTTGCGGGAGATGTCTGTAGCAAGGTCCTCGTGTAACCCACAGCCAGTGTACCCATGCCCTGTTGAAG
 ATGATCTACTGCTCCCACTGCCGGGCTCTCGTACTGTGAAGCCATGTTACAACCTACTGCTCAAACATCATGAGAGGC
 TGTTTGGCCAACCAAGGGGATCTCGATTTTGAATGGAACAATTTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTA
 GAGGGTCCCTTTCAACATTGAATCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGAT
 AATAGTGTTCAGTGTCTCAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGT
 TCCATCTCTGAAAGTGCCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACT
 AGTTTGGACCGACTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATCTGGTCTCCCTTCCGAGCAAC
 GTTTGCAACGATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTAC
 CTGTTTGCAGTGACAGGAAATGATTAGCCAACCAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACCAGAC
 ATACTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG
 GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAGTGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCAGAG
 TTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCTGGGGCA
 CAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGGAGATTAATTCTCAAACCTCTGAG
 AAAAGTGTTCATCAAAAAGTTAAAGGCACCAAGTTATCACTTTTCTACCATCCTAGTGACTTTGCTTTTTTAAATGAA
 TGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAAGTGTGACTTTGTTTTCTCATTAGTTTTGGG
 AGGAAAAGGGACTGTGCATTGAGTTGGTTCCTGCTCCCCCAAACCATGTTAAACGTGGCTAACAGTGTAGGTACAGAA
 CTATAGTTAGTTGTGCATTTGTGATTTTATCACTCTATTATTTGTTGTATGTTTTTTCTCATTTCTGTTTGTGGGT
 TTTTTTCCAACGTGTGATCTCGCCTTGTTCCTTACAAGCAAACCAAGGGTCCCTTCTTGGCACGTAACATGTACGTATT
 TCTGAAATATTAAATAGCTGTACAGAAGCAGTTTTATTTATCATGTTATCTTATTAAAGAAAAAGCCCAAAAGC

FIGURE 24

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRRLYVSKGFNKNDAPLHEINGDHLKICPQGST
CCSQEMEELKYSLSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENAEKSLNDMFVKTYGH
LYMQNSELFKDLFVELKRYVVGVNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTE
QLKPFGLDVPRLKLQVTRAFVAARTFAQGLAVAGDVVSKVSVVNPTAQCTHALLKMIYCSHCRGL
VTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMVAERLEGPFNIESVMDPIDVKISDAIMN
MQDNSVQVSQKVFQCGPPKPLPAGRISRISSESASFARFRPHHPEERPTTAAGTSLDRLVTDVK
EKLKQAKKEWSSLPSNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEVQVDTS
KPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAG
KSANEKADSAGVRPGAQAYLLTVFCILFLVMQREWR

Important features:

Signal peptide:

amino acids 1-22

ATP/GTP-binding site motif A (P-loop).

amino acids 515-524

N-glycosylation site.

amino acids 514-518

Glycosaminoglycan attachment sites.

amino acids 494-498, 498-502

N-myristoylation sites.

amino acids 63-69, 224-230, 276-282, 438-444, 497-503, 531-537

Glypicans proteins.

amino acids 54-75, 105-157, 238-280, 309-346, 423-460, 468-506

FIGURE 25

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGAC
CTGAGTCATCCCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACT
TACAGCTGCACCGACAGTTGCGATGAAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAA
TGCTGATGTCCATGGTCTCTAGCAGCCTGAATCCAGGGGTCGCCAGAGGCCACAGGGACCGAGGC
CAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAATGTGAGTGCAAAGATTGGTTCCTGAG
AGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGTGCCCCCTGTGATCATT
TCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATTCCAGA
GCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTGTAGGAGCTCTG
AGCGCCCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTC
TTCTTCTCCCACTCACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCA
TGTTTTTCAAGATCATTTTGTGTTGCTCTCTCTAGTGTCTTCTTCTCTCGTCAGTCTTAGCCT
GTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTACCTGAAAGATTCCAGGAAACTGTAGCTTCCT
AGCTAGTGTCATTTAACTTAAATGCAATCAGGAAAGTAGCAAACAGAAGTCAATAAATATTTTT
AAATGTCAAAAAAAAAAAAAAAAAA

FIGURE 26

MKVLISLLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFM
TVSGLPKKQCPDHFKNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL

Important features:

Signal peptide:

amino acids 1-22

N-myristoylation sites.

amino acids 27-33, 46-52

FIGURE 27

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAG
AGCTGGTCTGCCATGGACATCCTGGTCCCACCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCC
CCTGCACCTCATGGCTCTGCTGGGCTGCTGGCAGCCCCTGTGCAAAAGCTACTTCCCCTACCTGA
TGGCCGTGCTGACTCCCAAGAGCAACCGCAAGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAG
ATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTACTGGAGCTGGGCTGCGGAACCGGAGC
CAACTTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAAATCCCCACTTTGAGA
AGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGGCTCCT
GGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTG
CTCTGTGCAGAGCCCAAGGAAGGTCTGCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGC
TCTTTTTCTGGGAGCATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTTC
GAGCCACCTGGAAACACATTGGGGATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGA
GAACGCCCAGTTCTCCGAAATCCAATGGAACGACAGCCCCCTCCCTTGAAGTGGCTACCTGTTG
GGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCTCCAAGGCACTCATTTGCTCC
TTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGAGAGGGACCTTA
GCAGAATGAGAGAAGACATTCATGTACCACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGG
AAACACTAGGACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCAACGTTTGCCTC
CCAATGTTGTCCCTTTCCTTCGTTCCCATGGTAAAGCTCCTCTCGCTTTCCTCCTGAGGCTACAC
CCATGCGTCTCTAGGAACTGGTCACAAAAGTCATGGTGCCTGCATCCCTGCCAAGCCCCCCTGAC
CCTCTCTCCCCACTACCACCTTCTTCCTGAGCTGGGGGCACCAGGGAGAATCAGAGATGCTGGGG
ATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTTGTTCTCAAATATTTTTTAATAAATAGACGAA
ACCACG

FIGURE 28

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMAVLTPKSNRKMESKKRELFSSQIKGL
TGASGKVALLELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDM
RQLADGSMDVVVCTLVLCVQSPRKVLQEVRRVLRPGGVLFWEHVAEPYGSWAFMWQQVFPTW
KHIGDGCCLTRETWKDLENAQFSEIQMERQPPPLKWLPGPHIMGKAVKQSFSSKALICSFPSL
QLEQATHQPIYLPLRGT

Important features:

Signal peptide:

amino acids 1-23

Leucine zipper pattern.

amino acids 10-32

N-myristoylation sites.

amino acids 64-70, 78-84, 80-86, 91-97, 201-207

FIGURE 29

CAATGTTTGCCTATCCACCTCCCCCAAGCCCCTTTACCTATGCTGCTGCTAACGCTGCTGCTGCT
GCTGCTGCTGCTTAAAGGCTCATGCTTGGAGTGGGGACTGGTCGGTGCCCAGAAAGTCTCTTCTG
CCACTGACGCCCCCATCAGGGATTGGGCCTTCTTTCCCCCTTCCTTTCTGTGTCTCCTGCCTCAT
CGGCCTGCCATGACCTGCAGCCAAGCCCAGCCCCGTGGGGAAGGGGAGAAAGTGGGGGATGGCTA
AGAAAGCTGGGAGATAGGGAACAGAAGAGGGTAGTGGGTGGGCTAGGGGGGCTGCCTTATTTAAA
GTGGTTGTTTATGATTCTTATACTAATTATACAAAGATATTAAGGCCCTGTTCAATTAAGAAATT
GTTCCCTTCCCCTGTGTTCAATGTTTGTAAAGATTGTTCTGTGTAAATATGTCTTTATAATAAAC
AGTTAAAAGCTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 30

MLLLTLLLLLLLLLLKGSCLEWGLVGAQKVSSATDAPIRDWAFFPPSFLCLLPHPAMPMTCSQAQPRG
EGEKVGDG

Important features:

Signal peptide:

amino acids 1-15

Growth factor and cytokines receptors family:

amino acids 3-18

FIGURE 31

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCGATT
 CCTCCAAGCAAGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCT
 ATTTGCATCTGTTTTGATAAATGATGTTGACACCCTCCACCGAATTCTAAGTGGAATCATGTCGG
 GAAGAGATAACAATCCTTGGCCTGTGTATCCTCGCATTAGCCTTGTCTTTGGCCATGATGTTTACC
 TTCAGATTCATCACCACCCTTCTGGTTCACATTTTCATTTTCATTGGTTATTTTGGGATTGTTGTT
 TGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCAGCATAGAATTGGACA
 CAGAAAGGGAAAATATGAAGTGCGTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGGCAGTG
 CTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCAC
 AAATAAAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCA
 TTTTCTTCTGGGTCTCTGGGTGGCTGTGCTGCTGAGCCTGGGAACTGCAGGAGCTGCCCAGGTT
 ATGGAAGGCGGCCAAGTGGAATATAAGCCCCCTTCGGGCATTTCGGTACATGTGGTTCGTACCATTT
 AATTGGCCTCATCTGGACTAGTGAATTCATCCTTGCGTGCCAGCAAATGACTATAGCTGGGGCAG
 TGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATCATCCCATCCTTTTCGTCTCTC
 TCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCTCTGTGGTGAG
 GATTCCGAGAATCATTGTCTATGTACATGCAAAACGCACTGAAAGAACAGCAGCATGGTGCAATTGT
 CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTC
 AACCAGAATGCATATACTACAACCTGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGC
 ATTCAAAATCTTGTCCAAGAACTCAAGTCACTTTACATCTATTAAGTCTTTGGAGACTTCATAA
 TTTTCTAGGAAAGGTGTTAGTGGTGTGTTTCACTGTTTTTTGGAGGACTCATGGCTTTTAACTAC
 AATCGGGCATTCCAGGTGTGGGCAGTCCCTCTGTTATTGGTAGCTTTTTTTGCCTACTTAGTAGC
 CCATAGTTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACTTTTCTGTGTTTTGCTGTTGATC
 TGGAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTTTCGTA
 AAAAGGAGCAACAAATTAAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAGGAATGAGGA
 GGGAACAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATT
 TCCTTCTAAGAGCCATTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTT
 TTAAAAGACCTAATAAACCTATTCTTCCTCAAAA

FIGURE 32

MSGRDTILGLCILALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTNDLSIE
LDTERENMKCVLGFAIVSTGITAVLLVLI FVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFA
ILIFFWVLWVAVLLSLGTAGAAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQOMTIA
GAVVTCYFNRSKNDPPDHPILSSLSILFFYHQGTVVKGSFLISVVRIPIIVMYMQNALKEQQHG
ALSRYLFRCCYCCFWCLDKYLLHLNQAYTTTAINGTDFCTSAKDAFKILSKNSSHFTSINCFGD
FIIFLGKVLVVCFTVFGGLMAFNYNRAFQVWAVPLLLVAFFAYLVAHSFLSVFETVLDALFLCFA
VDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR

Important features:

Signal peptide:

amino acids 1-20

Putative transmembrane domains:

amino acids 35-54, 75-97, 126-146, 185-204, 333-350, 352-371

N-glycosylation sites.

amino acids 204-208, 295-299, 313-317

N-myristoylation sites.

amino acids 147-153, 178-184, 196-202, 296-275, 342-348

FIGURE 33

[illegible]

FIGURE 34

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAG
CQDPKYHVGTDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYNGVQSLSLPRWR
ESFIVLESKPKKGVITYPSALTYSSSKSPAAQAGETTKAYQRPPIPGTTAQPVTLMQLLAVTVAVA
TPTTLPRPSPSAASTTSIPRPQSVGHRSEQEMDLWSTATYTSQNRPRADPGIQRQDPSGAAFQKP
VGADVSLGLVPKEELSTQSLEPVSLGDPNCKIDLFLIDGSTSIGKRRFRIQKQLLADVAQALDI
GPAGPLMGVVQYGDNPATHTNLKTHNSRDLKTAIEKITQRGGLSNVGRAISFVTKNFFSKANGN
RSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENEKQYVVEPNFANKAVCRTNG
FYSLHVQSWFGLHKTLPQPLVKRVCDTDLACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTN
LTKEFEISDTRIGAVQYTYEQRLFEGFDKYSSKPDILNAIKRVGYWSSGTSTGAAINFALEQL
FKKSKPNKRKLMILITDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARHSEFF
VDEFDNLHQYVPRIIQNICTEFNSQPRN

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 181-200

N-glycosylation sites.

amino acids 390-394, 520-524

N-myristoylation sites.

amino acids 23-29, 93-99, 115-121, 262-268, 367-373, 389-395,
431-437, 466-472, 509-515, 570-576, 571-577, 575-581, 627-633

Amidation site.

amino acids 304-308

FIGURE 35

CCGAGCACAGGAGATTGCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGC
CAAACCATGTCTTTTTTCTGTTTTTCAGAGTAGTTCACAACAGATCTGAGTGTTTTAATTAAGCATGGAAT
ACAGAAAACAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCCACAGTTTTCTTAGCTCCCTGGACCC
GGTTGACCTGTTGGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCCGAGTGTA
AAGAACCTTCGGCTCGCGTGCTTCTGAGCTGCTGTGGATGCGCTCGGCTCTCTGGACTGTCTTCCGAGTA
GGATGTCACTGAGATCCCTCAAATGGAGCCTCCTGCTGCTGCTCACTCCTGAGTTCTTTGTGATGTGGTAC
CTCAGCCTTCCCCACTACAATGTGATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTA
CAGACAAGACTTTCACCTTCACACTTCGAGAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTCT
TGGTGACCTCCACCCCTTCAGATGTGAAAGCCAGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCT
TGGTGGGGATATGAGGTTCTTACATTTTTCTTATTAGGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGC
ATTGTCTTAGAGGATGAACACCTTCTTTATGGTGACATAATCCGACAAGATTTTTTAGACACATATAATA
ACCTGACCTTGAAAACCATATGGCATTGAGGTGGGTAACTGAGTTTTGCCCAATGCCAAGTACGTAATG
AAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTATCTTTTAAACCTAAACCACTCAGA
GAAGTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTTTACCAAAAACCCATATTT
CTTACCAGGAGTATCCTTTCAAGGTGTTCCCTCCATACTGCAGTGGGTGGGTATATAATGTCCAGAGAT
TTGGTGCCAAGGATCTATGAAATGATGGGTACGTAAACCCATCAAGTTTGAAGATGTTTATGTGGGAT
CTGTTTGAATTTATTAAAAGTGAACATTCATATCCAGAAGACACAAATCTTTTCTTCTATATAGAATCC
ATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACTTTT
TGCGAGGTGATGCTAAGGAACACCACATGCCATTATTAACTTCACATTCTACAAAAGCCTAGAAGGACAG
GATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTCAGTGTGCTGGCTT
ACACTGAAGTGAAGTCAAGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGG
CCCTTCAAAGATGATATGTGGAGGAATTAAATATAAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGG
ACCAAACAATTTGGACATGTCATTCTGTAGACTAGAAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCA
CTAGGCTGTAAAAACAAAACAATGTAGAGTTTTATTTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTA
TATCTTATGTGGATTACCAATTTAAAAATATATGTAGTTCTGTGTCAAAAACTTCTTCACTGAAGTTATA
CTGAACAAAATTTTACCTGTTTTTGGTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTTACAGTTATT
ATTATTTAAATTAAGTCAACTTTGTGTTTTTAAATGTTTTGACGATTTCAATACAAGATAAAAAGGATAG
TGAATCATTCTTTACATGCAAACATTTTCCAGTACTTAACTGATCAGTTTATTATTGATACATCACTCCA
TTAATGTAAAGTCATAGGTCATTATTGCATATCAGTAATCTCTTGGACTTTGTTAAATATTTTACTGTGGT
AATATAGAGAAGAATTAAAGCAAGAAAATCTGAAAA

FIGURE 36

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHF
TLREHSNCSHQNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLA
LSLEDEHLLYGDIIRQDFLDTYNNLTTLKTIMAFRWVTEFCPNAKYVMKTDTDVFINTGNLVKYL
NLNHSEKFFTGYPLIDNYSYRGFYQKTHISYQYEPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV
KPIKFEDVYVGICLNLLKVNHIHIPEDTNLFFLYRIHLDVCQLRRVIAAHGFSSKEIITFWQVMLR
NTTCHY

Important features:

Type II transmembrane domain:

amino acids 20-39

N-glycosylation sites.

amino acids 72-76, 154-158, 198-202, 212-216, 326-330

Glycosaminoglycan attachment site.

amino acids 239-243

Ly-6 / u-PAR domain proteins.

amino acids 23-37

N-myristoylation site.

amino acids 271-277

FIGURE 37

CGCTCGGGCACCAGCCGCGGCAAGGATGGAGCTGGGTTGCTGGACGCAGTTGGGGCTCAGTTTCTTCAGCTCCTTCTCATC
TCGTCTTGGCCAAGAGAGTACACAGTCATTAATGAAGCCTGCCCTGGAGCAGAGTGAATATCATGTGTCGGGAGTGTGTG
AATATGATCAGATTGAGTGCCTCTGCCCCGGAAGAGGGAAGTCTGGGTTATACCATCCCTTGCTGCAGGAATGAGGAGAA
TGAGTGTGACTCCTGCCTGATCCACCCAGGTTGTACCATCTTTGAAAACGCAAGAGCTGCCGAAATGGCTCATGGGGGGGT
ACCTTGGATGACTTCTATGTGAAGGGTTCTACTGTGCAGAGTGCCGAGCAGGCTGGTACGGAGGAGACTGCATGCCATGTG
GCCAGGTTCTGCGAGCCCCAAAGGGTCAGATTTTGTGGAAAGCTATCCCTAAATGCTCACTGTGAATGGACCATTTCATGC
TAAACCTGGGTTTGTTCATCCAACTAAGATTGTGTCATGTTGAGTCTGGAGTTTACTACATGTGCCAGTATGACTATGTTGAG
GTTCTGTATGGAGACAACCGCGATGGCCAGATCATCAAGCGTGTCTGTGGCAACGAGCGGCCAGCTCCTATCCAGAGCATAG
GATCCTCACTCCACGTCCTCTTCCACTCCGATGGCTCCAAGAATTTTGACGGTTTCCATGCCATTTATGAGGAGATCACAGC
ATGCTCCTCATCCCTTTGTTTCCATGACGGCACGTGCGTCTTGACAAGGCTGGATCTTACAAGTGTGCCTGCTTGGCAGGC
TATACTGGGCAGCGCTGTGAAAATCTCCTTGAAGAAAGAACTGCTCAGACCTGGGGGGCCAGTCAATGGGTACCAGAAAA
TAACAGGGGGCCCTGGGCTTATCAACGGACGCCATGCTAAAATGGCACCGTGGTGTCTTTCTTTTGTAAACACTCCTATGT
TCTTAGTGGCAATGAGAAAAAGAACTTGCCAGCAGAATGGAGAGTGGTCAGGGAAACAGCCCATCTGCATAAAAGCCTGCCGA
GAACCAAAGATTTTCAGACCTGGTGAGAAGGAGAGTTCTTCCGATGCAGGTTCAAGGGAGACACCATTACACCAGCTAT
ACTCAGCGGCTTTCAGCAAGCAGAACTGCAGAGTGCCCTACCAAGAAGCCAGCCCTTCCCTTTGGAGATCTGCCCATGGG
ATACCAACATCTGCATACCCAGCTCCAGTATGAGTGCATCTACCCCTTCTACCGCCGCTGGGCAGCAGCAGGAGGACATGT
CTGAGGACTGGGAAGTGGAGTGGCGGGCACCATCCTGCATCCCTATCTGCGGGAAAATTGAGAACATCACTGCTCCAAAGA
CCCAAGGGTTGCGCTGGCCGTGGCAGGCAGCCATCTACAGGAGGACCAGCGGGGTGCATGACGGCAGCCTACACAAGGGAGC
GTGGTTCCTAGTCTGCAGCGGTGCCCTGGTGAATGAGCGCACTGTGGTGGTGGCTGCCCACTGTGTTACTGACCTGGGGAAG
GTCACCATGATCAAGACAGCAGACCTGAAAGTTGTTTTGGGGAAATTCTACCGGGATGATGACCGGGATGAGAAGACCATCC
AGAGCCTACAGATTTCTGCTATCATTCTGCATCCCAACTATGACCCCATCTGCTTGATGCTGACATCGCCATCCTGAAGCT
CCTAGACAAGGCCCGTATCAGCACCCGAGTCCAGCCCATCTGCCTCGCTGCCAGTCCGGATCTCAGCACTTCCTTCCAGGAG
TCCCACATCACTGTGGCTGGCTGGAATGTCTTGGCAGACGTGAGGAGCCCTGGCTTCAAGAACGACACACTGCGCTCTGGGG
TGGTCAGTGTGGTGGACTCGCTGCTGTGTGAGGAGCAGCATGAGGACCATGGCATCCAGTGAGTGTCACTGATAACATGTT
CTGTGCCAGCTGGGAACCCACTGCCCCCTCTGATATCTGCACTGCAGAGACAGGAGGCATCGCGGCTGTGTCTTCCCGGGA
CGAGCATCTCCTGAGCCACGCTGGCATCTGATGGGACTGGTCAGCTGGAGCTATGATAAAACATGCAGCCACAGGCTCTCCA
CTGCCCTTACCAAGGTGCTGCCTTTTAAAGACTGGATTGAAAGAAATATGAAATGAACCATGCTCATGCACTCCTTGAGAAG
TGTTTCTGTATATCCGTCTGTACGTGTGTCATTGCGTGAAGCAGTGTGGGCTGAAGTGTGATTGGGCTGTGAAGTGGCT
GTGCCAGGGCTTCTGACTTCAGGGACAAAACCTCAGTGAAGGGTGAGTAGACCTCCATTGCTGGTAGGCTGATGCCGCGTCCA
CTACTAGGACAGCCAATTGGAAGATGCCAGGGCTTGCAAGAAGTAAGTTTCTTCAAAGAAGACCATATACAAAACCTCTCCA
CTCCACTGACCTGGTGGTCTTCCCCAACTTTCAGTTATACGAATGCCATCAGCTTGACAGGGAAGATCTGGGCTTCATGAG
GCCCCTTTGAAGCTCTCAAGTTCTAGAGAGCTGCCTGTGGGACAGCCAGGGCAGCAGAGCTGGGATGTGGTGCATGCCTT
TGTGTACATGGCCACAGTACAGTCTGGTCTTTTCCCTTCCCCATCTCTTGTACACATTTTAAATAAAATAAGGGTTGGCTTCT
GAACTACAAAAA
AAAAA

FIGURE 38

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYT
IPCCRNEENECDSCLIHPGCTIFENCKSCRNGSWGGLDDFYVKGIFYCAECRAGWYGGDCMRGQ
VLRAPKGQILLESYPLNAHCEWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGQII
KRVCGNERPAPIQSIGSSSLHVLFHSDGSKNFDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKC
ACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLINGRHAKIGTVVSFFCNNSYVLSGNE
KRTCQONGEWSGKQPICIKACREPKISDLVRRRVLPMPVQSRETPHQLYSAAFSKQKLQSAPTK
KPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGIENITAP
KTQGLRWPWQAAIYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAAHCVTDLGKVTMIKTADL
KVVLGKFYRDDDRDEKTIQSLQISAILHPNYDPILLDADIAILKLLDKARISTRVQPICLAASR
DLSTSFQESHITVAGWNVLADVRSFGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCA
SWEPTAPSDICTAETGGIAAVSFPGRASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLFPKDWI
ERNMK

Important features of the protein:

Signal peptide:

amino acids 1-23

EGF-like domain cysteine pattern signature.

amino acids 260-272

N-glycosylation sites.

amino acids 96-100, 279-283, 316-320, 451-455, 614-618

N-myristoylation sites.

amino acids 35-41, 97-103, 256-262, 284-290, 298-304, 308-314,
474-480, 491-497, 638-644, 666-672

Amidation site.

amino acids 56-60

Serine proteases, trypsin family.

amino acids 489-506

CUB domain proteins profile.

amino acids 150-167

FIGURE 39

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTAATC
TGAAGGTTCTCAGTCAAATTCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTTGGCTGG
TTTGGGCCCCCTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGCTTCTGTTGC
TGGTCTTGCCCTTGGCTCAGTCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCTGTATTCAGAACTCTGTA
AAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAGACGGCTGTGCGAGCC
TCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAGACGAGCCTGGCCTAGACA
ACCTTGCCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACTCTGGCCGAGCAACCGAACTA
GGGCACGGCCCCCTTGTAGAGATCCACTATTAGAAGCAGATCATTTAAAAAAATAAATCGAGCTTTGAGTGTCTTCGAA
GGACAAAGAGCGGGAGTGCAATTGCCAACCATGCCAGCAGGGCAGGAAAATTCTGAAAACACCACTGCCCTGAAG
TCTTTCCAAGGTTGTACCACTGATTCCAGATGGTGAAATTACCAGCATCAAGATCAATCGAGTAGATCCCAGTGAAA
GCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCCATATCATTATCCAACACATTTATCGTGATGGGG
TGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTCTAAAGGTCAACGGGATGGACATCAGCAATGTCCCTC
ACAACTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGGTGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCCGCA
GCAGGAACAATGGACAGGCCCCGGATGCCTACAGACCCGAGATGACAGCTTTCATGTGATTCTCAACAAAAGTAGCC
CCGAGGAGCAGCTTGAATAAACTGGTGCGCAAGGTGGATGAGCCTGGGGTTTTTCATCTCAATGTGCTGGATGGCG
GTGTGGCATATCGACATGGTCAGCTTGAGGAGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCA
GCCCAGAAAAGTGCGGCTCATCTGATTGAGGCCAGTGAAAGACGTGTTACCTCGTCGTGTCCCGCCAGGTTCCGCAGC
GGAGCCCTGACATCTTTCAGGAAGCCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACA
CTCCCCAGCCCCCTCCATCCTACAATTACTTGTTCATGAGAAGGTGGTAAATATCCAAAAGACCCCGGTGAATCTCTCG
GCATGACCGTCGACGGGGAGCATCACATAGAGAATGGGATTTGCCTATCTATGTCATCAGTGTGAGCCCCGAGGAG
TCATAAGCAGAGATGGAAGAATAAAAAAGGTGACATTTTGTGTAATGTGGATGGGGTGAAGTACAGAGGTCAGCC
GGAGTGAGGCAGTGGCATTATTGAAAAGAACATCATCCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGC
CCCAGGAAGACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTACTGGTCCCCATCCT
GGGTGATGTGGCTGGAATTACCACGGTGCTTGATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTC
TGGGCTTCTGCATTGTAGGAGGTTATGAAGAATACAATGGAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA
CACCAGCATACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTGCTGTCAATGGTAGAAGTACATCAGGAATGA
TACATGCTTGCTTGGAAGACTGCTGAAAGAACTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGGCCTGGCACTT
TTTTATAGAATCAATGATGGGTGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACTATATTTATC
TTGTGAGTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTCAGGAAAAGTATGATCATCTAATGAAAGCCAGTT
ACACCTCAGAAAATATGATTCAAAAAAATTAATACTACTAGTTTTTTTTTCAGTGTGGAGGATTTCTCATTACTCTAC
AACATTGTTTATATTTTTCTATTCAATAAAAAGCCCTAAAACAATAAATGATTGATTGTATACCCCACTGAATT
CAAGCTGATTTAAATTAAAAATTTGGTATATGCTGAAGTCTGCCAAGGTACATTATGGCCATTTTTTAATTTACAGCT
AAAAATATTTTTTAAATGCATTGCTGAGAAACGTTGCTTTTCATCAAACAAGAATAAATATTTTTTCAGAAGTTAAA

FIGURE 40

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPS
PEVSAAATISLMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINR
ALSVLRRTKSGSAVANHADQGRENSENTTAPEVFPRLYHLIPDGEITSIKINRVDPSESLSIRLV
GGSETPLVHIIIQHIYRDGVIARDGRLLPGDIILKVNGMDISNVPHNYAVRLLRQPCQVLWLTVM
REQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIKLVRKVDEPGVFIFNVLDGGVAYRHG
QLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQVRQSPDIFQEAGWNSNGSWSPG
PGRSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPIYVISVEPGGVISRDGR
IKTGDILLNVGDVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPP
SDWSPSWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDG
RIRCGDILLAVNGRSTSGMIHACLARLLKELKGRITLTIVSWPGTFL

Important features:

Signal peptide:

amino acids 1-15

N-glycosylation sites.

amino acids 108-112, 157-161, 289-293, 384-388

Tyrosine kinase phosphorylation sites.

amino acids 433-441, 492-500

N-myristoylation sites.

amino acids 51-57, 141-147, 233-239, 344-350, 423-429, 447-453,
467-473, 603-609

FIGURE 41

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTT
 CTTGCCTGCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGCTTCAACCTGACT
 TTCCACCTTTCCTACAAATTCCGATTACTGTTGCTGTTGACTTTGTGCCTGACAGTGGTTGGGTGGGC
 CACCAGTAACTACTTCGTGGGTGCCATTCAAGAGATTCCATAAGCAAAGGAGTTCATGGCTAATTTCC
 ATAAGACCCTCATTTTGGGGAAGGGAAAACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGTT
 GACAACTGTCCTTCTGTGTCTCCTTACCTCAGAGGCCAGAGCAAGCTCATTTTCAAACCAGATCTCAC
 TTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGGCCGGTATCGCCCTCAGGAATGTAAAG
 CTTTACAGAGGGTCGCCATCCTCGTTCCCCACCGGAACAGAGAGAAACACCTGATGTACCTGCTGGAA
 CATCTGCATCCCTTCCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCACCAGGCTGAAGG
 TAAAAAGTTTAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAAATTGGG
 ACTGCTTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAG
 CATCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGTTACGTTACAGTGGATATTTTGG
 GGGTGTACTGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTGCGGGAT
 GGGGAGGCGAAGACGATGACCTCAGACTCAGGGTTGAGCTCCAAAGAATGAAAATTTCCCGGCCCTG
 CCTGAAGTGGGTAAATATACAATGGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACG
 GATGAAGCTCTTACACCAAGTGTACGAGTCTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAAT
 TAGTATCTGTGGAACACAATCCTTTATATATCAACATCACAGTGGATTTCTGGTTTTGGTGCATGACCC
 TGGATCTTTTGGTGATGTTTGAAGAAGTATTCTTTGTTTGCAATAATTTTGGCCTAGAGACTTCAA
 ATAGTAGCACACATTAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTTTCCTTTTGTATTTTCT
 TAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAGTCATTTTGAT
 CATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGATAA
 AATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGAT
 AAAAGGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAA
 AGGTACGAAGATACAATACTGTTATTATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGT
 GAGAAGGCGTCCACAAAAGAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTTGGGAATGAAGA
 GGTAGCAGGAGGGTGGAGTGTGCGCTGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGC
 CTTCAGGGGAGGACCTGCCCAGGTATGCCTTCCAGTGATGCCACCAGAGAATACATTCTCTATTAGT
 TTTTAAAGAGTTTTTGTAAATGATTTTGTACAAGTAGGATATGAATTAGCAGTTTACAAGTTTACAT
 ATTAACATAATAAATATGTCTATCAAATACCTCTGTAGTAAATGTGAAAAGCAAAA

FIGURE 42

MGFNLT FHLSYKFRLLLLLTLCLTVVGWATSNYFVGAIQEIPKAKEFMANFHKTLLLGKGKTLTN
EASTKKVELDNCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYPQECKALQRVAILVPH
RNREKHLMYLLEHLHPFLQRQQLDYGIYVIHQAEKKFNRAKLLNVGYLEALKEENWDCFI FHDV
DLVPENDFNLYKCEEHPKHLVVGRNSTGYRLRYSGYFGGVTALSREQFFKVNGFSNNYWGWGGED
DDLRLRVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAERMKLLHQVSRVWRTDGLSSCSYKLV
SVEHNPLYINITVDFWFGA

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 4-8, 220-224, 335-339

Xylose isomerase proteins.

amino acids 191-202

FIGURE 44

MALSSQIWAACLLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPI
CIFCCGCCHRSKCGMCCKT

Important features:

Signal peptide:

amino acids 1-24

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 58-59

N-myristoylation site.

amino acids 44-50

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 1-12

FIGURE 45

GTGGCTTCATTTTCAGTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCC
TCATCTATATCCTTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGT
TCCGTTGGTGGGGCCGTGACTTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTG
GACCTTCAACACAACCCCTCTTGTCACCATACAGCCAGAAGGGGGCACTATCATAGTGACCCAAA
ATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTACTCCCTGAAGCTCAGCAAACCTGAAG
AAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCAGCAGCCCTCCACCCA
GGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCAGAGCA
ATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATT
TATACCTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTC
CTGGAGATGGGGAGAAAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACT
TCTCAAGCCCCATCCTTGCCAGGAAGCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCTCCATG
GTCCTCCTGTGTCTCCTGTTGGTGCCCTCCTGCTCAGTCTCTTTGTACTGGGGCTATTTCTTTG
GTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAAGAGAGTGGACATTTGTCGGG
AAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCCTCACACTAAT
AGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT
GGAAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTA
TCTTAGACAGCAGTGCACTCCCCCTAAGTCTCTGCTCA

FIGURE 46

MAGSPTCLTLIYILWQLTGSAASGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQP
EGGTIIIVTQNRNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSK
PKVTMGLQSNKNGTCVTNLTCCMEHGEEVDVIYTWKALGQAANESHNGSILPISWRWGESDMTFIC
VARNPVSRNFFSPILARKLCEGAADDPDSSMVLLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIE
EKKRVDICRETPNICPHSGENTYDITPHTNRTILKEDPANTVYSTVEIPKKMENPHSLLTMPDT
PRLFAYENVI

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 224-250

Leucine zipper pattern.

amino acids 229-251

N-glycosylation sites.

amino acids 98-102, 142-146, 148-152, 172-176, 176-180, 204-208,
291-295

FIGURE 47

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGG
ATTCAGCCTGCTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCT
TAGTTGAGGAAGACCAATTTTCTCAAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATT
ATAGGAGCAGGTCTGATGGCCATTCCAGCAACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTG
CTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTTCAGTGTGATCACAGTCATTGGTGCTC
TGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAAGGTCCTCTCATGTGTAATTCTCCAAGC
AACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTCATCCAGAATCCTTCAA
CTTGCACTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTTCAATAAACCCACCAGTAACG
ACACCATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAAACAAACATAGG
CTTATCCACTTCTCAGTATTTTTAGGTCTATTGCTTGTGGAATTCTGGAGGTCCTGTTTGGGCT
CAGTCAGATAGTCATCGGTTTCCTTGGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTG
TGTAGTTTAATGGGAATAAAATGTAAGTATCAGTAGTTTGAAAAAAAAA

FIGURE 48

MTCCEGWTSCNGFSLLVLLLLLVVLLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPA
TTMSLTARKRACCNNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSL
KNISDIHPESFNLQWFFNDSCAPPTGFNKPTSNDTMASGWRASSFHFDSEENKHRLIHFSVFLGL
LLVGILEVLFLGLSQIVIGFLGCLCGVSKRRSQIV

Important features:

Transmembrane domains:

amino acids 10-31 (type II), 50-72, 87-110, 191-213

N-glycosylation sites.

amino acids 80-84, 132-136, 148-152, 163-167

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 223-227

N-myristoylation sites.

amino acids 22-28, 54-60, 83-89, 97-103, 216-222

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 207-218

TNFR/NGFR family cysteine-rich region protein.

amino acids 4-12

FIGURE 49

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGA
AGGAGCAGTGAGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAG
ATCCGTGGGCTGCAGACCCCCGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGA
CATGGAGAGAGTGACCCTGGCCCTTCTCCTACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACC
CATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAAAAACCTGCAGCTGAGCGGACTGATC
TGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAAATGCAAATACAAGAG
CAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTCTGCCA
CTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACC
TCCTCCCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTAGGCCCCCT
TTCTGATCAGGAGGCTTCTTTATGAATTAACTCGCCCCACCACCCCTCA

FIGURE 50

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKS
SQKQHSPVPEKAIPILITPGSATTC

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 36-59

N-myristoylation sites.

amino acids 41-47, 45-51, 84-90

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7.

amino acids 54-67

FIGURE 51

[illegible]

FIGURE 52

MKFQGPLACLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAI GKEAGGAAGSKVS
 EALGQGTREAVGTGVRQVPFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHGADAVRGSWQGVF
 GHSGAWETSGGHGIFGSGGGLGGQGQGNPGGLGTPWVHGYPGNSAGSFGMNPQGAPWGQGGNGGPPNF
 GTNTQGAVAQPGYGSVRASNQNEGCTNPPPSGSGGGSSNSGGGSGSQSGSSGSGSNGDNNGSSSSGGS
 SSGSSSGSSSGGSSGGSSGGSSGNSGGSRGDSGSESSWGSSTGSSSGNHGSGGGNGHKPGCEKPGNE
 ARGSGESGIQGFRGQGVSSNMREISKEGNRLLGSGSDNYRGQSSWGSGGGDAVGGVNTVNSETSPGM
 FNFDTFWKNFKSKLGFINWDAINKDQRSSRIP

Signal peptide:

amino acids 1-21

N-glycosylation site.

amino acids 265-269

Glycosaminoglycan attachment site.

amino acids 235-239, 237-241, 244-248, 255-259, 324-328, 388-392

Casein kinase II phosphorylation site.

amino acids 26-30, 109-113, 259-263, 300-304, 304-308

N-myristoylation site.

amino acids 17-23, 32-38, 42-48, 50-56, 60-66, 61-67, 64-70, 74-80,
90-96, 96-102, 130-136, 140-146, 149-155, 152-158, 155-161,
159-165, 163-169, 178-184, 190-196, 194-200, 199-205, 218-224,
236-242, 238-244, 239-245, 240-246, 245-251, 246-252, 249-252,
253-259, 256-262, 266-272, 270-276, 271-277, 275-281, 279-285,
283-289, 284-290, 287-293, 288-294, 291-297, 292-298, 295-301,
298-304, 305-311, 311-317, 315-321, 319-325, 322-328, 323-329,
325-331, 343-349, 354-360, 356-362, 374-380, 381-387, 383-389,
387-393, 389-395, 395-401

Cell attachment sequence.

amino acids 301-304

FIGURE 53

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGG
CTGGGCCTCAGACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCT
ACTCGCCCGCATCCTGGCTTGGACCTATGCCTTCTATAACAACCTGCCGCCGGCTCCAGTGTTTCC
CACAGCCCCCAAACGGAACCTGGTTTTGGGGTCACCTGGGCCTGATCACTCCTACAGAGGAGGGC
TTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGGGCTTTACGGTATGGCTGGGTCCCAT
CATCCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCAATGCCTCAGCTGCCA
TTGCACCCAAAGGATAATCTCTTCATCAGGTTTCTGAAGCCCTGGCTGGGAGAAGGGATACTGCTG
AGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCACATCCT
GAAGTCCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGG
CCTCAGAGGGCAGCAGTCGTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTA
CAGAAATGCATCTTCAGCTTTGACAGCCATTGTCAGGAGAGGGCCAGTGAATATATTGCCACCAT
CTTGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCCAGCATATCCTCCAGCACATGGACTTTCTGT
ATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCCTGGTGCATGACTTCACAGAC
GCTGTTCATCCGGGAGCGGCGTTCGACCCCTCCCCACTCAGGGTATTGATGATTTTTTCAAAGACAA
AGCCAAGTCCAAGACTTTGGATTTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG
CATTGTCAGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACG
GCCAGTGGCCTCTCCTGGGTCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCG
ACAGGAGGTGCAAGAGCTTCTGAAGGACCGGATCCTAAAGAGATTGAATGGGACGACCTGGCCC
AGCTGCCCTTCTGACCATGTGCGTGAAGGAGAGCCTGAGGTTACATCCCCAGCTCCCTTCATC
TCCCGATGCTGCACCCAGGACATTGTTCTCCCAGATGGCCGAGTCATCCCCAAAGGCATTACCTG
CCTCATCGATATTATAGGGGTCCATCACAAACCAACTGTGTGGCCGGATCCTGAGGTCTACGACC
CCTTCCGCTTTGACCCAGAGAACAGCAAGGGGAGGTCACCTCTGGCTTTTATTCTTTCTCCGCA
GGGCCCAGGAACTGCATCGGGCAGGCGTTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGAT
GCTGCTGCACTTCCGGTTCCTGCCAGACCACACTGAGCCCCGCAGGAAGCTGGAATTGATCATGC
GCGCCGAGGGCGGGCTTTGGCTGCGGGTGGAGCCCCCTGAATGTAGGCTTGACAGTGACTTTCTGAC
CCATCCACCTGTTTTTTTTGCAGATTGTCATGAATAAAACGGTGCTGTCAA

FIGURE 54

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPQPPKRNWFWGHLG
LITPTEEGLKDSTQMSATYSQGFTVWLGPIIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKP
WLGEGILLSGGDKWSRHRMLTPAFHFNILKSYITIFNKSANIMLDKWQHLASEGSSRLDMFEHI
SLMTLDSLQKCIFSFDSHCQERPSEYIATILELSALVEKRSQHILQHMDFLYYLSHDGRRFHRAC
RLVHDFDTAVIRERRRTLPTQGIDDFKDKAKSKTLD FIDVLLLSKDEDGKALSDEDIRAEADTF
MFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDLAQLPFLTMCVKESLR
LHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDFFRFDPENSKGRSP
LAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLN
VGLQ

Important features:

Transmembrane domains:

amino acids 13-32 (type II), 77-102

Cytochrome P450 cysteine heme-iron ligand signature.

amino acids 461-471

N-glycosylation sites.

amino acids 112-116, 168-172

FIGURE 55

[illegible]

FIGURE 56

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCIIQSLALTWYSLSFIPFAR
DAVKKCFVCLA

Important features:

Signal peptide:

amino acids 1-33

Type II fibronectin collagen-binding domain protein.

amino acids 30-72

FIGURE 57

[illegible]

FIGURE 58

MLCLCLYVPVIGEAEQTEFOYFESKGLPAELKSI FKLSVFIPSQEFSTYRQWKQKIVQAGDKDLDG
QLDFEEFVHYLQDHEKKLRLVFKILDKKNDGRIDAQEIMQSLRDLGVKISEQQA EKILKSMDKNG
TMTIDWNEWRDYHLLHPVENIPEIILYWKHSTIFDVGENLTVPDEFTVEERQTGMWWRHLVAGGG
AGAVSRTCTAPLDRLKVLMOVHASRSNNMGIVGGFTQMIREGGARSLWRGNNGINVLKIAPESAIAK
FMAYEQIKRLVGSDQETLRIHERLVAGSLAGAI AQSSIYPMEVLKTRMALRKTGQYSGMLDCARR
ILAREGVAAFYKGYVPNMLGIIPYAGIDLAVYETLKNAWLQHYAVNSADPGVFVLLACGTMSSTC
GQLASYPLALVRTRMQAQASIEGAPEVTMSSLFKHILRTEGAFGLYRGLAPNFMKVIPAVSISYV
VYENLKITLGVQSR

Important features:

Signal peptide:

amino acids 1-16

Putative transmembrane domains:

amino acids 284-304, 339-360, 376-394

Mitochondrial energy transfer proteins signature.

amino acids 206-215, 300-309

N-glycosylation sites.

amino acids 129-133, 169-173

Elongation Factor-hand calcium-binding protein.

amino acids 54-73, 85-104, 121-140

FIGURE 59

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGC
 TTCCCTGGGGCAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTG
 CACTCATCATTGGCTTTGGTATTTAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCT
 GGGAACATTGGGGAGGATGGAATCCTGAGCTGCACTTTTGAACCTGACATCAAACCTTTCTGATAT
 CGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTCCATGAGTTCAAAGAAGGCAAAGATG
 AGCTGTGCGAGCAGGATGAAATGTTGAGAGGCCGACAGCAGTGTGCTGATCAAGTGATAGTT
 GGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGTTATAT
 CATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGG
 AAGTGAATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCC
 CAGCCCACAGTGGTCTGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATAC
 CAGCTTTGAGCTGAACTCTGAGAATGTGACCATGAAGGTTGTGTCTGTGCTCTACAATGTTACGA
 TCAACAACACATACTCCTGTATGATTGAAAATGACATTGCCAAAGCAACAGGGGATATCAAAGTG
 ACAGAATCGGAGATCAAAAGGCGGAGTCACCTACAGCTGCTAAACTCAAAGGCTTCTCTGTGTGT
 CTCTTCTTTCTTTGCCATCAGCTGGGCACTTCTGCCTCTCAGCCCTTACCTGATGCTAAAATAAT
 GTGCCTTGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTTAC
 CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTG
 AGCAAACAAGAGCAAGAAACAAAAAGAAGCCAAAGCAGAAGGCTCCAATATGAACAAGATAAAT
 CTATCTTCAAAGACATATTAGAAGTTGGGAAAATAATTCATGTGAACTAGACAAGTGTGTTAAGA
 GTGATAAGTAAAATGCACGTGGAGACAAGTGCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGT
 CACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTGTCTCTGAATTTTTAGTTATATGTGC
 TGTAAATGTTGCTCTGAGGAAGCCCCCTGGAAAGTCTATCCCAACATATCCACATCTTATATCCAC
 AAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAGGGGCG
 GCTGCATTTTAGTAATGGGTCAAATGATTCACTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTC
 TCTTCCCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGT
 CGGGGACACCGATTTTATAAATAAACTGAGCACCTTCTTTTTAAACAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 60

MASLGQILFWSIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLS
DIVIQWLKEGVLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKC
YIITSKGKGNANLEYKTGAFSMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVS
NTSFELNSENVTMKVVSVLNVNTINNTYSCMIENDIAKATGDIKVTESI KRRSHLQLLNSKASL
CVSSFFAISWALLPLSPYLMLK

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 258-281

N-glycosylation sites.

amino acids 112-116, 160-164, 190-194, 196-200, 205-209, 216-220,
220-224

N-myristoylation sites.

amino acids 52-58, 126-132, 188-194

FIGURE 61

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCCAGGAGCTGCAGGACAAG
CACCAGGAGCCCCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGT
GGGCTACCCCCCTGGTGGTGGTTATGGGGGTCCTGCCCCCTGGAGGGCCTTATGGACCACCAGCTGG
TGGAGGGCCCTATGGACACCCCCAATCCTGGGATGTTCCCCCTCTGGAACCTCAGGAGGACCATATG
GCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCCAAGTTCCTACGGTGCCAGCAGCCT
GGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTACTCCTGGTTCCAGTC
GGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAACTGCA
ATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCA
GGCCGCATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGAAGAACCTCTT
CCAGCAGTATGACCGGGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCC
AAATGGGCTACAACCTGAGCCCCCAGTTCACCCAGCTTCTGGTCTCCCGCTACTGCCCACGCTCT
GCCAATCCTGCCATGCAGCTTGACCGCTTCATCCAGGTGTGCACCCAGCTGCAGGTGCTGACAGA
GGCCTTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCTCAGCTTCGAGGACTTCGTCA
CCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCACCAGGGACCTT
TCCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTTCTGTCCCTCTAGAAGAAC
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCCTGCATCATAGCCACCA
AATAGTGAGGACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGA
ATGTCCTGATGGCCATGAGCAGTTGAGTGGCACAGCCTGGCACCAGGAGCAGGTCTTGTAATGG
AGTTAGTGTCCAGTCAGCTGAGCTCCACCTGATGCCAGTGGTGAGTGTTTCATCGGCCTGTTACC
GTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAACGAGCCCATTTTCTCCAAAGTGAAT
CTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCTGCCACACCCATAAAT
CCTTGTTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTGGGCAT
CTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCCTGCCATGCTCTGCTCGGCT
TCAGTCTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTTAATTTGCATTTTTTTTC
ATTTGGGGCCAAAAGTCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

FIGURE 62

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYG
HPNPGMFPSGTPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDS DH
SGYISMKELKQALVNCNWSSFND ETC LMMINMFDKTKSGRIDVYGFSALWKFIQQWKNLFQQYDR
DRSGSISYTELQQALSQMGYNLSPQFTQLLVSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREK
DTAVQGNIRLSFEDFVTMTASRML

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 147-150

Casein kinase II phosphorylation sites.

amino acids 135-138, 150-153, 202-205, 271-274

N-myristoylation sites.

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-
66, 70-75, 78-83, 83-88, 87-92, 110-115

FIGURE 63

[illegible]

FIGURE 64

MQGRVAGSCAPLGLLLVCLHLPGLFARSIGVVEEKVSQNFGTNLPQLGQPSSTGPSNSEHPQPAL
DPRSNDLARVPLKLSVPPSDGFPPAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAEDRLGEA
LPEELSYLSSAAALAPGSGPLPGESSPDATGLSPEASLLHQDSESRRRLPRSNSLGAGGKILSQRP
PWSLIHRVLPDHPWGTLNPSVSWGGGGPGTGWGTRPMPHPEGIWGINNQPPGTSGWGNINRYPGGS
WGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFPFPGVLRPPGSSWNIPAGFPNPPSPRLQWG

Important features of the protein:

Signal peptide:

amino acids 1-26

Casein kinase II phosphorylation sites.

amino acids 56-59, 155-158

N-myristoylation sites.

amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263,
259-264, 269-274, 270-275, 280-285, 281-286, 305-310

FIGURE 65

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTC
TGGGCTGCCCCTTGTCTCCTCTTGACCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGA
CTTTGCAACTGAAGCTGAAGGAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAA
TTGCTTGAAAAGCTCTGCCTCCTCCTCCATCTCCCTTCAGGGACCAGCGTCACCCTCCACCATGC
AAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCATTGAAGCCTGTGTCCTTCTTGGCCC
GGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAGCAGGCCCCCACCCTC
CTGAGTGGCAATAAATAAAATTCGGTATGCTG

FIGURE 66

MGSGLPLVLLLTLLGSSHGTGPGMTLQKLKESFLTNSSYESSFLELLEKLCLLLHLPSGTSVTL
HHARSQHHVVCNT

Important features:

Signal peptide:

amino acids 1-19

N-glycosylation site.

amino acids 37-41

N-myristoylation sites.

amino acids 15-21, 19-25, 60-66

FIGURE 67

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGC
CAGGTGCCCCGTGCGCAGGTGCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGCGAGAAGCCCC
TTCCTCGGCGCTGCCAACCCGCCACCCAGCCCATGGCGAACCCCGGGCTGGGGCTGCTTCTGGCG
CTGGGCCTGCCGTTCTGCTGGCCCGCTGGGGCCGAGCCTGGGGGGCAAATACAGACCACTTCTGC
AAATGAGAATAGCACTGTTTTGCCTTCATCCACCAGCTCCAGCTCCGATGGCAACCTGCGTCCGG
AAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTCCTGGCTGTGGGGCTG
GCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCCAGTAGCGA
GGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCT
GCCTGCCCCATCTAGGTCCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAA
GGCAGTGCCCTCTCTGGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAA
AGACTCTGCCCCTGAGGTCAAGAGAGGATGGGGCTATTCACTTTTATATATTTATATAAAATTAG
TAGTGAGATGTAAAAAAAAAAAAAAAAAAAAA

FIGURE 68

MANPGLGLLLALGLPFLLARWGRAWGQIQTTSANENSTVLPSSSTSSSSSDGNLRPEAITAIIVVFS
LLAALLLAVGLALLVRKLREKRQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

Important features:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 56-80

N-glycosylation site.

amino acids 36-40

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 86-90

Tyrosine kinase phosphorylation site.

amino acids 86-94

N-myristoylation sites.

amino acids 7-13, 26-32

FIGURE 69

FIGURE 70

MGLFRGFVFLLVLCLLHQSNSTSFIKLNNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFE
ATEKRFFFKNVSIIPENWKENPQYKRPKHENHKKHADVIVAPPTLPGRDEPYTKQFTECGEKGEY
IHFTPDLLLGGKKQNEYGPPGKLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCISAGISGRN
RVYKCQGGSCLSRACRIDSTTKLYGKDCQFFPDQVQTEKASIMFMQSIDSVEFCNEKTHNQEAP
SLQNIKCNFRSTWEVISNSEDFKNTIPMVTPPPPPVFSLKISQRIVCLVLDKSGSMGGKDRLNR
MNQAAKHFLQLQTVENGSWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGLPTYPLGGTSICSGIK
YAFQVIGELHSQLDGEVLLLLTDGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMSKITGG
SHFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLTLNSNAWMNDTVIIDSTVGKDTFFL
ITWNSLPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVTSR
AANSSVPPITVNAKMNDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGA
GADSFKNDBGVYSRYFTAYTENGRYSLKVRAHGGANTARLKLRPPLNRAAYIPGWVNGEIEANPP
RPEIDEDTQTTLEDFSRASGGAFVVSQVPSLPLPDQYPPSQITDLDATVHEDKIILWTAPGDN
FDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKESFAFKPENISEENATHIFIAI
KSIDKSNLTSKVSNIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTLVLSVIGSVVI
VNFILSTTI

Signal peptide:

amino acids 1-21

Putative transmembrane domains:

amino acids 284-300, 617-633

Leucine zipper pattern.

amino acids 469-491, 476-498

N-glycosylation site.

amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592,
628-632, 811-815, 832-836, 837-841, 852-856, 896-900

FIGURE 71

CTCCTTAGGTGGAACCCCTGGGAGTAGAGTACTGACAGCAAAGACCGGGAAGACCATACGTCCTCCCGGGCAGGGGTGA
 CAACAGGTGTCTATCTTTTGTATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAGACGCCAAGGTAAATTTGACCCA
 GAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCCAGTTATGCCAGGATTACTAGAGAGTGTCA
 ACTCAACCAGCAAGCGGCTCCTTCGGCTTAACTTGTGGTTGGAGGAGAGAACCCTTTGTGGGGCTGCGTTCTCTTAGCA
 GTGCTCAGAAGTGACTTGCCTGAGGGTGGACCAGAAAGAAAGGTCCCTCTTGCTGTTGGCTGCACATCAGGAA
 GSCTGTGATGGGAATGAAGGTGAAACTTGGAGATTCACTTCAGTCATTGCTTCTGCTGCAAGATCATCCTTTAAA
 AGTAGAGAAGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAGGAAATGGATGCAAGCAGCTC
 CGGGGGCCCCAAACGCATGCTTCTGTGCTTAGCCAGGGAGGCCCTTCCGTGGGGGCCCCGGCTTTGAGGGATGCC
 ACCGGTTCTGGACGCATGGCTGATTCTGATGATGATGTTTCGCCGGGGGCTGCTTGCCTGGATTTCCTGGGGTGGTG
 GTTTTGTGGTCTCCTCTGCTGTGCTATCTGTCTGTACATGTTGGCCTGCACCCCAAAGGTGACAGGAGCAG
 CTGGCACTGCCAGGGCCAAACAGCCCCACGGGAAGGAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCAC
 CGCAACTACGTGAGCAGCCTGAAGCGGCAGATCGCACAGCTCAAGGAGGAGCTGCAGGAGAGGAGTGAAGCAGCTCAGG
 AATGGGCAGTACCAAGCCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAACCCAGGCCGACCTC
 CTGGCCTTCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGCAGTG
 CCTTTCGATAGCTTTACTCTACAGAAGGTGTACAGCTGGAGACTGGCCTTACCCGCCACCCCGAGGAGAAGCCTGTG
 AGGAAGGACAAAGCGGGATGAGTTGGTGGAGCCATTGAATCAGCCTTGGAGACCCTGAACAATCCTGCAGAGAACAGC
 CCCAATCACCGCTCCTTACAGCGCCTCTGATTTCATAGAAGGGATCTACCGAACAGAAAGGGACAAAGGGACATTGTAT
 GAGCTCACCTTCAAAGGGGACCACAAACACGAATTCAAACGGCTCATCTTATTTGACCATTCAGCCCCATCATGAA
 GTGAAAAATGAAAAGCTCAACATGGCCAAACAGCCTTATCAATGTTATCGTGCCTCTAGCAAAAAGGGTGGACAAGTTC
 CGGCAGTTTATGCGAATTTTCAAGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGG
 AAAGAAGAAATAATGAAGTCAAAGGAATACTTGAACAACCTTCCAAAGCTGCCAATTCAGGAACCTTTACCTTCATC
 CAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGGAAGCAACGCTCCTTCTC
 TTTTCTGTGATGTGGACATCTACTTCACATCTGAATTCCTCAATACGTGTAGGCTGAATACACAGCCAGGGAAGAAG
 GTATTTTATCCAGTTCTTTTTCAGTCAGTACAACTCCTGGCATAATATACGGCCACCATGATGCAGTCCCTCCCTTGGAA
 CAGCAGCTGGTCATAAAGAAAGAACTGGATTTTGGAGAGACTTTGGATTTGGGATGACGTTGTCAGTATCGGTGAGAC
 TTCATCAATATAGTGGGTGATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTC
 CACAGCAACCTCATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAG
 CTGACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCCACGGCCAGCTGGGCATGCTG
 GTGTTGAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAGACAAGTAGCAAAAAACATGAACCTCCAGA
 GAAGGATTGTGGGAGACACTTTTTCTTCTTTTGAATTAAGTGGCTGCAACAGAGAAAGACTTCCATAAA
 GGACGACAAAAGAAATTGGACTGATGGGTGAGAGATGAGAAAGCCTCCGATTTCTCTGTTGGGCTTTTACAACAGA
 AATCAAAATCTCCGCTTTGCTGCAAAAAGTAACCCAGTTGCACCCCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTG
 AGATTATAAGCCTAATGGTGTGGAGGTTTGTGTTGTTTACAATACTGAGACCTGTTGTTTTGTGTGCTCATTGA
 AATATTCATGATTAAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA
 GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAATATCAGAAGGCAGGAGAGGATAGGCTTATTATGATACT
 AGTGAGTACATTAAAGTAAATAAAATGGACCAGAAAAGAAAAGAACCATAAATATCGTGTCAATTTTCCCAAGAT
 TAACCAAAAATAATCTGCTTATCTTTTGGTTGCTTTTAACTGTCTCCGTTTTTTCTTTTATTTAAAAATGCACT
 TTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTACCACCTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTAC
 ATTTTTATATTTTTTAAGAAGATACTTTGAGATGCATTATGAGAATTTTCAAGTCAAAGCATCAAAATGATGCCATAT
 CCAAGGACATGCCAAATGCTGATTCTGTGAGGCACTGAATGTGAGGCATTGAGACATAGGGAAGGAATGGTTTGTACT
 AATACAGACGTACAGATACTTTCTCTGAAGAGTATTTTGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATGAC
 ACTTTCTGCTTTACAGAAAAGGAACTCATTACAGACTGGTGATATCGTGATGTACCTAAAAGTCAAGAAACCATTTT
 CTCCTCAGAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGTGTGAACCAACAATCTCTTTTC
 AAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATGGAGAAAAATATATATATATATATATATATTGT
 GAAAGATCAATCCATCTGCCAGAATCTAGTGGGATGGAAGTTTTGTACATGTTATCCACCCAGGCCAGGTGGAAG
 TAACTGAATATTTTTTAAATTAAGCAGTTCTACTCAATCACAAGATGCTTCTGAAAATTCGATTTTATTACCATTT
 CAAACTATTTTTTAAAAATAAATACAGTTAAACATAGAGTGGTTTCTTCATTTCATGTGAAAATTTATTAGCCAGCACCAG
 ATGCATGAGCTAATTATCTCTTTGAGTCCTTGCTTCTGTTGCTCACAGTAACTCATTGTTTAAAGCTTCAAGAAC
 ATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTGTACTGGTAGTTTATGAAATTTAATTAACACAGG
 CCATGAATGGAAGGTGGTATTCACAGCTAATAAAATATGATTTGTGGATATGAA

FIGURE 72

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEW
EQHRNYVSSLKRQIAQLKEELQERSEQLRNGQYQASDAAGLGLDRSPPEKTQADLLAFLHSQV
AEVNAGVKLATEYAAVPFDSFTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALET
ENSPNHRPYTASDFIEGIYRTERDKGTLYELTFKGDHKHEFKRLILFRPFSPIMKV
KNEKLNMAN
TLINVIPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVYFGKEEINEVKGILENTSKA
ANFRNF
TFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNLCRLNTQPGKKV
FYFVLF
SQY
NPGIYGHHDVAPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKGWGGED
VHLYR
KYLHSNLIVRTPVRGLFHLWHEKRCMDELTPQYKCMQSKAMNEASHGQLGMLVFRHE
IEAHL
RKQKQKTSSKKT

Important features:

Signal peptide:

amino acids 1-27

N-glycosylation sites.

amino acids 315-319, 324-328

N-myristoylation sites.

amino acids 96-102, 136-142, 212-218, 311-317, 339-345, 393-399

Amidation site.

amino acids 377-381

FIGURE 73

[illegible]

FIGURE 74

MLFSALLLEVIWILAADGGQHWTYEGPHGQDHPASYPECGNNAQSPIDIQTDSVTFDPLPALQ
PHGYDQPGTEPLDLHNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKSPGGSEHQINSEAT
FAELHIVHYDSYDSLSEAAERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPP
FNLRELLPKQLGQYFRYNGSLTTPPCYQSVLWTVFYRRSQISMEQLEKLQGTLFSTEEEPSKLLV
QNYRALQPLNQRMVFASFQAGSSYTTGEMLSLGVGILVGCLCLLLAVYFIARKIRKKRLENRKS
VVFTSAQATTEA

Important features of the protein:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 291-310

N-glycosylation site.

amino acids 213-216

Eukaryotic-type carbonic anhydrases proteins

amino acids 197-245, 104-140, 22-69

FIGURE 75

TGCCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTGTGTC
 TCTGGTGGTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGTCCTACA
 ATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTTTCATCACAA
 ATTGGCCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC
 AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCTGTTTCCATGCAACAAATATACTCCAATCTGA
 AGTATAACGTGCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAACCAC
 ACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTTCGTCCC
 AGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGATCAATCAT
 CAGAGTTCAAGGCTAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGTGTTCCTTTTT
 TCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACACCCAGCAAATTT
 GATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGAAAAAATCGTGATTA
 ACTTTATCACCCCTCAATATCTCGGATGATTCTAAAATTTCTCATCAGGATATGAGTTTACTGGGA
 AAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAACCTGAGGCCCCCTCAGGA
 GGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGAAATTTTTTGTGACTCTGAAG
 AAAACACGGAAGGTACTTCTCTCACCAGCAAGAGTCCCTCAGCAGAACAAATACCCCCGGATAAA
 ACAGTCATTGAATATGAATATGATGTCAGAACCACTGACATTTGTGCGGGGCCTGAAGAGCAGGA
 GCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATTATTGGAGTCGCAGGCAGCGTTGGCAG
 TCTTGGGCCCCGAAACGTTACAGTACTCATAACCCCTCAGCTCCAAGACTTAGACCCCCTGGCG
 CAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGAAGAGCCATCGACGACCCTGGTCGACTGGGA
 TCCCCAAACTGGCAGGCTGTGTATTCTTCGCTGTCCAGCTTCGACCAGGATTCAGAGGGCTGCG
 AGCCTTCTGAGGGGGATGGGCTCGGAGAGGAGGGTCTTCTATCTAGACTCTATGAGGAGCCGGCT
 CCAGACAGGCCACCAGGAGAAAATGAAACCTATCTCATGCAATTCATGGAGGAATGGGGTTATA
 TGTGCAGATGGAAAACTGATGCCAACACTTCCTTTTGCCCTTTTGTTCCTGTGCAACAAGTGAG
 TCACCCCTTTGATCCCAGCCATAAAGTACCTGGGATGAAAGAAGTTTTTCCAGTTTGTGAGTGT
 CTGTGAGAATTACTTATTTCTTTTCTCTATTCTCATAGCACGTGTGTGATTGGTTCATGCATGTA
 GGTCTCTTAACAATGATGGTGGGCCTCTGGAGTCCAGGGGCTGGCCGGTTGTTCTATGCAGAGAA
 AGCAGTCAATAAATGTTTGCCAGACTGGGTGCAGAAATTATTCAGGTGGGTGT

FIGURE 76

MSYNGLHQRVFKELKLLTLCSSQIGPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIY
SNLKYNVSVLNTKSNRTWSQCVTNHTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLK
DQSSEFKAKIIFWYVLPISITVFLFSVMGYSIYRYIHVGKEKHPANLILYGNFDRFFVPAEK
IVINFITLNISSDDSKISHQDMSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFC
DSEENTEGTSLTQQESLSRTIPDPKTVIEYEYDVRTTDCAGPEEQELSLQEEVSTQGTLLSQA
ALAVLGPQTLQYSYTPQLQDLPLAQEHTDSEEGPEEEPSTTLVDWDPQTGRLCIPSLSSFDQDS
EGCEPSEGDGLGEEGLLSRLYEAPDRPPGENETYLMQFMEEWGLYVQMEN

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 140-163

N-glycosylation sites.

amino acids 71-74, 80-83, 89-92, 204-207, 423-426

FIGURE 77

GAGGAGCGGGCCGAGGACTCCAGCGTGCCCGAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACAC
 CTGGGAAGATGCGCCGGCCCGTGACCTTCACCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATC
 CAAGCCACCCTCAGTCCCAGTTCAGTTCTCATCCTCGGCCCAAAGTCATCAAAGAAAAGCTGAC
 ACAGGAGCTGAAGGACCACAACGCCACCAGCATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGC
 GGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCTGGTGAACACCGTCCTGAAGCACATC
 ATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCCCTCGGCCAATGACCA
 GGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAAGACCA
 TCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGC
 CCCACCCGCTGGTCCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAAGTGTGTA
 TAAGCTCTCCTTCCTGGTGAACGCCTTAGCTAAGCAGGTTCATGAACCTCCTAGTGCCATCCCTGC
 CCAATCTAGTGAAAAACCAGCTGTGTCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGAC
 CTCCTGCAGCTGGTGAAGGTGCCCATTTCCTCAGCATTGACCGTCTGGAGTTTGACCTTCTGTA
 TCCTGCCATCAAGGGTGACACCATTAGCTCTACCTGGGGGCCAAGTTGTTGGACTCACAGGGAA
 AGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCTTGACAATGCCACCCTGGACAACATCCCCG
 TTCAGCCTCATCGTGAGTCAGGACGTGGTGAAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA
 ATTCATGGTCCTGTTGGACTCTGTGCTTCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGC
 TGATCAATGAAAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGAC
 ACTCCCGAGTTTTTTATAGACCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTT
 TCCCTCCAGTGAAGCCCTCCGCCCTTTGTTTACCCTGGGCATCGAAGCCAGCTCGGAAGCTCAGT
 TTTACACCAAAGGTGACCAACTTATACTCAACTGAATAACATCAGCTCTGATCGGATCCAGCTG
 ATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCATCACTGAGATCATCCA
 CTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGTGAAGG
 CCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGTGCTTACTCCAGCCTCC
 TTGTGGAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGATGGCAGCCATCAGGGAAGGCTGG
 GTCCCAGCTGGGAGTATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTG
 CCTGTGAAAAA

FIGURE 78

MAGPWTFLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDHNATSILQQLPLLSAMREK
PAGGIPVLGSLVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGENTPLVKTIVE
FHMTTEAQATIRMDTSASGPTRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNL
VKNQLCPVIEASFNGMYADLLQLVKVPISLSIDRLEFDLLYPAIKGDITQLYLGAKL LDSQGKVT
KWFNNSAASLTMP TLDNIPFSLIVSQDVVKA AVAVLSPEEFMVLLDSVLPESAHLKSSIGLIN
EKAADKLGSTQIVKILTQDTP EFFIDQGHAKVAQLIVLEVFPSS EALRPLFTLGIEASSEAQFYT
KGDQLILNLN NISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSLVKALG
FEAAESSLT KDALVLTPASLWKPPSPVSQ

Important features of the protein:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 48-51, 264-267, 401-404

Glycosaminoglycan attachment site.

amino acids 412-415

LBP / BPI / CETP family proteins.

amino acids 407-457

FIGURE 79

GAGAGAAGTCAGCCTGGCAGAGAGACTCTGAAATGAGGGATTAGAGGTGTTCAAGGAGCAAGAGC
 TTCAGCCTGAAGACAAGGGAGCAGTCCCTGAAGACGCTTCTACTGAGAGGTCTGCCATGGCCTCT
 CTTGGCCTCCAACCTTGTGGGCTACATCCTAGGCCTTCTGGGGCTTTTGGGCACACTGGTTGCCAT
 GCTGCTCCCCAGCTGGAAAACAAGTTCTTATGTGGTGCCAGCATTGTGACAGCAGTTGGCTTCT
 CCAAGGGCCTCTGGATGGAATGTGCCACACACAGCACAGGCATCACCCAGTGTGACATCTATAGC
 ACCCTTCTGGGCCTGCCCCTGACATCCAGGCTGCCCAGGCCATGATGGTGACATCCAGTGCAAT
 CTCCTCCCTGGCCTGCATTATCTCTGTGGTGGGCATGAGATGCACAGTCTTCTGCCAGGAATCCC
 GAGCCAAAGACAGAGTGGCGGTAGCAGGTGGAGTCTTTTTCATCCTTGGAGGCCTCCTGGGATTC
 ATTCTGTTGCCTGGAATCTTCATGGGATCCTACGGGACTTCTACTCACCCTGGTGCCTGACAG
 CATGAAATTTGAGATTGGAGAGGCTCTTTACTTGGGCATTATTTCTTCCCTGTTCTCCCTGATAG
 CTGGAATCATCCTCTGCTTTTCTGCTCATCCAGAGAAATCGCTCCAATACTACGATGCCTAC
 CAAGCCCAACCTCTTGCCACAAGGAGCTCTCCAAGGCCTGGTCAACCTCCCAAAGTCAAGAGTGA
 GTTCAATTCCTACAGCCTGACAGGGTATGTGTGAAGAACCAGGGGCCAGAGCTGGGGGGTGGCTG
 GGTCTGTGAAAAACAGTGGACAGCACCCCGAGGGCCACAGGTGAGGGACACTACCACTGGATCGT
 GTCAGAAGGTGCTGCTGAGGATAGACTGACTTTGGCCATTGGATTGAGCAAAGGCAGAAATGGGG
 GCTAGTGTAACAGCATGCAGGTTGAATTGCCAAGGATGCTCGCCATGCCAGCCTTTCTGTTTTCC
 TCACCTTGCTGCTCCCCCTGCCCTAAGTCCCCAACCCCTCAACTTGAAACCCCATTCCTTAAGCCA
 GGACTCAGAGGATCCCTTTGCCCTCTGGTTTACCTGGGACTCCATCCCCAAACCCACTAATCACA
 TCCCACTGACTGACCCTCTGTGATCAAAGACCCTCTCTCTGGCTGAGGTGGCTCTTAGCTCATT
 GCTGGGGATGGGAAGGAGAAGCAGTGGCTTTTGTGGGCATTGCTCTAACCTACTTCTCAAGCTTC
 CCTCCAAAGAACTGATTGGCCCTGGAACCTCCATCCCACTCTTGTTATGACTCCACAGTGTCCA
 GACTAATTTGTGCATGAACTGAAATAAAACCATCCTACGGTATCCAGGGAACAGAAAGCAGGATG
 CAGGATGGGAGGACAGGAAGGCAGCCTGGGACATTTAAAAAATA

FIGURE 8o

MASLGLQLVGYILGLLGLLGTLVAMLLPSWKTSSYVGASIVTAVGFSKGLWMECATHSTGITQCD
IYSTLLGLPADIQAAQAMMVTSSAISSLACIISVVGMRCTVFCQESRAKDRVAVAGGVFFILGGL
LGFIPVAWNLHGILRDFYSPLVPDSMKFEIGEALYLGIISSLFSLIAGIILCFSCSSQRNRSNYY
DAYQAQPLATRSSPRPGQPPKVKSEFNSYSLTGYV

Important features of the protein:

Signal peptide:

amino acids 1-24

Transmembrane domains:

amino acids 82-102, 117-140, 163-182

N-glycosylation site.

amino acids 190-193

PMP-22 / EMP / MP20 family proteins.

amino acids 46-59

FIGURE 81

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTT
 CCCGCGTTCTCTTTCCACCTTTCTCTTCTTCCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCCT
 GCCCACCGCTGCTTCTTCCGCCCCCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGG
 GTTGATCTGTGGCCCCCTGTGCCTCCGTGTCCTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGG
 ACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCGAGGTGAGGGTCCTCTCCTCCTTGCTGGGA
 CTCGCGCTGCTCTGGTTCCCCCTGGACTCCCACGCTCGAGCCCGCCCAGACATGTTCTGCCTTTT
 CCGTGGGAAGAGATACTCCCCCGGCGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCTGATGT
 ACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCTT
 GTCCACTGCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACAC
 TCCCTCTGGACTCCGGGCCCCACCAAAGTCCTGCCAGCACAAACGGGACCATGTACCAACACGGAG
 AGATCTTCAGTGCCCATGAGCTGTTCCCCTCCCGCCTGCCCAACCAGTGTGTCTCTGCAGCTGC
 ACAGAGGGCCAGATCTACTGCGGCCTCACAACCTGCCCCGAACCAGGCTGCCAGCACCCCTCCC
 ACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGAGCAATCGGATGAAGAGGACA
 GTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGCTGGGAGAAAG
 AGAGGCCCGGGCACCCAGCCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT
 CAGACCCAAGGGAGCAGGCAGCACAACTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCT
 GTGTGCATGGCGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGC
 CCCTTGCCCTGCATCCTATGCACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCC
 CACCGAGTACCCCTGCCGTACCCCGAGAAAGTGGCTGGGAAGTGCTGCAAGATTTGCCCAGAGG
 ACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAGGTGTCCCAAGGCACCGGGCCGGGTC
 CTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCGCTTTGCCCTGGAACACGAGGC
 CTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAACTGAGGCTCAGAGAG
 GTGAAGTACCTGGCCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAG
 GAAGCAAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCCACGAAGGTCCT
 GGAACGTCTTCCTAGCCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAG
 ACATAACAAAGACCTTAACAGTTGCAGATATGAGCTGTATAATTGTTGTTATTATATATTAATAAA
 TAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAAAAAAAAAA

FIGURE 82

MVPEVRVLSSLLGLALLWFPLDSHARARPDMFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSE
GAHVSCYRLHCPPVHCPQPVTEPQQCCPKCVEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELF
PSRLPNQCVLCSCTEGQIYCGLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVR
HPQDPCSSDAGRKRGPPTAPTGLSAPLSFIPRHFRPKGAGSTTVKIVLKEKHKKACVHGGKTY
HGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPTTEYPCRHPKVKAGKCKICPEDKADPGHSE
ISSTRCPKAPGRVLVHTSVSPSPDNLRFALEHEASDLVEIYLWKLVKDEETEAQRGEVPGPRPH
SQNLPLDSDQESQEARLPERGTALPTARWPPRRSLERLPSDPGAEGHGQSRQSDQDITKT

Signal peptide:

amino acids 1-25

FIGURE 83

GACAGCTGTGTCTCGATGGAGTAGACTCTCAGAACAGCGCAGTTTGCCCTCCGCTCAGCGCAGAGCCTCTCC
 GTGGCTTCCGCACCTTGAGCATTAGGCCAGTTCTCCTCTTCTCTCTAATCCATCCGTACCTCTCCTGTCA
 TCCGTTTCCATGCCGTGAGGTCCATTACAGAACACATCCATGGCTCTCATGCTCAGTTTGGTTCTGAGTC
 TCCTCAAGCTGGGATCAGGGCAGTGGCAGGTGTTTGGGCCAGACAAGCCTGTCCAGGCCTTGGTGGGGAG
 GACGCAGCATTCTCCTGTTTCTGTCTCCTAAGACCAATGCAGAGGCCATGGAAGTGCGGTCTTTCAGGGG
 CCAGTTCTCTAGCGTGGTCCACCTCTACAGGGACGGGAAGGACCAGCCATTTATGCAGATGCCACAGTATC
 AAGGCAGGACAAAACCTGGTGAAGGATTCTATTGCGGAGGGGCGCATCTCTCTGAGGCTGGAAAACATTACT
 GTGTTGGATGCTGGCCTCTATGGGTGCAGGATTAGTTCCCAGTCTTACTACCAGAAGGCCATCTGGGAGCT
 ACAGGTGTCAGCACTGGGCTCAGTTCCTCTCATTTCCATCACGGGATATGTTGATAGAGACATCCAGCTAC
 TCTGTCACTCCTCGGGCTGGTTCCCCCGGGCCACAGCGAAGTGGAAAGGTCCACAAGGACAGGATTTGTCC
 ACAGACTCCAGGACAAACAGAGACATGCATGGCCTGTTTGATGTGGAGATCTCTCTGACCGTCCAAGAGAA
 CGCCGGGAGCATATCCTGTTCCATGCGGCATGCTCATCTGAGCCGAGAGGTGGAATCCAGGGTACAGATAG
 GAGATACCTTTTTCGAGCCTATATCGTGGCACCTGGCTACCAAAGTACTGGGAATACTCTGCTGTGGCCTA
 TTTTTTGGCATTGTTGGACTGAAGATTTTCTTCTCCAAATTCAGTGGAAAATCCAGGCGGAACTGGACTG
 GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCCGAAACACGCAGTGGAGGTGACTCTGGATCCAG
 AGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAACCTGTAACCCATAGAAAAGCTCCCCAGGAGGTG
 CCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCAGAGTTTCCAAGCAGGGAACATTA
 CTGGGAGGTGGACGGAGGACACAATAAAAGGTGGCGCGTGGGAGTGTGCCGGGATGATGTGGACAGGAGGA
 AGGAGTACGTGACTTTGTCTCCCGATCATGGGTACTGGGTCTCAGACTGAATGGAGAACATTTGTATTTT
 ACATTAAATCCCCGTTTTATCAGCGTCTTCCCCAGGACCCACCTACAAAATAGGGGTCTTCTTGACTA
 TGAGTGTGGGACCATCTCCTTCTTCAACATAAATGACCAGTCCCTTATTTATACCCTGACATGTCGGTTTG
 AAGGCTTATTGAGGCCCTACATTGAGTATCCGTCCCTATAATGAGCAAAATGGAACCTCCCATAGTCATCTGC
 CCAGTCACCCAGGAATCAGAGAAAGAGGCCTCTTGGCAAAGGGCCTCTGCAATCCAGAGACAAGCAACAG
 TGAGTCCTCCTCACAGGCAACCACGCCCTTCTCCCCAGGGGTGAAATGTAGGATGAATCACATCCCACAT
 TCTTCTTTAGGGATATTAAAGTCTCTCTCCCAGATCCAAAGTCCCGCAGCAGCCGGCCAAGGTGGCTTCCA
 GATGAAGGGGGACTGGCCTGTCCACATGGGAGTCAGGTGTCATGGCTGCCCTGAGCTGGGAGGGAAGAAGG
 CTGACATTACATTTAGTTTGCTCTCACTCCATCTGGCTAAGTGATCTTGAAATACCACCTCTCAGGTGAAG
 AACCGTCAGGAATTCCTATCTCACAGGCTGTGGTGTAGATTAAAGTAGACAAGGAATGTGAATAATGCTTAG
 ATCTTATTGATGACAGAGTGTATCCTAATGGTTTGTTCATTATATTACATTTTCAGTAAAAAAA

FIGURE 84

MAIMLSLVLSLLKLGSGQWQVFGPDKPVQALVGEDAAFSCLSPKTNAEAMEVRFFRGQFSSVVH
LYRDGKDQPFMQMPQYQGRTKLVKDSIAEGRISLRLENITVLDAGLYGCRISQSYQKAIWELQ
VSALGSVPLISITGYVDRDIQLLCQSSGWFPPTAKWKGPQGQDLSTDSRTNRDMHGLFDVEISL
TVQENAGSISCSMRHAHLSREVESRVQIGDTFFEPISWHLATKVLGILCCGLFFGIVGLKIFFSK
FQWKIQAELDWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQSFQAGKHYWEVDGGHNKRWRVGVCRDDVDRRKEYVTLSPDHGYWVLRNLNGEHLFT
LNPRFISVFPRTPTKIGVFLDYECGTISFFNINDQSLIYTLTCRFEGLLRPYIEYPSYNEQNGT
PIVICPVTQESEKEASWQRASAIPESTNSSESSQATTPLPRGEM

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 239-255

FIGURE 85

AACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCAGACATGCTGCTGCTGCTGCTGCCCT
 GCTCTGGGGGAGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCCGTGA
 CGGTGCAGGAAGGCCTGTGTGTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATT
 TACCCTGGCCCAGTAGTTTCATGGCTACTGGTTCCGGGAAGGGGCCAATACAGACCAGGATGCTCC
 AGTGGCCACAAACAACCCAGCTCGGGCAGTGTGGGAGGAGACTCGGGACCGATTCCACCTCCTTG
 GGGACCCACATACCAAGAATTGCACCCTGAGCATCAGAGATGCCAGAAGAAGTGATGCGGGGAGA
 TACTTCTTTCGTATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCTGTGAA
 TGTGACAGCCTTGACCCACAGGCCCAACATCCTCATCCCAGGCACCCTGGAGTCCGGCTGCCCC
 AGAATCTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATA
 GGGACCTCCGTGTCCCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCTCATCCCACA
 GCCCCAGGACCATGGCACCAGCCTCACCTGTGAGGTGACCTTCCCTGGGGCCAGCGTGACCACGA
 ACAAGACCGTCCATCTCAACGTGTCCTACCCGCCTCAGAACTTGACCATGACTGTCTTCCAAGGA
 GACGGCACAGTATCCACAGTCTTGGGAAATGGCTCATCTCTGTCACTCCCAGAGGGCCAGTCTCT
 GCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGCCTGAGCTGGA
 GAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC
 CTGAGGGATGCAGCTGAATTCACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCT
 GAACGTCTCCCTGCAGAGCAAAGCCACATCAGGAGTGACTCAGGGGGTGGTCGGGGGAGCTGGAG
 CCACAGCCCTGGTCTTCCTGTCCTTCTGCGTCATCTTCGTTGTAGTGAGGTCCTGCAGGAAGAAA
 TCGGCAAGGCCAGCAGCGGGCGTGGGAGATACGGGCATAGAGGATGCAAACGCTGTGAGGGGTTC
 AGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGACAGTCCCCCAGACCAGCCTCCCCCAG
 CTTCTGCCCCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTCAGCTTCCAGATGGTG
 AAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATCCACAG
ATGAGAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCCTCCAGGCAAGGGAGAAGTCA
 GAGGCTGATTCTTGTAGAATTAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATG
 TGCAGAGTGAAAAGCACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCC
 TCCCTTTTATTTTTTTAACTAAAAGACAGACAAATTCCTA

FIGURE 86

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFREG
ANTDQDAPVATNNPARAVWEETRDRFHLLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNY
KHHRLSVNVTALTHRPNILIPGTLESQCPQNLTCVWPWACEQGTTPPMISWIGTSVSPDPSTTRS
SVLTLLIPQPQDHGTSLTTCQVTFPGASVTTNKTVHLNVSYPPQNLMTVTFQGDGTVSTVLGNGSSL
SLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPSNPGVLELPWVHLRDAAEFTCRAQNP
LGSQQVYLNVSLSQKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRKKSARPAAGVGDGTGIE
DANAVRGSASQGPLTEPWAEDSPPDQPPPASARSSVGEGLQYASLSFQMKPWDSRGQEATDTE
YSEIKIHR

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 351-370

FIGURE 87

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTAC
CCAAGGAAAGTGAGCTGAGACTCAGACAAGATTACAATGAACCAACTCAGCTTCCTGCTGTTTC
TCATAGCGACCACCAGAGGATGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGT
TCTTCGTCTCCATCTCTGCCCAGAAGCTGCAAGGAAATCAAAGACGAATGTCTAGTGCATTTGA
TGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTACCAGACCTTCTGTGACATGACCTCTG
GGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCATGAGAATGACATGCGTGGGAAGTGCACGGTG
GGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAACTGGGC
CAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACT
ACGACATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCAATAAGTCCCCCATGCAGCACTGG
AGAAACAGCTCCCTGCTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCT
GTTTGGCATCTACCAGAAATATCCAGTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCC
CGGTGATCCCTGTGGTCTATGATTTTGGCGACGCCCAGAAAACAGCATCTTATTACTCACCTAT
GGCCAGCGGGAATTCAGTGCGGGATTTGTTTCAGTTCAGGGTATTTAATAACGAGAGAGCAGCCAA
CGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATTGGTGGAGGAG
GATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT
GGAATCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCG
TTGAGAGTTTTGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAA
CAACTTACCCAGTAGCTAGAATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGA
AAAAAA

FIGURE 88

MNQLSFLLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVI
YQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAAT
SDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLEFGIYQKYPVKYG
EGKCWTDNGPVIPIVVYDFGDAQKTASYSPYQGREFTAGFVQFRVFNNERAANALCAGMRVTGCN
TEHHCIGGGGYFPEASPOQCGDFSGFDWSGYGTHVGYSRSSREITEAAVLLFYR

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation site.

amino acids 163-167

Glycosaminoglycan attachment sites.

amino acids 74-78, 289-293

N-myristoylation sites.

amino acids 76-82, 115-121, 124-130, 253-259, 292-298

FIGURE 89

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGAC
CGCCGCCCTTGTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACG
CTCCTGGCGCATCTGGTGGTCGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTG
CCTGCCTCTCACGTTACCCCCGAGGAGTATGACAAGCAGGACATTACAGCTGGTGGCCGCGCTCT
CTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTTCCTCTCAGGAGTCTCCATGTTCAAC
AGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGCCCTGTCCTTCTTCAT
ATTCGAGCGTTGGGAGTGCACTACGTATTGGTACATTTTGTCTTCTGCAGTGCCCTTCCAGCTG
TCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCCTTCTTGATTACCTTCA
TGACGGGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAG
GCATAGGCTTCGGTTTTCCCCTCGGAAACTGCTTCTGCTGGAGGATATGTGTTGGAATAATTACG
TCTTGAGTCTGGGATTATCCGCATTGTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACA
TTAAGACTTATATACAGTTTTAGGGGACAATTAAAAAAAAAAAAA

FIGURE 90

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFA
VELAGFLSGVSMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCALPAVTEMALFV
TVFGLKKKPF

Transmembrane domain:

amino acids 12-28 (type II), 51-66, 107-124

FIGURE 91

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGAC
 TCGCTGCTGCTTCGTGTTCCCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATG
 GTCCTCCCGGCTCAGAGGACCCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCT
 CGGAAGCGGGGCCACATCTCACCTAAGTCCCGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCT
 GGCCCCGCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCCCCAACCGCCCGAACCACAGCCCC
 CACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCCAACATCAAGACGGTG
 GCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTCCACTT
 CCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTGCCCCCAGTAAAGCTGTAG
 AGTTCCACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATG
 GAGTGGGAGAAGGTAGAACGGGGCCGCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTG
 CTCCCGAGACCACGCTCAGAGCTCAGCCACCTGGAGCTGCTCCAGCCCTTCAAAGTCGTCTGTG
 TCTACATCGCCTTCTACAGCACGGACTATCGGCTGGTCCAGAAGGTGTGCCAGATTACAACCTAC
 CATAGTGATACCCCTACTACCCATCTGGGTGACCCCGGGGCAGGCCACAGAGGCCAGGCCAGGGC
 TGGAAGGACAGGCCTGCCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGGGTGGGCCTC
 AGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG
 AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGG
 AGGAGGAGTGGGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGG
 TCCCCGAGGCCTGTGGGCAGGCCGATCAGTGTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGC
 CCTTGGTTCTTGCCATCCTGAGGAAAGATAGCAACAGGGAGGGGGAGATTTTCATCAGTGTGGACA
 GCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAGGAGCCAGTCAGCAGGGTGGGGTGGG
 GCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTTGTCGTGTGCTGAGCATGG
 CATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCTCCAGC
 CAGGCCACCCCTTTCCAAAATTCCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATG
 GCACACCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATC
 CGCGTGCTGTGTGTCCCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCG
 GAGAGGGGTCCCTCAACAGTCAGCCTCACCTGTGACCGGGGTCTCTCCGGATCTGGATGGCGC
 CGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGGCCGGCCGAGAGCATGTGCTGGATCTGTTC
 TGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGAAACCGCTGATTGCTGACTTT
 TGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

FIGURE 92

MQLTRCCFVFLVQGSLYLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTL
LGLLAPPGEAWGILGPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTF
SVHFQHNATGQGNISISLVPPSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDP
AKICSRDHAQSSATWSCSQPFKVVCVYIAFYSTDYRLVQKVCPDYNHSDTPYYPSG

Important features of the protein:

Signal peptide:

amino acids 1-14

N-glycosylation sites.

amino acids 62-65, 127-130, 137-140, 143-146

2-oxo acid dehydrogenases acyltransferase

amino acids 61-71

FIGURE 93

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCCTGCGCTCGCC
CTTTATGTCTTCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTT
CTGGTTGGTGTCTCTACTGATTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACA
AAGATGGACCAACACAGAAATATCTGCTGATCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAA
ATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCAGTGAAGGTTTGAAGAGTATAAACCC
AGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGGGCTTTGGAATCATGA
GTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCATTCAT
GGAGATTCTCCTCAATTCTTCCTTTATTTCAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGT
ATTCTGGGGCATTGTATTTTTTGATGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTC
TCCTGACCCACCTGCTGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCG
TCAGCATTTATAATCCTGGTGTGCTCATGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCG
AAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGAACTTTCTTCTTTACAACCAGCGCTCCAGAT
AACTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTAGAGGAAGCACAACTGTGCCT
TTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

FIGURE 94

MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFEMARVIIDNKDG
PTQKYLLIFGAFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLLAYVSGLGFGIMSGV
FSFVNTLSDSLGPGTVGIHGDSPOFFLYSAFMTLVIILLHVFWGIVFFDGCEKKKWGILLIVLLT
HLLVSAQTFISSYYGINLASAFIILVLMGTWAFLAAGGSCRSLKLCLLCQDKNFFLYNQSR

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domains:

amino acids 32-51, 119-138, 152-169, 216-235

Glycosaminoglycan attachment site.

amino acids 120-123

Sodium:neurotransmitter symporter family protein

amino acids 31-65

FIGURE 95

AATTTTTCACCAGAGTAAACTTGAGAAACCAACTGGACCTTGAGTATTGTACATTTTGCCTCGTG
GACCCAAAGGTAGCAATCTGAAACATGAGGAGTACGATTCTACTGTTTTGTCTTCTAGGATCAAC
TCGGTCATTACCACAGCTCAAACCTGCTTTGGGACTCCCTCCCACAAAAGTGGCTCCGGATCAGG
GAACACTACCAAACCAACAGCAGTCAAATCAGGTCTTTCCTTCTTTAAGTCTGATACCATTAACA
CAGATGCTCACACTGGGGCCAGATCTGCATCTGTTAAATCCTGCTGCAGGAATGACACCTGGTAC
CCAGACCCACCCATTGACCCTGGGAGGGTTGAATGTACAACAGCAACTGCACCCACATGTGTTAC
CAATTTTTGTACACAACCTGGAGCCCAGGGCACTATCCTAAGCTCAGAGGAATTGCCACAAATC
TTCACGAGCCTCATCATCCATTTCCTTGTTCCCGGGAGGCATCCTGCCACCCAGTCAGGCAGGGGC
TAATCCAGATGTCCAGGATGGAAGCCTTCCAGCAGGAGGAGCAGGTGTAAATCCTGCCACCCAGG
GAACCCAGCAGGCCGCCTCCCAACTCCCAGTGGCACAGATGACGACTTTGCAGTGACCACCCCT
GCAGGCATCCAAAGGAGCACACATGCCATCGAGGAAGCCACCACAGAATCAGCAAATGGAATTCA
GTAAGCTGTTTCAAATTTTTTCAACTAAGCTGCCTCGAATTTGGTGATACATGTGAATCTTTATC
ATTGATTATATTATGGAATAGATTGAGACACATTGGATAGTCTTAGAAGAAATTAATTCTTAATT
TACCTGAAAATATTCTTGAAATTTAGAAAATATGTTCTATGTAGAGAATCCCAACTTTTAAAAA
CAATAATTCAATGGATAAATCTGTCTTTGAAATATAACATTATGCTGCCTGGATGATATGCATAT
TAAACATATTTGGAAAAGTGGAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 96

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQQSNQVFPSLSLIPLTQM
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTTLSSEE
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGS LPAGGAGVNPATQGTPAGRLPTPSG
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

Signal peptide:

amino acids 1-16

FIGURE 97

GCTCAAGTGCCCTGCCTTGCCCCACCCAGCCAGCGCTGGCCAGAGCCCTCTGGAGAAGGAGCTCT
CTTCTTGCTTGGCAGCTGGACCAAGGAGCCAGTCTTGGGCGCTGGAGGGCCTGTCTTGACCAATG
GTCCCTGCCTGGCTGTGGCTGCTTTGTGTCTCCGTCCCCCAGGCTCTCCCCAAGGCCCAGCCTGC
AGAGCTGTCTGTGGAAGTTCAGAAAACATATGGTGGAAATTTCCCTTTATACCTGACCAAGTTGC
CGCTGCCCCGTGAGGGGGCTGAAGGCCAGATCGTGCTGTGAGGGGACTCAGGCAAGGCAACTGAG
GGCCCATTTGCTATGGATCCAGATTCTGGCTTCTGTGCTGGTGACCAGGGCCCTGGACCGAGAGGA
GCAGGCAGAGTACCAGCTACAGGTCACCCTGGAGATGCAGGATGGACATGTCTTGTGGGGTCCAC
AGCCTGTGCTTGTGCACGTGAAGGATGAGAATGACCAGGTGCCCCATTTCTCTCAAGCCATCTAC
AGAGCTCGGCTGAGCCGGGGTACCAGGCCTGGCATCCCCCTTCTCTTCTTGGAGCTTCAGACCG
GGATGAGCCAGGCACAGCCAACTCGGATCTTCGATTCCACATCCTGAGCCAGGCTCCAGCCCAGC
CTTCCCCAGACATGTTCCAGCTGGAGCCTCGGCTGGGGGCTCTGGCCCTCAGCCCCAAGGGGAGC
ACCAGCCTTGACCACGCCCTGGAGAGGACCTACCAGCTGTTGGTACAGGTCAAGGACATGGGTGA
CCAGGCCTCAGGCCACCAGGCCACTGCCACCGTGGAAGTCTCCATCATAGAGAGCACCTGGGTGT
CCCTAGAGCCTATCCACCTGGCAGAGAATCTCAAAGTCCTATACCCGCACCACATGGCCCAGGTA
TACTGGAGTGGGGGTGATGTGCACTATCACCTGGAGAGCCATCCCCCGGGACCCTTTGAAGTGAA
TGAGAGGGGAAACCTCTACGTGACCAGAGAGCTGGACAGAGAAGCCCAGGCTGAGTACCTGCTCC
AGGTGCGGGCTCAGAATTCCTATGGCGAGGACTATGCGGGCCCTCTGGAGCTGCACGTGCTGGTG
ATGGATGAGAAATGACAACGTGCCTATCTGCCCTCCCCGTGACCCCAAGTCAGCATCCCTGAGCT
CAGTCCACCAGGTACTGAAGTGACTAGACTGTGAGCAGAGGATGCAGATGCCCCCGGCTCCCCCA
ATTCCCACGTTGTGTATCAGCTCCTGAGCCCTGAGCCTGAGGATGGGGTAGAGGGGAGAGCCTTC
CAGGTGGACCCCACTTCAGGCAGTGTGACGCTGGGGGTGCTCCCACTCCGAGCAGGCCAGAACAT
CCTGCTTCTGGTGTGCTGGCCATGGACCTGGCAGGCGCAGAGGGTGGCTTCAGCAGCACGTGTGAAG
TCGAAGTCGCAGTCACAGATATCAATGATCACGCCCTGAGTTCATCACTTCCCAGATTGGGCCCT
ATAAGCCTCCCTGAGGATGTGGAGCCCGGGACTCTGGTGGCCATGCTAACAGCCATTGATGCTGA
CCTCGAGCCCGCCTTCCGCTCATGGATTTTGCCATTGAGAGGGGAGACACAGAAGGGACTTTTG
GCCTGGATTGGGAGCCAGACTCTGGGCATGTTAGACTCAGACTCTGCAAGAACCTCAGTTATGAG
GCAGCTCCAAGTCATGAGGTGGTGGTGGTGGTGCAGAGTGTGGCGAAGCTGGTGGGGCCAGGCCC
AGGCCCTGGAGCCACCGCCACGGTGACTGTGCTAGTGGAGAGAGTGATGCCACCCCCCAAGTTGG
ACCAGGAGAGCTACGAGGCCAGTGTCCCCATCAGTGCCCCAGCCGGCTCTTCTCTGCTGACCATC
CAGCCCTCCGACCCCATCAGCCGAACCCCTCAGGTTCTCCCTAGTCAATGACTCAGAGGGCTGGCT
CTGCATTGAGAAATTTCTCCGGGGAGGTGCACACCGCCAGTCCCTGCAGGGCGCCCAGCCTGGGG
ACACCTACACGGTGCTTGTGGAGGGCCAGGATACAGCCCTGACTCTTGCCCTGTGCCCTCCCAA
TACCTCTGCACACCCCGCCAAGACCATGGCTTGATCGTGAGTGGAACCCAGCAAGGACCCCGATCT
GGCCAGTGGGCACGGTCCCTACAGCTTCACCCCTTGGTCCCAACCCACGGTGCAACGGGATTGGC
GCCTCCAGACTCTCAATGGTTCCCATGCCTACCTACCTTGGCCCTGCATTGGGTGGAGCCACGT
GAACACATAATCCCCGTGGTGGTGCAGCCACAATGCCAGATGTGGCAGCTCCTGGTTCGAGTGAT
CGTGTGTGCTGCAACGTGGAGGGGCAGTGATGCGCAAGGTGGGCGCATGAAGGGCATGCCCA
CGAAGCTGTGCGCAGTGGGCATCCTTGTAGGCACCCTGGTAGCAATAGGAATCTTCTCATCCTC
ATTTTACCCACTGGACCATGTCAAGGAAGAAGGACCCGGATCAACCAGCAGACAGCGTGCCCT
GAAGGCGACTGTCTGAATGGCCAGGCAGCTCTAGCTGGGAGCTTGGCCTCTGGCTCCATCTGAG
TCCCCTGGGAGAGAGCCAGCACCCCAAGATCCAGCAGGGGACAGGACAGAGTAGAAGCCCCTCCA
TCTGCCCTGGGGTGGAGGCACCATCACCATCACCAGGCATGTCTGCAGAGCCTGGACACCAACTT
TATGGACTGCCCATGGGAGTGCTCCAAATGTGAGGGTGTGTGCCCAATAATAAAGCCCCAGAGAA
CTGGGCTGGGCCCTATGGGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAG

FIGURE 98

MVPAWLWLLCVSVPQALPKAQPAELSVEVPENYGGNFPLYLTKLPLPREGAEGQIVLSGDSGKAT
EGPFAMDPDSGFLLVTRALDREEQAEYQLQVTLEMQDGHVLWGPQPVLVHVKDENDQVPHFSAI
YRARLSRGTRPGIPFLFLEASDRDEPGTANSDLRFHILSQAPAQPSDFMQLEPRLGALALSPKG
STSLDHALERTYQLLVQVKDMGDQASGHQATATVEVSI IESTWVSLEPIHLAENLKVLYPHHMAQ
VHWSGGDVHYHLESHPPGPFEVNAEGNLYVTRELDREAQAEYLLQVRAQNSHGEDIAAPLELHVL
VMDENDNVPICPPRDPTVSIPELSPPGTEVTRLSAEDADAPGSPNSHVYQLLSPEPEDGVEGRA
FQVDPTSGSVTLGVLPLRAGQNILLVLAMDLAGAEGGSSTCEVEVAVTDINDHAPEFITSQIG
PISLPEDVEPGTLVAMLTADLEPAFRLMDFAIERGDTEGTFGLDWEPDSGHVRLRLCKNLSY
EAAPSHEVVVVVQSVAKLVGPGPGGATATVTVLVERVMPPPKLDQESYEASVPI SAPAGSFLLT
IQPSDPISRTLRFSLVNDSEGWLCIEKFSGEVHTAQSLQGAQPGDTYTVLVEAQDTALT LAPVPS
QYLCTPRQDHGLIVSGPSKDPDLASGHGPYSFTLGPNPTVQRDWRLQTLNGSHAYLTALHWVEP
REHIIPVVVSHNAQMWQLLVRVIVCRCNVEGQCMRKVGRMKGMPTKLSAVGILVGT LVAIGIFLI
LIETHWTMSRKKDPDQPADSVPLKATV

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 762-784

FIGURE 99

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCCACGCCTG
 AGTCCAAGATTCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTA
 TCTCTTCACCTTCAAGTCCCCTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTAC
 ATCTAGGACCCAGGCATCTTGCTTTCCAGCCACAAAGAGACAGATGAAGATGCAGAAAGGAAATG
 TTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGCAACAAATTCCAATGAGACTAGCACC
 TCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCACCAACTCTGGGTCCAG
 TGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGGGGTCA
 GCATAGTCACCAACTCTGAGTTCCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAG
 TTCAGCACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGG
 GGCCAGCACAGCCACCAACTCTGAGTCCAGCACACCCCTCCAGTGGGGCCAGCACAGTCCACCAACT
 CTGGGTCCAGTGTGACCTCCAGTGGAGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCC
 AGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAGCACACTCTCCAGTGGGGCCAGCACAGCCAC
 CAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACT
 GCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG
 AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCA
 GCACAGCCACCAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAG
 TCCAGCACGACCTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG
 GGCTAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCGGCACAGCCACCAACT
 CTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACACCCTCC
 AGTGGGGCCAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGGGGCCAACACAGCCAC
 CAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACAA
 CCTCCAGTGGGGTCCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACA
 GCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCTAG
 CACAGTGTCCAGTGGGATCAGCACAGTCCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCA
 ACACAGCCACCAACTCTGGGTCCAGTGTGACCTCTGCAGGCTCTGGAACAGCAGCTCTGACTGGA
 ATGCACACAACTTCCCATAGTGCATCTACTGCAGTGAGTGAGGCAAAGCCTGGTGGGTCCCTGGT
 GCCGTGGGAAATCTTCCTCATCACCTGGTCTCGGTTGTGGCGGCCGTGGGGCTCTTTGCTGGGC
 TCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGTCTACCACCCT
 CATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG
 GTGGAGTCCTAACTGGTTCTGGAGGAGACCAGTATCATCGATAGCCATGGAGATGAGCGGGAGGA
 ACAGCGGGCCCCTGAGCAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCA
 CCCAAGACCTGGTTTCTTTTCATTTCATCCCAGGAGACCCCTCCCAGCTTTGTTTGAGATCCTGAA
 AATCTTGAAGAAGGTATTCTTCACCTTTCTTGCCTTTACCAGACACTGGAAAGAGAATACTATAT
 TGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACACACACGACAAAGAGAAGCTGTGCTTG
 CCCCAGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAACCTCCATGCTGGACTC
 CATCTGGCATTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAAAAAA
 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 100

MKMQKGNVLLMFGLLLHLEAATNSNETSTSANTGSSVISSGASTATNSGSSVTSSGVSTATISGS
SVTSNGVSIIVTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESSTPSS
GASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTSSGASTA
TNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGAGTATNSES
STTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTATNSDSSTTSS
GAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNSESSTVSSGASTA
TNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESSTVSSGISTVTNSES
STTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGSLVPWEIFLITLVSVVA
AVGLFAGLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAPHRPRWSPNWFRRPVSSI
AMEMSGRNSGP

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 510-532

FIGURE 101

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCCTA
 CGGAGCCCCAGCTTGCCACGCACCCCACTCGGCGTCGCGCGGCGTGCCCTGCTTGTCACAGGTG
 GGAGGCTGGAACATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCAACAAAT
 GGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCATGGTCAATT
 AACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCTTTGGAATCAT
 GGTGTCATGGAAAGGGATTTACTTTTACTGACTCTGTTTTGGGGAAGCTTTTTTGAAGCATTT
 TCATGCTGAGTCCCTTTTTACCTTTGATGTTTGTAACCCATCTTGGTATCGCTGGATCAACAAC
 CGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCATGTTTGGTGTA AAAAGT
 GATTATAACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCATTATCATGAACCATCGGACAA
 GAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCTACCTCAGATTGGAGAAAATT
 TGCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTGGTGGGCCATGCAGGCTGCTGCCTATAT
 CTTCAATTCATAGGAAATGGAAGGATGACAAGAGCCATTTCTGAAGACATGATTGATTACTTTTGTG
 ATATTCACGAACCACTTCAACTCCTCATATTCCCAGAAGGGACTGATCTCACAGAAAACAGCAAG
 TCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTCAGAAATATGAATATGTTTTACATCCAAG
 AACTACAGGCTTTACTTTTGTGGTAGACCGTCTAAGAGAAGGTAAGAACCTTGATGCTGTCCATG
 ATATCACTGTGGCGTATCCTCACAACATTCTCAATCAGAGAAGCACCTCCTCCAAGGAGACTTT
 CCCAGGGAAATCCACTTTCACGTCCACCGGTATCCAATAGACACCCTCCCCACATCCAAGGAGGA
 CCTTCAACTCTGGTGCCACAAACGGTGGGAAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAG
 GGGAGAAGAATTTTTATTTTACCGGACAGAGTGTCAATCCACCTTGCAAGTCTGAACCTCAGGGTG
 CTTGTGGTCAAATTGCTCTCTATACTGTATTGGACCCTGTTAGCCCTGCAATGTGCCTACTCAT
 ATATTTGTACAGTCTTGTTAAGTGGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGA
 GAATATTTGGTGGACTGGAGATCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACAT
 TTAAATTCAAAGAAAAATGAGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTG
 GAAATGTTCTAAACCTTTCTAAGCTCAGATGCATTTTTGCATGACTATGTCGAATATTTCTTACT
 GCCATCATTATTTGTTAAAGATATTTTGCATTAAATTTTGTGGGAAAAATATTGCTACAATTTTT
 TTTAATCTCTGAATGTAATTTGATACTGTGTACATAGCAGGGAGTGATCGGGGTGAAATAACTT
 GGGCCAGAATATTATTAAACAATCATCAGGCTTTTAAA

FIGURE 102

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSGIMVSWKGIYFILTLFWGSFFGSIEMLS
FLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTRMDWM
FLWNCLMRYSYLRLEKICLKASLKGVPFGFWAMQAAAYIFIHRKWKDDKSHFEDMIDYFCDIHEP
LQLLIFPEGTDLTENSRSNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLDAVHDITVA
YPHNIPQSEKHLQGDFFPREIHFFVHRYPIDTLPTSKEDLQLWCHKRWEEKEERLRSFYQGEKNF
YFTGQSVIPPCKSELRLVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITIVIFVLQERIFGG
LEIIEELACYRLLHKQPHLNSKKNE

Important features of the protein:

Signal peptide:

amino acids 1-22

Transmembrane domains:

amino acids 44-63, 90-108, 354-377

FIGURE 103

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGACC
TCCAAATCATCCATCCACCCCTGCTGTCATCTGTTTTTCATAGTGTGAGATCAACCCACAGGAATA
TCCATGGCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTGGCAAGT
CACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTGCTCCCTCT
TTCCTGAGACCAGTGCAGAGGCTATGGAAGTGCAGTTTCTCAGGAATCAGTTCCATGCTGTGGTC
CACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCGAGGGAGAAGTGA
GTTTGTGAAGGACTCCATTGCAGGGGGGCGTGTCTCTAAGGCTAAAAACATCACTCCCTCGG
ACATCGGCCTGTATGGGTGCTGGTTCAGTTCCAGATTTACGATGAGGAGGCCACCTGGGAGCTG
CGGTGGCAGCACTGGGCTCACTTCTCTCATTTCATCGTGGGATATGTTGACGGAGGTATCCA
GTTACTCTGCCTGTCTCAGGCTGGTTCCTCCAGCCACAGCCAAGTGGAAAGGTCCACAAGGAC
AGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAGCCTGTATGATGTGGAGATCTCC
ATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCATCCACCTTGCTGAGCAGAGTCATGA
GGTGGAAATCCAAGGTATTGATAGGAGAGACGTTTTTCCAGCCCTCACCTTGGCGCCTGGCTTCTA
TTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGTGTGTCATGGGGATGATAATTGTTTTCTTC
AAATCCAAAGGAAAAATCCAGGCGGAACCTGGACTGGAGAAGAAAGCACGGACAGGCAGAATTGAG
AGACGCCCGGAAACACGCAGTGGAGGTGACTCTGGATCCAGAGACGGCTCACCCGAAGCTCTGCG
TTTTCTGATCTGAAAAGCTGTAACCCATAGAAAAGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGA
TTTACAAGGAAGAGTGTGGTGGCTTCTCAGGGTTTTCCAAGCAGGGAGACATTACTGGGAGGTGGA
CGTGGGACAAAATGTAGGGTGGTATGTGGGAGTGTGTGCGGATGACGTAGACAGGGGGAAGAACA
ATGTGACTTTGTCTCCCAACAATGGGTATTGGGTCTCAGACTGACAACAGAACATTTGTATTTCT
ACATTCAATCCCCATTTTATCAGCCTCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTCT
GGACTATGAGGGTGGGACCATCTCCTTCTTCAATACAAATGACCAGTCCCTTATTATACCCTGC
TGACATGTCAGTTTGAAGGCTTGTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAG
GGGACTCCCATATTTCATATGTCCAGTGTCTGGGGATGAGACAGAGAAGACCCCTGCTTAAAGGGC
CCCACACCACAGACCCAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCT
CCGGAGCCTGCGCACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCC
TGAGCCCTGCAGCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAG
TCAGAAGCCATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAA
CTCCATCCAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTTGCTAGTCACGG
ACAGTGATTCTGCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTT
TGAGGGCACAGTGTGTTGCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTT
TGCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACACC
TGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTTCACTGATTCTATAAGCCCAGCAT
TACCTGATACCAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCATATCC
CTCATTAACACAGACACAAAAATTCTAAATAAAATTTTAACAAATTAACCTAAACAATATATTTA
AAGATGATATATACTACTCAGTGTGGTTTGTCCACAAATGCAGAGTTGGTTTAATATTTAAAT
ATCAACCAGTGTAATTACGACATTAATAAAGTAAAAAGAAAACCATAAAAA

FIGURE 104

MAFVLILVLSFYELVSGQWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAVVH
LYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEATWELR
VAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLYDVEISI
IVQENAGSILCSIHLAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVMGMIIVFFK
SKGKIQAELOWRRKHGQAEIRDARKHAVEVTLPETAHPKLCVSDLKTVTHRKAQEVPHSEKRF
TRKSVVASQGFQAGRHYEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYWVLRLTTEHLYFT
FNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLLRPYIQHAMYDEEKG
TPIFICPVSWG

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 131-150, 235-259

FIGURE 105

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAGGAAAAGAG
 TTTGTTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTGGCAGTGTGCATTGGA
 CTCCTGTTTCATTATGTGAGATATAATCAAAAGAAGACCTACAATTACTATAGCACATTGTCATTTACAAC
 TGACAAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAACAATTTTACAGAAATGAGCCAGAGACTTGAAT
 CAATGGTGAAAAATGCATTTTATAAATCTCCATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTC
 AGTCAACAGAAGCATGGAGTGTGGCTCATATGCTGTTGATTGTAGATTTCCTCTACTGAGGATCCTGA
 AACTGTAGATAAAATTGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAG
 ATCCTCACTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTCGGGA
 ACACGAAGAAGTAAAACCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGAGGGTGAATG
 GCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTTAATTAATGCCACATGGC
 TTGTGAGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAGATGGACTGCTTCCTTTGGAGTAACA
 ATAAAACCTTCGAAAATGAAACGGGTCTCCGGAGAATAATTGTCCATGAAAAATACAAACCCCATCACA
 TGACTATGATATTTCTCTTGCAGAGCTTTCTAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCT
 TCCCTGATGCATCCTATGAGTTTCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAAT
 GATGGTTACAGTCAAAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACC
 TCAAGCTTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGCAT
 GCCAGGGTGACTCTGGAGGACCACTGGTTAGTTTCTAGATGCTAGAGATATCTGGTACCTTGCTGGAATAGTG
 AGCTGGGGAGATGAATGTGCGAAACCCAAACAAGCCTGGTGTTTATACTAGAGTTACGGCCTTGCGGGACTG
 GATTACTTCAAAACTGGTATCTAAGAGACAAAAGCCTCATGGAACAGATAACATTTTTTTTTGTTTTTTG
 GGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGAAGACTTGCAAAACAGCTAGATTTGACTGATCTCA
 ATAAACTGTTTGCTTGATGCATGTATTTCTTCCCAGCTCTGTTCCGCACGTAAGCATCTGCTTCTGCCA
 GATCAACTCTGTCTGTGAGCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTAC
 ATTACAGCCTGTATTCATTTGTTCTCTAGAAGTTTGTGAGAATTTGACTTGTGACATAAATTTGTAAT
 GCATATATACAATTTGAAGCACTCCTTTTCTTCAGTTCCCTCAGCTCCTCTCATTTCAGCAAATATCCATTT
 TCAAGGTGCAGAACAAAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCTACATTTTATTGGCACAGAA
 AAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTATTATGAAAGGTCAAGCAAAGACA
 GCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACCTAAGTTAAGGAAGTCCAGAAAGAAGCCAAG
 ATATATCCTTATTTTCATTTCCAAACAACTACTATGATAAATGTGAAGAAGATTCTGTTTTTTTGTGACCT
 ATAATAATTATACAACTTCATGCAATGTACTTGTCTAAGCAAATTAAAGCAAATATTTATTTAACATTG
 TTAAGGATGTCAACATATAACAATAAAATATAAATCACCCA

FIGURE 106

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDKLY
AEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFHSTED
PETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQSLRIVGG
TEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIKPSKMKRGL
RRIIVHEKYKHPSHDYDISLAELSSPVPTYNAVHRVCLPDASYEFQPGDVMFVTGFGALKNDGYS
QNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSSGGLVSSDARDIWYLAG
IVSWGDECAKPNKPGVYTRVTALRDWITSKTI

Transmembrane domain:

amino acids 21-40 (type II)

FIGURE 107

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCCTG
 CCCCAGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACCAGGGCCCAGC
 GCCGACGATCGCTGCCGTTTTTGGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTCTCCCTT
 ACGGGGCTCACAATGCGCCAGAGAAGATTCCGTGAAGTGCTGCGCTGCCGTGCTCTACGCCCTCAA
 TCTGCTCTTTTGGTTAATGTCCATCAGTGCTGTTGGCAGTTTCTGCTTGGATGAGGGACTACCTAA
 ATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTGACTTACTTTTCCT
 GTGGTTCATCCGGTCATGATTGCTGTTTGTCTGTTTCCTTATCATTGTGGGGATGTTAGGATATTG
 TGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTGGAAGTTTGCTTGTCAATTTTCT
 GTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTATGGTTCCAGTACAATGGTCA
 GATATGGTCACCTTTGAAAGCCAGGATGACAAATTATGGATTACCTAGATATCGGTGGCTTACTCA
 TGCTTGGAAATTTTTTTCAGAGAGAGTTTAAAGTGCTGTGGAGTAGTATATTTCAGTGACTGGTTGG
 AAATGACAGAGATGGACTGGCCCCAGATTCTTGCTGTGTTAGAGAATTTCCAGGATGTTCCAAA
 CAGGCCCACCAGGAAGATCTCAGTGACCTTTATCAAGAGGGTTGTGGGAAGAAAATGTATTCCTT
 TTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTTCTGGGAATCTCCATTGGGGTGACACAAA
 TCCTGGCCATGATTCTCACCATTACTCTGCTCTGGGCTCTGTATTATGATAGAAGGGAGCCTGGG
 ACAGACCAATGATGTCTTGAAGAATGACAACCTCTCAGCACCTGTCATGTCCCTCAGTAGAACCT
 GTTGAAACCAAGCCTGTCAAGAATCTTTGAACACACATCCATGGCAAACAGCTTTAATACACACT
 TTGAGATGGAGGAGTTATAAAAAAGAAATGTCACAGAAGAAAACCACAAACTTGTTTTATTGGACT
 TGTGAATTTTTGAGTACATACTATGTGTTTCAGAAATATGTAGAAATAAAAAATGTTGCCATAAAA
 TAACACCTAAGCATATACTATTCTATGCTTTAAAAATGAGGATGGAAAAGTTTCATGTCATAAGTC
 ACCACCTGGACAATAATTGATGCCCTTAAAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCC
 TGTGTATGACTTTTACTGAACACAGTTATGTTTTGAGGCAGCATGGTTTGATTAGCATTTCCGCA
 TCCATGCAAACGAGTCACATATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCAA
 CTAGTATATAAAGTACTAATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTA
 CTCAGCGATCTATTCTTCTGATGCTAAATAAATTATATATCAGAAAACCTTTCAATATTGGTGACT
 ACCTAAATGTGATTTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACAT
 TGTCTTAAGCTGATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGAT
 TTCAGTTCTGATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCTGTATAGCATCATT
 ATTTTTAGCCTTTCCCTGTTAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAAC
 TGTATTTTAAATACTTAACCACTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTC
 AGAATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTGAGA
 AAGGACTTGTATGCTGTTTTTCTCCCAATGAAGACTCTTTTGACACTAAACACTTTTTAAAAA
 GCTTATCTTTGCCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATCTACAGAAAA
 TAGTGTCTTTTTTCTCCAGAAAAATGCTTGTGAGAATCATTAAAACATGTGACAATTTAGAGATT
 CTTTGTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTTTTTTACAA
 GAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTATTTTGTGTTAT
 TTCTCAGAATATGGAAAGAAAATTAAATGTGTCAATAAATATTTTCTAGAGAGTAA

FIGURE 108

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTLTAE TRVEEAVILTYFPVVHP
VMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWSDMVT
LKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTMDWPPDSCCVREFPGCSKQAHQ
EDLSDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDRREPGTDQM
MSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

Signal peptide:

amino acids 1-33

Transmembrane domains:

amino acids 12-35, 57-86, 94-114, 226-248

FIGURE 109

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTTCTCCTCTGATAAAGCCCCTACCAGTGCT
 GATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAGTGCTTGAAAGAGAAGGGGACAAAGGAACA
 CCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTCCAGAAGGATGCCTCCATTCTGCTTCTCACCTG
 CCTCTTCATCACAGGCACCTCCGTGTCACCCGTGGCCCTAGATCCTTGTCTGCTTACATCAGCCTGAATGAGC
 CCGGAGGAACACTGACCACCAGTTGGATGAGTCTCAAGGTCTCTCTATGTGACAACCATGTGAATGGGGAG
 TGGTACCACTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAACCCA
 CGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCAGGCTTGTGCCAGCT
 TCAATGGGAACGTGCTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTGCCCTGGAGGCTACTATGTGTATCGT
 CTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGGTCATTTTTATGACATCTGCGACGAGGACTGCCATGG
 CAGCTGCTCAGATACCAGCGAGTGACATGCGCTCCAGGAAGTGTGCTAGGCCCTGACAGGCAGACATGCTTTG
 ATGAAAATGAATGTGAGCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTCTACCGCTGT
 GAGTGTGGGGTTGGCCGTGTGCTAAGAACTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAATAACAA
 TGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCCCCGGGGCCTGGTGTGT
 CTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAAATCAAATGCCATTGAAGTGAACATCCCCAGGGAG
 CTGGTTGGTGGCCTGGAGCTCTTCTGACCAACACCTCTGCCGAGGAGTGTCCAACGGCACCCATGTCAACAT
 CCTCTTCTCTCTCAAGACATGTGGTACAGTGGTGCATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGA
 CAGGTCTACCCAAGCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTGCTGATCCCGGTG
 ACCTGCGAGTTTCCACGCCCTGTACACCATTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCCACTGGAAT
 CATGAGCCGAAATCATGGGATCTTCCCATTCACTCTGGAGATCTTCAAGGACAATGAGTTTGAAGAGCCTTACC
 GGGAGCTCTGCCCACCCCTCAAGCTTCGTGACTCCCTCTACTTTGGCATTGAGCCCGTGGTGCACGTGAGCGGC
 TTGGAAGCTTGGTGGAGAGCTGCTTTGCCACCCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCAT
 CCGGGATGGCTGTGTTTCAGATGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGG
 TCCCTGTCTTCAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGTG
 TTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGCGAATGCGTCTGGGGCAGGAGGAGGACTCAGC
 CGGTCTACAGGGCCAGACGCTAACAGGCGGCCGATCCGCATCGACTGGGAGGACTAGTTCGTAGCCATACCTC
 GAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTCCCCCACC GCCCTCTAAGAACATCTGCCAACAGC
 TGGGTTACAGACTTCACACTGTGAGTTCAGACTCCAGCACCACTCACTCTGATTCTGGTCCATTCACTGGGCA
 CAGGTACAGCACTGCTGAACAATGTGGCCTGGGTGGGTTTCATCTTTCTAGGGTTGAAAACATAAAGTGTCCA
 CCCAGAAAGACACTCACCCCATTTCCCTCATTTCTTCTTACACTTAAATACCTCGTGTATGGTGAATCAGAC
 CACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAGTTACTGAAATTATGA
 CTTAAATACCAATGACTCCTTAAATATGTAAATTATAGTTATACCTTGAAATTTCAATTCAAATGCAGACTAA
 TTATAGGGAATTTGGAAGTGTATCAATAAACAGTATATAATTTT

FIGURE 110

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFTGMAGDAMP
TFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPGGYVYRLTKPSVCFHV
YCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQNNGGCSEICVNLKNSYRCECGVGRV
LRSDGKTCEDVEGCHNNNGGCSHSCLGSEKGYQCECPRGLVLSEDNHTCQVPVLCKSNAIEVNIPRELVGG
LELFLTNTSCRGVSNNGTHVNILFSLKTCGTVDVVDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVT
CEFPRLYTISEGYVPNLRNSPLEIMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKLRLDSLYFGIEPVVHV
SGLESIVESCFATPTS KIDEVLKYILIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRV
LVCGVLDERSRCAQGCCHRRMRGAGGEDSAGLQGQTLTGGP IRI DWED

Important features of the protein:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

Tyrosine kinase phosphorylation sites.

amino acids 411-418, 443-451

N-myristoylation sites.

amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,
522-528, 531-537

Aspartic acid and asparagine hydroxylation site.

amino acids 197-209

ZP domain proteins.

amino acids 431-457

Calcium-binding EGF-like proteins.

amino acids 191-212, 232-253

FIGURE 111

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCTGCACTCGG
 GCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACCTGTGTGGGGAGGCCCT
 CCTGCTGCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACCGGGAGGATCACAGAGCCAGCATGT
 TACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCGATGTCAAACCCCTGCGCAAACCCCGTATCCCC
 ATGGAGACCTTCAGAAAGGTGGGGATCCCCATCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTGT
 GGTGTCTCATCAAGGTGATTCTGGATAAATACTACTTCTCTGCGGGCAGCCTCTCCACTTCATCCCGA
 GGAAGCAGCTGTGTGACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTC
 CCCGAAGGGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCAC
 AGGGAAGTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTAGGCAGATGG
 GCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAATCACAGAAAACAGCCAG
 GAGCTTCGCATGCGGAACCAAGTGGGCCCTGTCTCTCAGGCTCCCTGGTCTCCCTGCACTGTCTTGCCTG
 TGGGAAGAGCCTGAAGACCCCCGTGTGGTGGGTGGGGAGGAGGCCTCTGTGGATTCTTGGCCTTGGCAGG
 TCAGCATCCAGTACGACAAACAGCACGTCTGTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCA
 GCCCCTGCTTCAGGAAACATACCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAG
 CTTCCTCATCCCTGGCTGTGGCCAAGATCATCATCATTGAATTCAACCCCATGTACCCCAAAGACAATGACA
 TCGCCCTCATGAAGCTGCAGTTCCTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCTTT
 GATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGAATGGAGGGAA
 GATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGTGCAATGCAGACGATGCGT
 ACCAGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGGAAGGGGTGTGGACACCTGCCAGGGT
 GACAGTGGTGGGCCCCTGATGTACCAATCTGACCAGTGGCATGTGGTGGGCATCGTTAGCTGGGGCTATGG
 CTGCGGGGGCCCCGAGCACCCCAAGGATATACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCT
 GGAAGGCTGAGCTGTAATGCTGCTGCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCT
 GGGGATCCCCCAAAGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCCACAGCCTCAGCAT
 TTCTTGGAGCAGCAAAGGGCCTCAATTCCTGTAAGAGACCCTCGCAGCCCAGAGGCGCCAGAGGAAGTCA
 GCAGCCCTAGCTCGGCCCACTTGGTGCTCCCAGCATCCCAGGGAGAGACACAGCCCACTGAACAAGGTCT
 CAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGCAGGCTGTCTTGTAAAAGCC
 CAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCAGCCCTGTCCGTCTTACCCATCCCCAA
 GCCTACTAGAGCAAGAAACCAGTTGTAATATAAAATGCACTGCCCTACTGTTGGTATGACTACCGTTACCT
 ACTGTTGTCATTGTTATTACAGCTATGGCCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAA
 AAAA

FIGURE 112

MLQDPDSQPLNSLDVKPLRKPRIPMETFRKVGIPIIIIALLSLASIIIVVVLIKVILDKYYFLCG
QPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSACFDN
FTEALAEACRQMGYSRAVEIGPDQDLVDVEITENSQELMRNSSGPCLSGSLVSLHCLACGKSL
KTPRVVGEEASVDSWPWQVSIQYDKQHVCSSILDPHWLTAHCFRKHTDVFNWKVRAGSDKL
GSFPSLAVAKIIIIIEFNPMYPKDNNDIALMKLQFPLTFSGTVRPICLPFFDEELTPATPLWIIGWG
FTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTCCGDSGGPLMYQS
DQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL

Transmembrane domain:

amino acids 32-53 (typeII)

FIGURE 113

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGATTA
TAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACTTTTTTCAGCAACTAAAAAGCCACAGGAGT
TGAAGTGTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATTAAAAATC
TGTTTTTTGTTTCTCTTGTAAGTACCTTTACCTTCCTAACACAGAGGATCTGTCACTGTGGCTCT
GGCCCAAACCTGACCTTCACTCTGGAACGAGAACAGAGGTTTCTACCCACACCGTCCCCTCGAAG
CCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAACTGTCTCACGTCT
GGAGGCACTGACTCGGGCAGTGCAGGTAGCTGAGCCTCTTGGTAGCTGCGGCTTTCAGGTGGGC
CTTGCCCTGGCCGTAGAAGGGATTGACAAGCCCGAAGATTTTCATAGGCGATGGCTCCCCTGCCCC
AGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGGGCGGTGGACACCTGCTCA
GAAGCAGTGGGTGAGACATCACGCTGCCCCGCCATCTAACCTTTTCATGTCTGCACATCACCTG
ATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCTTGAGTGAGCTGTGGCTCAGACC
CAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGACTTGCAATTCTCCTGGAACATGAGG
GAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACCTGTGCCAAATTATGGGTGAGAAAAGATG
GAGGTGTTGGGTTATCACAAGGCATCGAGTCTCCTGCATTCACTGGACATGTGGGGGAAGGGCTG
CCGATGGCGCATGACACACTCGGGACTCACCTCTGGGGCCATCAGACAGCCGTTTCCGCCCCGAT
CCACGTACCAGCTGCTGAAGGGCAACTGCAGGCCGATGCTCTCATCAGCCAGGCAGCAGCCAAAA
TCTGCGATCACCAGCCAGGGGCAGCCGTCTGGGAAGGAGCAAGCAAAGTGACCATTCTCCTCCC
CTCCTTCCCTCTGAGAGGCCCTCCTATGTCCCTACTAAAGCCACCAGCAAGACATAGCTGACAGG
GGCTAATGGCTCAGTGTTGGCCCAGGAGGTGAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCT
GTGCGAACACGGAAATGCCTCCAGTAAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGG
CTCAATTTAAATCATGTTCTAGTAATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGTT
CAAATGATCTCCAAGGGCCCTTATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCA
AACCTAAGAACCAGGTGCATTAGAATCAGTTATTGCGGGTGTGGTGGCCTGTAATGCCAACAT
TTTGGGAGGCCGAGGCGGGTAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAACATGG
TGAAACCCCTGTCTCTACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATC
CCAGCTACTCGGGAGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACA
GGAGAATCACTTACGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTA
TGTTTATTTGTAA

FIGURE 114

MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTASPC
WPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD

Signal peptide:

amino acids 1-15

FIGURE 115

CAGCAGTGGTCTCTCAGTCCTCTCAAAGCAAGGAAAGAGTACTGTGTGCTGAGAGACCAATGGCAA
AGAATCCTCCAGAGAATTGTGAAGACTGTCACATTCTAAATGCAGAAGCTTTTAAATCCAAGAAA
ATATGTAAATCACTTAAGATTTGTGGACTGGTGTGGTATCCTGGCCCTAACTCTAATTGTCCT
GTTTTGGGGGAGCAAGCACTTCTGGCCGGAGGTACCCAAAAAGCCTATGACATGGAGCACACTT
TCTACAGCAATGGAGAGAAGAAGAAGATTTACATGGAAATTGATCCTGTGACCAGAACTGAAATA
TTCAGAAGCGGAAATGGCACTGATGAAACATTGGAAGTGCACGACTTTAAAAACGGATACACTGG
CATCTACTTCGTGGGTCTTCAAAAATGTTTTATCAAACTCAGATTAAAGTGATTCCCTGAATTTT
CTGAACCAGAAGAGGAAATAGATGAGAATGAAGAAATTACCACAACCTTTCTTTGAACAGTCAGTG
ATTTGGGTCCCAGCAGAAAAGCCTATTGAAAACCGAGATTTTCTTAAAAATTCCAAAATTCTGGA
GATTTGTGATAACGTGACCATGTATTGGATCAATCCCACTCTAATATCAGTTTCTGAGTTACAAG
ACTTTGAGGAGGAGGGAGAAGATCTTCACTTTCCTGCCAACGAAAAAAGGGATTGAACAAAAT
GAACAGTGGGTGGTCCCTCAAGTGAAAGTAGAGAAGACCCGTCACGCCAGACAAGCAAGTGAGGA
AGAACTTCCAATAAATGACTATACTGAAAATGGAATAGAATTTGATCCCATGCTGGATGAGAGAG
GTTATTGTTGTATTTACTGCCGTCGAGGCAACCGCTATTGCCGCCGCGTCTGTGAACCTTTACTA
GGCTACTACCCATATCCATACTGCTACCAAGGAGGACGAGTCATCTGTCGTGTCATCATGCCTTG
TAACTGGTGGGTGGCCCGCATGCTGGGGAGGGTCTTAATAGGAGGTTTGAGCTCAAATGCTTAAAC
TGCTGGCAACATATAATAAATGCATGCTATTCAATGAATTTCTGCCTATGAGGCATCTGGCCCT
GGTAGCCAGCTCTCCAGAATTACTTGTAGGTAATTCCTCTCTTCATGTTCTAATAAACTTCTACA
TTATCACCAAAAAAAAAAAAAAAAAAAAA

FIGURE 116

MAKNPPENCEDCHILNAAEFKSKKICKSLKICGLVFGILALT LIVLFWGSKHFWPEVPKKAYDME
HTFYSNGEKKKIYMEIDPVTRTEIFRSGNGTDETLVHDFKNGYTGIIYFVGLQKCFIKTQIKVIP
EFSEPEEEIDENEEITTTFFEQSVIWWPAEKPIENRDFLKNKILEICDNVTMYWINPTLISVSE
LQDFEEEGEDLHFPANEKKGIEQNEQWVVPQVKVEKTRHARQASEEELPINDYTENGIEFDPMLD
ERGYCCIIYCRRGNRYCRRVCEPLLGYYPYPYCYQGGRVICRVIMPCNWWWVARMLGRV

Important features of the protein:

Signal peptide:

amino acids 1-40

Transmembrane domain:

amino acids 25-47 (type II)

N-glycosylation sites.

amino acids 94-97, 180-183

Glycosaminoglycan attachment sites.

amino acids 92-95, 70-73, 85-88, 133-136, 148-151, 192-195, 239-
242

N-myristoylation sites.

amino acids 33-38, 95-100, 116-121, 215-220, 272-277

Microbodies C-terminal targeting signal.

amino acids 315-317

Cytochrome c family heme-binding site signature.

amino acids 9-14

FIGURE 117

GAGCTCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGCAGGCGGCA
GGGCGGGCGGCCAGGATCCATGTCCACCACCACATGCCAAGTGGTGGCGTTCCCTCCTGTCCATCCTGGGGCT
GGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGGACCTGTACGACAACCCCGTCACCT
CCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGAGGCAGAGTTCAGGCTTCACCGAATGCAGGCCC
TATTTACCATCCTGGGACTTCCAGCCATGCTGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCT
GGGTGCCATTGGCCTCCTGGTATCCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTG
CCAAAGCCAACATGACACTGACCTCCGGGATCATGTTTCATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTG
TCTGTGTTTGCCAAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTGG
GATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGCGGCTCTGTTGTTGGGCTGGGTGCTGGAGGCC
TCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAGAAGAAACCAACTACAAA
GCCGTTTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTGGAGGCTTCAAGGCCAGCACTGGCCTT
TGGGTCCAACACCAAAAACAAGAAGATATACGATGGAGGTGCCCGCACAGAGGACGAGGTACAATCTTATC
CTTCCAAGCACGACTATGTGTAATGTCTCTAAGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCA
CCCCAAAAACAAGGAGATCCCATCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCT
CGATTTTCATCTTTGGAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACA
GCTGAGTTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAACT
TTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAGACTCCCCCTC
TTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAGAAAACCTTTTGAAAGGAAA
GAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTCCCCACCCCAACTTTGGCTAGTAATAA
ACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATTTGTAATCTCTCCAGCCCATGATCTCGGTTTTCTT
ACACTGTGATCTTAAAAGTTACCAAACCAAAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTG
TTGACATCTTCTTATTACAGCAACACCATTTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTTCTGT
CGCGGGTCAGAAATTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAA
AATAAATAATGTTTTAGTAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAGAAG
GAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAAGTACAAATTCC
ATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGGATCACTTGAGCCCAGAAGT
TCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAAATACAGAGAGAAAAAATCAGCCAGTCA
TGGTGGCATAACCTGTAGTCCCAGCATTCGGGGAGGCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGT
TGGGGCTGCAGTGAGCCATGATCACACCACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAA
AATAAAAAATAAATAATGGAACACAGCAAGTCCTAGGAAGTAGGTTAAACTAATTCTTTAA

FIGURE 118

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLRSCVRQSSGFTECRP
YFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFIVSGL
CAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGGVMMCIA
CRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQSYPSKHDY
V

Signal peptide:

amino acids 1-23

Transmembrane domains:

amino acids 81-100, 121-141, 173-194

FIGURE 119

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGGAGTCC
 AGCTGGCTAAAACTCATCCCAGAGGATAATGGCAACCCATGCCTTAGAAATCGCTGGGCTGTTTCTTG
 GTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTGCCTCAGTGGAGAGTGTGGCCTTCATT
 GAAAACAACATCGTGGTTTTTTGAAAACCTTCTGGGAAGGACTGTGGATGAATTGCGTGAGGCAGGCTAA
 CATCAGGATGCAGTGCAAAATCTATGATTCCCTGCTGGCTCTTTCTCCGGACCTACAGGCAGCCAGAG
 GACTGATGTGTGCTGCTTCCGTGATGTCTTCTTGGCTTTCATGATGGCCATCCTTGGCATGAAATGC
 ACCAGGTGCACGGGGGACAATGAGAAGGTGAAGGCTCACATTCTGCTGACGGCTGGAATCATCTTCAT
 CATCACGGGCATGGTGGTGCTCATCCCTGTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATA
 ACTCAATAGTGAATGTTGCCCCAAAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCA
 CTGGTGCTGATTGTTGGAGGAGCTCTGTTCTGCTGCGTTTTTTTGTGCAACGAAAAGAGCAGTAGCTA
 CAGATACTCGATACCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAGTCAACCGAGCG
 TCTACTCCAGAAGTCAGTATGTGTAGTGTGTATGTTTTTTTAACTTTACTATAAAGCCATGCAAATG
 ACAAAAATCTATATTACTTTCTCAAAATGGACCCCAAGAACTTTGATTTACTGTTCTTAACTGCCT
 AATCTTAATTACAGGAACTGTGCATCAGCTATTTATGATTCTATAAGCTATTTACGAGAATGAGATA
 TTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGGTTCAAGCATCTA
 CTCTTTTTATCATTACTTCAAAATGACATTGCTAAAGACTGCATTATTTTACTACTGTAATTTCTCC
 ACGACATAGCATTATGTACATAGATGAGTGTAACATTTATATCTCACATAGAGACATGCTTATATGGT
 TTTATTTAAATGAAATGCCAGTCCATTACACTGAATAAATAGAACTCAACTATTGCTTTTCAGGGAA
 ATCATGGATAGGGTTGAAGAAGGTACTATTAATTGTTTAAAAACAGCTTAGGGATTAATGTCCTCCA
 TTTATAATGAAGATTAAATGAAGGCTTTAATCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATAT
 GCTGTTTTTTAGCCTAGGAGTTAGAAATCCTAACTTCTTTATCCTCTTCTCCAGAGGCTTTTTTTTT
 CTTGTGTATTAAATTAACATTTTTTAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAACCTGCTT
 TTCCAGGGCTATACTCAGAAGAAAGATAAAAGTGTGATCTAAGAAAAAGTGATGGTTTTAGGAAAGTG
 AAAATATTTTTGTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGAT
 ATATTTTAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAGAGCAGAAAAATA
 TGTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTCACCT
 GCTCCTATGTGGGTACCTGAGTCAAATTTGTCATTTTTGTTCTGTGAAAAATAAATTTCTTCTTGTA
 CCATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATTCCAAATTTGATGAA
 ACTGACAATCCAATTTGAAAGTTTGTGTGACGCTGTCTAGCTTAAATGAATGTGTTCTATTTGCTT
 TATACATTTATATTAATAAATTGTACATTTTTCTAATT

FIGURE 120

MATHALEIAGLFLGGVGMVGTVAVTVMPPQWRVSAFIENNIVVFENFW EGLWMNCVRQANIRMQCK
IYDSSLALSPDLQAARGLMCAASVMSFLAFMAILGMKCTRCTGDNEKVKAHILLTAGIIFIITG
MVLVLPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCNEKSSSY
RYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

Signal peptide:

amino acids 1-17

Transmembrane domains:

amino acids 82-101, 118-145, 164-188

FIGURE 121

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGGAG
CCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGCAGCC
GGGAGCCATGCGACCCCAGGGCCCCGCCGCTCCCCGCAGCGGCTCCGCGGCCTCCTGCTGCTCC
TGCTGCTGCAGCTGCCCCGCGCGTTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAAGGCGCAG
CTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAGCAGGAGTGCC
TGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCGGGTACACCTGGGATCCAGGTGGGATG
GATTCAAAGGAGAAAAGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCTGGACACCCAACTAC
AAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAATTGCGGAGTGACATT
TACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTTCAGTGGCTCACTTCGGCTAAAATGCA
GAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTGAATGTTTCAGGACCTCTTCCC
ATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGAATTCAACAATTAATATTCATCG
CACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTGGATTAGTGGATGTTGCTATCTGGG
TTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTACTGGATGGAATTCAGTTTCTCGCATC
ATTATTGAAGAACTACCAAAATTAAATGCTTTAATTTTCATTTGCTACCTCTTTTTTTATTATGCC
TTGGAATGGTTCACTTAAATGACATTTTAAATAAGTTTATGTATACATCTGAATGAAAAGCAAAG
CTAAATATGTTTACAGACCAAAGTGTGATTTACACTGTTTTTAAATCTAGCATTATTCATTTTG
CTTCAATCAAAAGTGGTTTCAATATTTTTTTTAGTTGGTTAGAATACTTTCTTCATAGTCACATT
CTCTCAACCTATAATTTGGAATATTGTTGTGGTCTTTTGTTTTTCTCTTAGTATAGCATTTTTTA
AAAAATATAAAAGCTACCAATCTTTGTACAATTGTAAATGTTAAGAATTTTTTTTATATCTGT
TAAATAAAAATTATTTCCAACA

FIGURE 122

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVPGR
DGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAECTFTK
MRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECSGPLPIEAIYLDQGSPEMNSTINIHRTS
SVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIEELPK

Signal peptide:

amino acids 1-30

Transmembrane domain:

amino acids 195-217

FIGURE 123

GCTGAGCGTGTGCGCGGTACGGGGCTCTCCTGCCTTCTGGGCTCCAACGCAGCTCTGTGGCTGAA
 CTGGGTGCTCATCACGGGAAGTCTGGGCTATGGAATACAGATGTGGCAGCTCAGGTAGCCCCAA
 ATTGCCTGGAAGAATACATCATGTTTTTCGATAAGAAGAAATTGTAGGATCCAGTTTTTTTTTTTA
 ACCGCCCCCTCCCCACCCCCCAAAAACTGTAAAGATGCAAAAACGTAATATCCATGAAGATCC
 TATTACCTAGGAAGATTTTGATGTTTTGCTGCGAATGCGGTGTTGGGATTTATTTGTTCTTGGAG
 TGTTCTGCGTGGCTGGCAAAGAATAATGTTCCAAAATCGGTCCATCTCCCAAGGGGTCCAATTTT
 TCTTCCTGGGTGTGACGAGCCCTGACTCACTACAGTGCAGCTGACAGGGGTGTGATGCAACTG
 GCCCCAAGCCAAAGCAAAAGACCTAAGGACGACCTTTGAACAATACAAAGGATGGGTTTCAATG
 TAATTAGGCTACTGAGCGGATCAGCTGTAGCACTGGTTATAGCCCCACTGTCTTACTGACAATG
 CTTTCTTCTGCCGAACGAGGATGCCCTAAGGGCTGTAGGTGTGAAGGCAAAATGGTATATTGTGA
 ATCTCAGAAATTACAGGAGATACCCTCAAGTATATCTGCTGGTTGCTTAGGTTTGTCCCTTCGCT
 ATAACAGCCTTCAAAAACCTTAAGTATAATCAATTTAAAGGGCTCAACCAGCTCACCTGGCTATAC
 CTTGACCATAACCATATCAGCAATATTGACGAAAATGCTTTTAAATGGAATACGCAGACTCAAAGA
 GCTGATTCTTAGTTCCAATAGAATCTCCTATTTTCTTAACAATACCTTCAGACCTGTGACAAATT
 TACGGAAGCTTGGATCTGTCTATAATCAGCTGCATTCTCTGGGATCTGAACAGTTTCGGGGCTTG
 CGGAAGCTGCTGAGTTTACATTTACGGTCTAACTCCCTGAGAACCATCCCTGTGCGAATATTCCA
 AGACTGCCGCAACCTGGAACCTTTGGACCTGGGATATAACCGGATCCGAAGTTTAGCCAGGAATG
 TCTTTGCTGGCATGATCAGACTCAAAGAACTTCACCTGGAGCACAATCAATTTTCCAAGCTCAAC
 CTGGCCCTTTTTCGAAGTTGGTCAGCCTTCAGAACCTTTACTTGCAGTGGGAATAAAATCAGTGT
 CATAGGACAGACCATGTCTGGACCTGGAGCTCCTTACAAAGGCTTGATTTATCAGGCAATGAGA
 TCGAAGCTTTCAGTGGACCCAGTGTTCAGTGTGTCCGAATCTGCAGCGCCTCAACCTGGAT
 TCCAACAAGCTCACATTTATTGGTCAAGAGATTTTGATTCTTGATATCCCTCAATGACATCAG
 TCTTGCTGGGAATATATGGGAATGCAGCAGAAATATTGCTCCCTTGTAACCTGGCTGAAAAGTT
 TTAAAGGTCTAAGGGAGAATACAATTATCTGTGCCAGTCCCAAAGAGCTGCAAGGAGTAAATGTG
 ATCGATGCAGTGAAGAACTACAGCATCTGTGGCAAAAGTACTACAGAGAGGTTTGATCTGGCCAG
 GGCTCTCCCAAAGCCGACGTTTAAGCCCCAAGCTCCCCAGGCCGAAGCATGAGAGCAAACCCCTT
 TGCCCCCGACGGTGGGAGCCACAGAGCCCCGGCCAGAGACCGATGCTGACGCCGAGCACATCTCT
 TTCCATAAAATCATCGCGGGCAGCGTGGCGCTTTTCTGTCCGTGCTCGTCATCTGCTGGTTAT
 CTACGTGTCAATGGAAGCGGTACCCTGCGAGCATGAAGCAGCTGCAGCAGCGCTCCCTCATGCGAA
 GGCACAGGAAAAAGAAAAGACAGTCCCTAAAGCAAAATGACTCCCAGCACCCAGGAATTTTATGTA
 GATTATAAACCACCAACACGGAGACCAGCGAGATGCTGCTGAATGGGACGGGACCCTGCACCTA
 TAACAAATCGGGCTCCAGGGAGTGTGAGGTATGAACCATTGTGATAAAAAGAGCTCTTAAAAGCT
 GGGAAATAAGTGGTGCTTTATTGAACTCTGGTGACTATCAAGGGAACGCGATGCCCCCCTCCCC
 TTCCCTCTCCCTCTCACTTTGGTGGCAAGATCCTTCCTTGTCGGTTTTAGTGCATTATAATACT
 GGTCATTTTCTCTCATACATAATCAACCCATTGAAATTTAAATACCACAATCAATGTGAAGCTT
 GAACTCCGGTTTAAATATAATACCTATTGTATAAGACCCTTTACTGATTCCATTAATGTGCGATTT
 GTTTTAAGATAAACTTCTTTCATAGGTAAAAA

FIGURE 124

MGFNVIRLLSGSAVALVIAPTVLLTMLSSAERGCPCGCRCEGKMVYCESQKLQEIPSSISAGCLG
LSLRYNLSLQKLKYNQFKGLNQLTWLYLDHNHISNIDENAFNGIRRLKELILSSNRISYFLNNTFR
PVTNLRNLDLSYNQLHSLGSEQFRGLRKLLSLHLRSNSLRTIPVRI FQDCRNLELLDLGYNRIRS
LARNVFAGMIRLKEHLHLEHNQFSKLNLFPRVLSQLNLYLQWNKISVIGQTMSWTWSSLQRLDL
SGNEIEAFSGPSVFCVQPNLQRLNLDNKLTFIGQEILDSWISLNDISLAGNIWECSRNICSLVN
WLKSFKGLRENTIICASPKELQGVNVIDAVKNYSICGKSTTERFDLARALPKPTFKPKLPRPKHE
SKPPLPPTVGATEPGPETDADAEHISFHKIIAGSVALFLSVLVILLVIYVSWKRYPASMKQLQQR
SLMRHRHKKKRQSLKQMTPTSTQEFYVDYKPTNTETSEMLLNGTGPCCTYNKSGSRECEV

Important features of the protein:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 420-442

N-glycosylation sites.

amino acids 126-129, 357-360, 496-499, 504-507

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 465-468

Tyrosine kinase phosphorylation site.

amino acids 136-142

N-myristoylation sites.

amino acids 11-16, 33-38, 245-250, 332-337, 497-502, 507-512

FIGURE 125

CCGTTATCGTCTTGCGCTACTGCTGAATGTCCGTCCCGGAGGAGGAGGAGAGGCTTTTGCCGCTG
ACCCAGAGATGGCCCCGAGCGAGCAAATTCCTACTGTCCGGCTGCGCGGCTACCGTGGCCGAGCT
AGCAACCTTTCCCCTGGATCTCACAAAACTCGACTCCAAATGCAAGGAGAAGCAGCTCTTGCTC
GGTTGGGAGACGGTGCAAGAGAATCTGCCCCCTATAGGGGAATGGTGCGCACAGCCCTAGGGATC
ATTGAAGAGGAAGGCTTTCTAAAGCTTTGGCAAGGAGTGACACCCGCCATTTACAGACACGTAGT
GTATTCTGGAGGTCGAATGGTCACATATGAACATCTCCGAGAGGTTGTGTTTGGCAAAAGTGAAG
ATGAGCATTATCCCCCTTTGGAAATCAGTCATTGGAGGGATGATGGCTGGTGTATTGGCCAGTTT
TTAGCCAATCCAACCTGACCTAGTGAAGGTTCAAGTGCAAATGGAAGGAAAAAGGAACTGGAAGG
AAAACCATTGCGATTTTCGTGGTGTACATCATGCATTTGCAAAAATCTTAGCTGAAGGAGGAATAC
GAGGGCTTTGGGCAGGCTGGGTACCCAATATACAAAGAGCAGCACTGGTGAATATGGGAGATTTA
ACCACTTATGATACAGTGAAACACTACTTGGTATTGAATACACCACTTGAGGACAATATCATGAC
TCACGGTTTATCAAGTTTATGTTCTGGACTGGTAGCTTCTATTCTGGGAACACCAGCCGATGTCA
TCAAAGCAGAATAATGAATCAACCACGAGATAAACAAGGAAGGGGACTTTTGTATAAATCATCG
ACTGACTGCTTGATTCAGGCTGTTCAAGGTGAAGGATTCATGAGTCTATATAAAGGCTTTTTACC
ATCTTGGCTGAGAATGACCCCTTGGTCAATGGTGTCTGGCTTACTTATGAAAAATCAGAGAGA
TGAGTGGAGTCAGTCCATTTTAA

FIGURE 126

MSVPEEEERLLPLTQRWPRASKFLLSGCAATVAELATFPLDLTKTRLQMQGEAALARLGDGARES
APYRGMVRTALGIIEEGFLKLWQGVTPAIYRHVVYSGGRMVITYEHLREVVFVGKSEDEHYPLWKS
VIGGMMAGVIGQFLANPTDLVKVQMMEGKRKLEGKPLRFRGVHHAFAKILAEGGIRGLWAGWVP
NIQRAALVNMGDLTTYDTVKHYLVLNTPLEDNIMTHGLSSLCSGLVASILGTPADVIKSRIMNQP
RDKQGRGLLYKSSTDCLIQAVQGEFMSLYKGFLPSWLRMTPWSMVFWLTYEKIREMSGVSPF

Transmembrane domains:

amino acids 25-38, 130-147, 233-248

FIGURE 127

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCGAC
CCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGCGTGGGCCCATGGCCAGGCCCCGGCATGG
AGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGGCCTCGGGGGGCATCGGCGCGGGCCGTGGCC
CGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAACATCGAGGA
GCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCTACAGATGTGACCTAT
CAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGCGGTGTAGACATC
TGCATCAACAATGCTGGCTTGGCCCCGGCCTGACACCCTGCTCTCAGGCAGCACCAGTGGTTGGAA
GGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAAGCCTACCAGTCCATGA
AGGAGCGGAATGTGGACGATGGGCACATCATTAACATCAATAGCATGTCTGGCCACCGAGTGTTA
CCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTCCTGCGCTGACAGAGGGACT
GAGGCAAGAGCTTCGGGAGGCCCCAGACCCACATCCGAGCCACGTGCATCTCTCCAGGTGTGGTGG
AGACACAATTCGCCTTCAAACCTCCACGACAAGGACCCTGAGAAGGCAGCTGCCACCTATGAGCAA
ATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTTATCTACGTCCCTCAGCACCCCCGCACA
CATCCAGATTGGAGACATCCAGATGAGGCCCACGGAGCAGGTGACCTAGTGACTGTGGGAGCTCC
TCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTGCCTCTGGATTTTAGGTGTTGATTTCTGGAT
CACGGGATACCACTTCCTGTCCACACCCCGACCAGGGGCTAGAAAATTTGTTTGAGATTTTATA
TCATCTTGTCAAATTGCTTCAGTTGTAAATGTGAAAATGGGCTGGGGAAAGGAGGTGGTGTCCC
TAATTGTTTTACTTGTTAACTTGTTCTTGTGCCCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTG
TCTTCCCTTTGACATGGGAAAGGAGTTGTGGCCAAAATCCCCATCTTCTTGCACCTCAACGTCTG
TGGCTCAGGGCTGGGGTGGCAGAGGGAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCC
AGACTTCCTCCTCTGCCTGCCCCACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCC
AGCCAGTCTTGGCTTCTTGTCCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAG
CAGAACACCAGGGCCTGGCCCAGTGGATTTTCATGGTGATCATTAATAAAGAAAAATCGCAACCAA
AAAAAAAAAA

FIGURE 128

MARPGMERWRDRLALVTGASGGIGA AAVARALVQQGLKVVG CARTVGNIEELAAECKSAGYPGTLL
PYRCDLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFNVNVLALSICTR
EAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQTHIRATC
ISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQMRPTEQVT

Important features of the protein:

Signal peptide:

amino acids 1-17

N-myristoylation sites.

amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,
199-205

Short-chain alcohol dehydrogenase.

amino acids 30-42, 104-114

FIGURE 129

AAC TTCTAC ATGGGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACCAT
CATGTCCCTCCCACCCTCCTTTGACTGCGGGCCGTT CAGGTGCAGAGTCTCAGTTGCCCGGGAGC
ACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCCAGTTCTGGTTTCATGC
CAGCCTGTAAAAGGCCATGGAAC TTTGGGTGAATCACCGATGCCATTTAAGAGGGTTTTCTGCCA
GGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCA TTTCAGTAGCCACCAGCCACCTGTGG
CCGTTGAGTGCTTGAAAT TGAGGAACTGAGAAAATTAATTTCTCATGTATTTTTCTCATTTATTTA
TTAATTTTTTAACTGATAGTTGTACATATTTGGGGGTACATGTGATATTTGGATACATGTATACAA
TATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACATTTATTTTTTATTCTTT
TTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCCATCTCAGCTTACTGCAAC
CTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAAGTAGCTGGGACTACAGGCAT
GCACCACAATGCCCAACTAATTTTTGTATTTTTAGTAGAGACGGGGTTTTGCCATGTTGCCCAGG
CTGGCCTTGAAC TCTGGCCTCAAACAATCCACTTGCCTCGGCCTCCCAAAGTGTTATGATTACA
GGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTTCTTTGTGTTGGGAACTTTGAAATTAT
ACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTGCTATGGAACACTGGGACTTCTTCCCTCT
ATCTAACTGTATATTTGTACCAGTTAACCAACCGTACTTCATCCCCACTCCTCTCTATCCTTCCC
AACCTCTGATCACCTCATTCTACTCTCTACCTCCATGAGATCCACTTTTTTAGCTCCCACATGTG
AGTAAGAAAATGCAATATTTGTCTTTCTGTGCCTGGCTTATTTCACTTAACATAATGACTTCCTG
TTCCATCCATGTTGCTGCAAAATGACAGGATTCGTTCTTAATTTCAATTAAAATAACCACACATG
GCAAAAA

FIGURE 130

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQPV
KGHGTLGESPMFVKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

Important features of the protein:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 86-92

Zinc carboxypeptidases, zinc-binding region 2 signature.

amino acids 68-79

FIGURE 131

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACCATGATCAGCGCAGCCTGGAGC
 ATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAGTTATGGCTAAATCCTG
 TCCATCTGTGTGCTGCGATGCGGGTTTCATTTACTGTAATGATCGCTTTCTGACATCCATTCCAACAG
 GAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAACAACCAAATAAATAATGCTGGGATTCCCTTCAGAT
 TTGAAAACTTGCTGAAAGTAGAAAGAATATACCTATACCACAACAGTTTAGATGAATTTCTACCAACCT
 CCCAAAGTATGTAAAAGAGTTACATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAA
 AAATTCCCTATCTGGAAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCA
 TTCCGAGACAGCAACTATCTCCGACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAATTCCCTGGGGTTT
 GCCCAGGACTATAGAAGAACTACGCTTGATGATAATCGCATATCCACTATTTTCATCACCATCTCTTCAAG
 GTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATGGTTTAGGTGACAAAGTT
 TTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATTCCTGACTGCTGCACCAGTAAACCT
 TCCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATAACCACATCAATCGGGTGCCCCAAATGCTTTTT
 CTTATCTAAGGCAGCTCTATCGACTGGATATGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTT
 GATGATTTGGACAATATAACACAACCTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGAAGATGAAATG
 GGTACGTGACTGGTTACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAA
 AGGTTTCGTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAAGC
 ACCATTGAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCCAAGGACAGTGGCCAGCTCCAGTGAC
 CAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAGGGAGTCCCTCAAGAAAAA
 CAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCAATCTCTTGGAACTTGCTCTACCTATG
 ACTGCTTTGAGACTCAGCTGGCTTAACTGGGCCATAGCCCGGCATTGGATCTATAACAGAAACAATTGT
 AACAGGGGAACGCAGTGAGTACTTGGTCACAGCCCTGGAGCCTGATTCACCCTATAAAGTATGCATGGTTC
 CCATGGAAACCAGCAACCTCTACCTATTTGATGAAACTCCTGTTTGTATTGAGACTGAACTGCACCCCTT
 CGAATGTACAACCCTACAACCACCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACC
 TTTGGCTGCCATCATTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGTATG
 TTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAAAGGATGACTAT
 GCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAAACTTCTTTTCAGATGTTACCAAT
 AAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCATATTCCTCCTAATGGAATGAATCTGT
 ACAAAAAACAATCACAGTGAAAGCAGTAGTAACCGAAGCTACAGAGACAGTGGTATTCCAGACTCAGATCAC
 TCACACTCATGATGCTGAAGGACTCACAGCAGACTTGTGTTTTGGGTTTTTTAAACCTAAGGGAGGTGATG
 GT

FIGURE 132

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTLYL
QNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLSKIPYL
EELHLDDNSVSAVSIEEGAFRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRISTISSPSL
QGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLT AAPVNLPGTNLRKLYLQDNHIN
RVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKWVRDWLQSLPV
KVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQGQWPAPVTKQPD
IKNPKLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWLKLGHSPAFGSITET
IVTGERSEYLVTALEPDSPYKVCMPMETSNNLYLFDETPVC IETETAPLRMYNP TTTTLNREQEKE
PYKNPNLPLAAIIGGAVALVTIAL LALVCWYVHRNGSLFSRNCAYSKGRRRKDDYAEAGTKKDNS
ILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNL YKNNHSESSSNRSYRDSGIPDSDHSHS

Important features of the protein:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 531-552

N-glycosylation sites.

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

Tyrosine kinase phosphorylation site.

amino acids 515-522

N-myristoylation sites.

amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,
640-645

Amidation site.

amino acids 567-570

Leucine zipper pattern.

amino acids 159-180

Phospholipase A2 aspartic acid active site.

amino acids 34-44

FIGURE 133

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCCAGGCTTCTTGGCA
 GCCCTGCCGGGGCCACTTGTCTTCTATGTCTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAGGGCG
 TGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGAGCAGAG
 GCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACCGCTGTGGT
 CCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCTGATCCTGTCT
 GCGCTGGAGAGAGCCACCGTCTTCTAGAACAGAGGCTGCCTGAAATCAACCTGGATGGCATGGT
 GGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCGGGAGAAGTGGGCCCAGGAGCCCCCTGC
 TGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGCTGCCATCCAGAGATCC
 CTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCAGCTGACCCCTCCAGCCCCG
 GTTTTGGAAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTTGGTGTACCCCCACGTTCCGGG
 CCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGTGCAGCTGCTGGGAACCGGGACG
 GACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAGCCTCATGACCAAGCCCCGGCTGCTC
 AGGCTACTGCCTGTCCCACTGCTCTTCTTCTCTGGGCCAGAATGAGGGGATGCACACAGG
 GACCACTCCAACAGAGCCAGGACTATATCAACCTCTTCTGCGCCAACATGATGGACTTGAACCGC
 AGAGCTGAGGCCATCGGATACGCCTACCCTACCCGGGACATCTTCATGGAAAACATCATGTTCTG
 TGGAAATGGGCGGCTTCTCCGACTTCTACAAGCTCCGGTGGCTGGAGGCCATTCTCAGCTGGCAGA
 AACAGCAGGAAGGATGCTTCGGGGAGCCTGATGCTGAAGATGAAGAATTATCTAAAGCTATTCAA
 TATCAGCAGCATTTTTTCGAGGAGAGTGAAAGAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGT
 TGCTCAGGCTGGAGTACAGTGGCGCAATCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGC
 AATTCTCTTGCCCTCATCCTCCCGAGTAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAAT
 TTTTATATTTTTTTAGTAGAGACAGGGTTTCATCATGTTGCTCATGCTGGTCTCGAACTCCTGAT
 CTCAAGAGATCCGCCCACCTCAGGCTCCCAAAGTGTTGGGATTATTAGGTGTGAGCCACCGTGTCTG
 GCTGAAAAGCACTTTCAAAGAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTC
 ATGGGGGCTCTCTCCCCTAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGG
 GTGGCTTCTTATACATCCTGGCAGAATAACCCCCAGCAAACAGAGAGCCACACCCATCCACACCG
 CCACCACCAAGCAGCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCC
 TTTAGTCCTCATCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGG
 ATAAGCAAAGCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGG
 GGCCGGGAGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA

FIGURE 134

MSARGRWEGGRRACRGSGLGLARAQGAERVTSSSEQRPMASLGLLLLLLLLTALPPLWSSSLPGLD
TAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPLSLRV
GMLGEKLEAAIQRSYLHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQDSFSEE
RSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQGPLQQSQD
YINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILSQKQKEGCFG
EPDAEDEELSKAIQYQQHFSSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLPPGFKQFSCILIP
SSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRPPTSGSQSVGL

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 39-56

Tyrosine kinase phosphorylation sites.

amino acids 149-156, 274-282

N-myristoylation sites.

amino acids 10-16, 20-26, 63-69, 208-214

Amidation site.

amino acids 10-14

Glycoprotein hormones beta chain signature 1.

amino acids 230-237

FIGURE 135

GGTCTGAGTGCAGAGCTGCTGTCATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGCTG
CTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGGGAGG
GAGTGGGGTCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAGTTGTTCCAGGGGTGAAGC
CTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTGCGTTTTCCTTAAG
ACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAGTTGTATCTCC
AGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGAGAGCAAGATATG
TGAATTACATCAAAACATCAGAGGTTGTGAGACTGCCCTATCCTCTCCAAATGAAATCTTCAGGT
CCACCTTCTTACTTTATTAAAAGGGAATCGTGGGGCTGGACAGACTTTCTAATGAACCCAATGGT
TATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAGTGGTCAACACAAGTGATC
CTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATTCCAACCATGAGTTGCCTGAT
GTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTGGCAAATCTAGCAGCGGCAGCAG
TAAACAGGCAAAAGTGGGGCTGGCAAAAGGAGGTAGTCAGGCCGTCCAGAGCTGGCATTTCAC
AAACACGGCAACACTGGGTGGCATCCAAGTCTTGGAACCGTGTGAAGCAACTACTATAAACTT
GAGTCATCCCGACGTTGATCTCTTACAACGTGTGTATGTT
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGTAT
GAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTTATAGAAAATGCCATTAATAAATTAT
ATGAACTACTATACATTATGTATATTAATTAAACATCTTAATCCAGAAATCAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAA

FIGURE 136

MAAALWGFFPVLLLLLLLLSGDVQSSEVPGAAAEGSGSGVGIGDRFKIEGRAVVPGVKPDWISAA
RVLVDGEEHVGFLKTDGSFVVDIPSGSYVVEVVSPAYRFDVPRVDITSGKMRARYVNYIKTSE
VVRLPYPLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMVLPLLI FVLLPKVVNTSDPDMRREME
QSMNMLNSNHELDPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

Important features of the protein:

Signal sequence:

amino acids 1-23

Transmembrane domain:

amino acids 161-182

N-glycosylation site.

amino acids 184-187

Glycosaminoglycan attachment sites.

amino acids 37-40, 236-239

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 151-154

N-myristoylation sites.

amino acids 33-38, 36-41, 38-44, 229-234

Amidation site.

amino acids 238-241

ATP/GTP-binding site motif A (P-loop).

amino acids 229-236

FIGURE 137

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGGGA
CCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAAACAA
GTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTTCCAGGC
CTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACGACCTGTGC
CACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAGGCTTGGCAGT
TTTTCTTACTCCTGTGGTCTCCAGATTTCAGGCCTAAGATGAAAGCCTCTAGTCTTGCCTTCAGC
CTTCTCTCTGCTGCGTTTATCTCCTATGGACTCCTTCCACTGGACTGAAGACACTCAATTTGGG
AAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCTGAGATACGGGGCAGTG
TGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACTGAGTCTTTGCAAGACACA
AAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTCTATCTGGACAGGGTATTTAA
AAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGCAGCCTCGCCAATTCCTTTCTTA
CCATCAAGAAGGACCTCCGGCTCTCTCATGCCCACATGACATGCCATTGTGGGGAGGAAGCAATG
AAGAAATACAGCCAGATTCTGAGTCACTTTGAAAAGCTGGAACCTCAGGCAGCAGTTGTGAAGGC
TTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAGGAGACAGAATAGGAGGAAAGTGATGCTG
CTGCTAAGAATATTTCGAGGTCAAGAGCTCCAGTCTTCAATACCTGCAGAGGAGGCATGACCCCAA
ACCACCATCTCTTTACTGTACTAGTCTTGTGCTGGTCACAGTGTATCTTATTTATGCATTACTTG
CTTCCTTGCATGATTGTCTTTATGCATCCCCAATCTTAATTGAGACCATACTTGTATAAGATTTT
TGTAATATCTTTCTGCTATTGGATATATTTATTAGTTAATATATTTATTTATTTTTTGCTATTTA
ATGTATTTATTTTTTTACTTGGACATGAACTTTAAAAAAATTCACAGATTATATTTATAACCTG
ACTAGAGCAGGTGATGTATTTTTATACAGTAAAAAATAACCTTGTAATTTCTAGAAGAGTGG
CTAGGGGGGTTATTCATTTGTATTCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGAT
ATTTGAAATTGAACCAATGACTACTTAGGATGGGTTGTGGAATAAGTTTGTATGTGGAATTGCAC
ATCTACCTTACAATTACTGACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAG
CCAGGAATCCTACACGGCCAGCATGTATTTCTACAAATAAAGTTTTCTTTGCATACCAAAAAAA
AAAAA

FIGURE 138

MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSELRPEIF
SSREAWQFFLLLWSPDFRPKMKASSLAFSLLSAAFYLLWTPSTGLKTLNLGSCVIATNLQEIRNG
FSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVFKNYQTPDHYTLRKIS
SLANSFLTIIKKDLRLSHAHMTCHCGEEAMKKYSQILSHFEKLEPQAAVVKALGELDILLQWMEET
E

Important features of the protein:

Signal peptide:

amino acids 1-42

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 192-195, 225-228

N-myristoylation sites.

amino acids 42-47, 46-51, 136-141

FIGURE 139

CCTGGAGCCGGAAGCGCGGCTGCAGCAGGGCGAGGCTCCAGGTGGGGTTCGGTTCGCATCCAGCC
TAGCGTGTCCACGATGCGGCTGGGCTCCGGGACTTTCGCTACCTGTTGCGTAGCGATCGAGGTGC
TAGGGATCGCGGTCTTCCTTCGGGGATTCTTCCCGGCTCCCGTTCGTTTCCTCTGCCAGAGCGGAA
CACGGAGCGGAGCCCCCAGCGCCCGAACCCTCGGCTGGAGCCAGTTCTAACTGGACCACGCTGCC
ACCACCTCTCTTCAGTAAAGTTGTTATTGTTCTGATAGATGCCTTGAGAGATGATTTTGTGTTTG
GGTCAAAGGGTGTGAAATTTATGCCCTACACAACCTTACCTTGTGGAAGGAGCATCTCACAGT
TTTGTGGCTGAAGCAAAGCCACCTACAGTTACTATGCCTCGAATCAAGGCATTGATGACGGGGAG
CCTTCCTGGCTTTGTGACGTCATCAGGAACCTCAATTCTCCTGCACTGCTGGAAGACAGTGTGA
TAAGACAAGCAAAGCAGCTGGAAAAAGAATAGTCTTTTATGGAGATGAAACCTGGGTAAATTA
TTCCCAAAGCATTTTGTGGAATATGATGGAACAACCTCATTTTTCGTGTGAGATTACACAGAGGT
GGATAATAATGTCACGAGGCATTGGATAAAGTATTAAGAGAGGAGATTGGGACATATTAATCC
TCCACTACCTGGGGCTGGACCACATTGGCCACATTTAGGGCCCAACAGCCCCCTGATTGGGCAG
AAGCTGAGCGAGATGGACAGCGTGTGATGAAGATCCACACCTCACTGCAGTCGAAGGAGAGAGA
GACGCCTTTACCCAATTTGCTGGTTCTTTGTGGTGACCATGGCATGTCTGAAACAGGAAGTCACG
GGGCCTCCTCCACCGAGGAGGTGAATACACCTCTGATTTTAATCAGTTCTGCGTTTGAAAGGAAA
CCCGGTGATATCCGACATCCAAAGCACGTCCAATAGACGGATGTGGCTGCGACACTGGCGATAGC
ACTTGGCTTACCGATTCCAAAAGACAGTGTAGGGAGCCTCCTATTTCCAGTTGTGGAAGGAAGAC
CAATGAGAGAGCAGTTGAGATTTTACATTTGAATACAGTGCAGCTTAGTAAACTGTTGCAAGAG
AATGTGCCGTATATGAAAAAGATCCTGGGTTTGAGCAGTTTAAATGTGAGAAAGATTGCATGG
GAACTGGATCAGACTGTACTTGGAGGAAAAGCATTGAGAGTCTTATTCAACCTGGGCTCCAAGG
TTCTCAGGCAGTACCTGGATGCTCTGAAGACGCTGAGCTTGTCCCTGAGTGCACAAGTGGCCAG
TTCTCACCTTGCTCCTGCTCAGCGTCCCACAGGCACTGCACAGAAAGGCTGAGCTGGAAGTCCCA
CTGTCACTCCTGGGTTTCTCTGCTCTTTTATTTGGTGATCCTGGTTCTTTCGGCCGTTACGCT
CATTGTGTGCACCTCAGCTGAAAGTTCGTGCTACTTCTGTGGCCTCTCGTGGCTGGCGGCAGGCT
GCCTTTCGTTTACCAGACTCTGGTTGAACACCTGGTGTGTGCCAAGTGTGGCAGTGCCCTGGAC
AGGGGGCCTCAGGGAAGGACGTGGAGCAGCCTTATCCCAGGCCTCTGGGTGTCCCGACACAGGTG
TTCACATCTGTGCTGTGAGGTCAGATGCCTCAGTTCTTGGAAAGCTAGGTTCTGCGACTGTTAC
CAAGGTGATTGTAAAGAGCTGGCGGTACAGAGGAACAAGCCCCCAGCTGAGGGGGTGTGTGAA
TCGGACAGCCTCCCAGCAGAGGTGTGGGAGCTGCAGCTGAGGGAAGAAGAGACAATCGGCCTGGA
CACTCAGGAGGGTCAAAGGAGACTTGGTCGCACCACTCATCCTGCCACCCCCAGAATGCATCCT
GCCTCATCAGGTCCAGATTTCTTTCCAAGGCGGACGTTTCTGTTGGAATTCTTAGTCCTTGGCC
TCGGACACCTTCATTTCGTTAGCTGGGGAGTGGTGGTGAGGCAGTGAAGAAGAGGCGGATGGTCAC
ACTCAGATCCACAGAGCCCAGGATCAAGGGACCACTGCAGTGGCAGCAGGACTGTTGGGCCCCC
ACCCCAACCCTGCACAGCCCTCATCCCTCTTGGCTTGGCCGTCAGAGGCCCTGTGCTGAGTGT
CTGACCGAGACACTCACAGCTTTGTATCAGGGCACAGGCTTCCTCGGAGCCAGGATGATCTGTG
CCACGCTTGACCTCGGGCCCCATCTGGGCTCATGCTCTCTCCTGCTATTGAATTAGTACCTAG
CTGCACACAGTATGTAGTTACCAAAGAATAAACGGCAATAATTGAGAAAAAAA

FIGURE 140

MRLGSGTFATCCVAIEVLGIAVFLRGFFPAPVRSSARAEHGAEPPEPSAGASSNWTTLPPPLF
SKVVIVLIDALRDDDFVFGSKGVKFMPTTYLVEKGASHSFVAEAKPPTVTMPRIKALMTGSLPGF
VDVIRNLNSPALLEDVIRQAKAAGKRIVFYGDETWVKLFPKHFVEYDGTTSFFVSDYTEVDNNV
TRHLDKVLKRGDWDILILHYLGLDGHIGHISGPN SPLIGQKLSEMDSVLMKIHTSLQSKERETPLP
NLLVLCGDHGMSETGSHGASSTEEVNTPLILISSAFERKPGDIRHPKHVQ

Important features of the protein:

Signal peptide:

amino acids 1-34

Transmembrane domain:

amino acids 58-76

N-glycosylation sites.

amino acids 56-60, 194-198

N-myristoylation sites.

amino acids 6-12, 52-58, 100-106, 125-131, 233-239, 270-276,
275-281, 278-284

Amidation site.

amino acids 154-158

Cell attachment sequence.

amino acids 205-208

FIGURE 141

GGCACGAGGCAAGCCTTCCAGGTTATCGTGACGCACCTTGAAAGTCTGAGAGCTACTGCCCTACA
GAAAGTTACTAGTGCCCTAAAGCTGGCGCTGGCACTGATGTTACTGCTGCTGTTGGAGTACAAC
TCCCTATAGAAAACAACCTGCCAGCACCTTAAGACCACTCACACCTTCAGAGTGAAGAACTTAAAC
CCGAAGAAATTCAGCATTTCATGACCAGGATCACAAAGTACTGGTCCTGGACTCTGGGAATCTCAT
AGCAGTTCAGATAAAAACTACATACGCCCAGAGATCTTCTTTGCATTAGCCTCATCCTTGAGCT
CAGCCTCTGCGGAGAAAGGAAGTCCGATTCTCCTGGGGGTCTCTAAAGGGGAGTTTTGTCTCTAC
TGTGACAAGGATAAAGGACAAAGTCATCCATCCCTTCAGCTGAAGAAGGAGAACTGATGAAGCT
GGCTGCCCAAAGGAATCAGCACGCCGGCCCTTCATCTTTTATAGGGCTCAGGTGGGCTCCTGGA
ACATGCTGGAGTCGGCGGCTCACCCCGGATGGTTCATCTGCACCTCCTGCAATTGTAATGAGCCT
GTTGGGGTGACAGATAAATTTGAGAACAGGAAACACATTGAATTTTCATTTCAACCAGTTTGCAA
AGCTGAAATGAGCCCCAGTGAGGTCAGCGATTAGGAAACTGCCCCATTGAACGCCTTCCTCGCTA
ATTGAACTAATTGTATAAAAAACACCAAACCTGCTCACT

FIGURE 142

MLLLLLLEYNFPIENNCQHLKTTHTFRVKNLNPKKFSIHDQDHKVLVLD SGNLI AVPDKNYIRPEI
FFALASSLSSASA EKGSPILLGVSKGEFCLYCDKDKGQSHPSLQLKKEKLMKLAAQKESARRPFI
FYRAQVGSWNMLESAAHPGWFICTSCNCNEPVGVTDKFENRKHIEFSFQPVCKAEMSPSEVSD

cAMP- and cGMP-dependent protein kinase phosphorylation site.
amino acids 33-36

N-myristoylation site.
amino acids 50-55, 87-92

Interleukin-1
amino acids 37-182

FIGURE 143

[illegible]

FIGURE 144

MLGLPWKGGLSWALLLLLLLGSQILLIYAWHFHEQRDCDEHNVMARYLPATVEFAVHTFNQQSKDY
YAYRLGHILNSWKEQVESKTVFSMELLLGRTRCGKFEDDIDNCHFQESTELNNTFTCFFTISTRP
WMTQFSLLNKTCLGEGFH

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 117-121, 139-143

N-myristoylation site.

amino acids 9-15

FIGURE 145

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCGATGAGGAAG
CACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGGTCCA
GACGAGGGGCATCAAGCACAGAATCAAGTGGAAACCGGAAGGCCCTGCCCAGCACTGCCCAGATCA
CTGAGGCCCAGGTGGCTGAGAACCGCCCGGAGCCTTCATCAAGCAAGGCCGCAAGCTCGACATT
GACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCGATGGCATCCA
CTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCTGCATCAATGCCA
CCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACCAGCAGGTGCTCTGG
CGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGGAGAGGGGCGCAGGACT
TCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGATCTGGCTCATGGTGAAT
AAGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAATCCTGGCAAGTGACCCAGCT
CTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGCGCGGATGCACTCGCACTGCAAA
TGCCGCTCCACGTATGCGCCCTGGTATGTGCCTGCGTTCTGATAGATGGGGGACTGTGGCTTCT
CCGTCACTCCATTCTCAGCCCCCTAGCAGAGCGTCTGGCACACTAGATTAGTAGTAAATGCTTGAT
GAGAAGAACACATCAGGCACTGCGCCACCTGCTTCACAGTACTTCCCAACAACCTCTTAGAGGTAG
GTGTATTCCCGTTTTACAGATAAGGAAACTGAGGCCCAGAGAGCTGAAGTACTGCACCCAGCATC
ACCAGCTAGAAAGTGGCAGAGCCAGGATTCAACCCTGGCTTGTCTAACCCAGGTTTTCTGCTCT
GTCCAATTCAGAGCTGTCTGGTGATCACTTTATGTCTCACAGGGACCCACATCCAAACATGTAT
CTCTAATGAAATTGTGAAAGCTCCATGTTTAGAAATAAATGAAAACACCTGA

FIGURE 146

MRKHLSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQGRK
LDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNKLHQQ
VLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMLVK

Important features of the protein:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 157-171

N-glycosylation sites.

amino acids 98-102, 110-114

Tyrosine kinase phosphorylation site.

amino acids 76-83

N-myristoylation sites.

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

Amidation site.

amino acids 62-66

FIGURE 147

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCATTT
CCTGATGATTTATAGACTCAAAGAAAACTATGTTTCAGAAGCTCTCTTCTCTTCTGGCCTCCTCT
CTGTCTTCTTTCCCTCTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCAAGCTGG
AAATCTTTCATTTTGCTTGTCAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTTTGAAATTT
CAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCATTGATGTGCTGAGG
TTTGGGGT

FIGURE 148

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGTCE
GLFYEYIA

Important features of the protein:

Signal peptide:

amino acids 1-25

N-myristoylation site.

amino acids 62-68

FIGURE 149

GTCTCCGCGTCACAGGAAC TTCAGCACCCACAGGGCGGACAGCGCTCCCCTCTACCTGGAGACTTGAC
TCCCGCGCGCCCCAACCCCTGCTTATCCCTTGACCGTCGAGTGTGAGAGATCCTGCAGCCGCCAGTCC
CGGCCCTCTCCCGCCCCACCCACCCTCCTGGCTCTTCCTGTTTTTACTCCTCCTTTTCATT CATA
ACAAAAGCTACAGCTCCAGGAGCCCAGCGCCGGGCTGTGACCCAAGCCGAGCGTGGAAGAATGGGGTT
CCTCGGGACCGGCACTTGGATTCTGGTGT TAGTGCTCCCGATTCAAGCTTTCCTCAAACCTGGAGGAA
GCCAAGACAAATCTCTACATAATAGAGAATTAAGTGCAGAAAGACCTTTGAATGAACAGATTGCTGAA
GCAGAAGAAGACAAGATTAAAAAACATATCCTCCAGAAAACAAGCCAGGTGAGAGCAACTATTCTTT
TGTTGATAACTTGAACCTGCTAAAGGCAATAACAGAAAAGGAAAAAATTGAGAAAGAAAGACAATCTA
TAAGAAGCTCCCCACTTGATAATAAGTTGAATGTGGAAGATGTTGATTCAACCAAGAATCGAAAAC TG
ATCGATGATTATGACTCTACTAAGAGTGGATTGGATCATAAATTTCAAGATGATCCAGATGGTCTTCA
TCAACTAGACGGGACTCCTTTAACCGCTGAAGACATTGTCCATAAAATCGCTGCCAGGATTTATGAAG
AAAATGACAGAGCCGTGTTTGACAAGATTGTTTCTAACTACTTAATCTCGGCCTTATCACAGAAAGC
CAAGCACATACACTGGAAGATGAAGTAGCAGAGGTTTTACAAAAATTAATCTCAAAGGAAGCCAACAA
TTATGAGGAGGATCCCAATAAGCCCACAAGCTGGACTGAGAATCAGGCTGGAAAAATACCAGAGAAAG
TGA CTCCAATGGCAGCAATTCAAGATGGTCTTGCTAAGGGAGAAAACGATGAAACAGTATCTAACACA
TTAACCTTGACAAATGGCTTGGAAGGAGAACTAAAACCTACAGTGAAGACAACCTTTGAGGAACTCCA
ATATTTCCCAAATTTCTATGCGCTACTGAAAAGTATTGATT CAGAAAAAGAAGCAAAAGAGAAAGAAA
CACTGATTACTATCATGAAAACACTGATTGACTTTGTGAAGATGATGGTGAAATATGGAACAATATCT
CCAGAAGAAGGTGTTTCCTACCTTGAAAACCTGGATGAAATGATTGCTCTTCAGACCAAAAAACAAGCT
AGAAAAAATGCTACTGACAATATAAGCAAGCTTTTCCCAGCACCATCAGAGAAGAGTCATGAAGAAA
CAGACAGTACCAAGGAAGAAGCAGCTAAGATGGAAAAGGAATATGGAAGCTTGAAGGATTCCACAAAA
GATGATAACTCCAACCCAGGAGGAAAGACAGATGAACCCAAAGGAAAAACAGAAGCCTATTTGGAAGC
CATCAGAAAAAATATTGAATGGTTGAAGAAACATGACAAAAGGGAAATAAAGAAGATTATGACCTTT
CAAAGATGAGAGACTTCATCAATAAACAAGCTGATGCTTATGTGGAGAAAGGCATCCTTGACAAGGAA
GAAGCCGAGGCCATCAAGCGCATTTATAGCAGCCTGTAAAAAATGGCAAAGATCCAGGAGTCTTTCAA
CTGTTTCAGAAAACATAATATAGCTTAAAACACTTCTAATTCTGTGATTAAATTTTTTGACCCAAGG
GTTATTAGAAAGTGCTGAATTTACAGTAGTTAACCTTTTACAAGTGGTTAAACATAGCTTTCTTCCC
GTAAAACTATCTGAAAGTAAAGTTGTATGTAAGCTGAAAAAAAAAAAAAAAAAAAAA

FIGURE 150

MGFLGTGTWILVLVLPIQAFPKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPG
QSNYSFVDNLNLLKAITEKEKIEKERQSIRSSPLDNKLNVEDVDSTKNRKLIDDYDSTKSGLDHK
FQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE
VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPMAAIQDGLAKGENDETVSNTLTLTNGLE
RRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLITIMKTLIDFVKMMVKYGTISPEEGV
SYLENLDEMIALQTKNKLEKNATDNISKLFPAPEKSHEETDSTKEEAAMEKEYGSLKDSTKDD
NSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDK
EEAEAIKRIYSSL

N-glycosylation sites:

amino acids 68-71, 346-349, 350-353

Casein kinase II phosphorylation site:

amino acids 70-73, 82-85, 97-100, 125-128, 147-150, 188-191, 217-
220, 265-268, 289-292, 305-308, 320-323, 326-329, 362-365, 368-
341, 369-372, 382-385, 386-389, 387-390

N-myristoylation sites:

amino acids 143-148, 239-244

FIGURE 151

CGGCTCGAGGCTCCCGCCAGGAGAAAGGAACATTCTGAGGGGAGTCTACACCCTGTGGAGCTCAA
GATGGTCCTGAGTGGGGCGCTGTGCTTCCGAATGAAGGACTCGGCATTGAAGGTGCTTTATCTGC
ATAATAACCAGCTTCTAGCTGGAGGGCTGCATGCAGGGAAGGTCATTAAAGGTGAAGAGATCAGC
GTGGTCCCCAATCGGTGGCTGGATGCCAGCCTGTCCCCCGTCATCCTGGGTGTCCAGGGTGGAA
CCAGTGCCTGTCATGTGGGTGGGGCAGGAGCCGACTCTAACACTAGAGCCAGTGAACATCATGG
AGCTCTATCTTGGTGCCAAGGAATCCAAGAGCTTACCTTCTACCGGCGGGACATGGGGCTCACC
TCCAGCTTCGAGTCGGCTGCCTACCGGGCTGGTTCCTGTGCACGGTGCCTGAAGCCGATCAGCC
TGTCAGACTCACCCAGCTTCCCGAGAATGGTGGCTGGAATGCCCCATCACAGACTTCTACTTCC
AGCAGTGTGACTAGGGCAACGTGCCCCCAGAACTCCCTGGGCAGAGCCAGCTCGGGTGAGGGGT
GAGTGGAGGAGACCCATGGCGGACAATCACTCTCTCTGCTCTCAGGACCCCCACGTCTGACTTAG
TGGGCACCTGACCACTTTGTCTTCTGGTTCCAGTTTGGATAAATTCTGAGATTTGGAGCTCAGT
CCACGGTCTCCCCACTGGATGGTGCTACTGCTGTGGAACCTTGTA AAAACCATGTGGGGTAA
CTGGGAATAACATGAAAAGATTTCTGTGGGGGTGGGGTGGGGGAGTGGTGGGAATCATTCCTGCT
TAATGGTAAC TGACAAGTGTTACCCTGAGCCCCGAGGCCAACC CATCCCCAGTTGAGCCTTATA
GGGTGAGTAGCTCTCCACATGAAGTCTGTCACTCACCCTGTGCAGGAGAGGGAGGTGGTCATA
GAGTCAGGGATCTATGGCCCTTGGCCAGCCCCACCCCTTCCCTTTAATCCTGCCACTGTCATA
TGCTACCTTTCTATCTCTTCCCTCATCATCTTGTTGTGGGCATGAGGAGGTGGTGATGTCAGAA
GAAATGGCTCGAGCTCAGAAGATAAAAAGATAAGTAGGGTATGCTGATCCTCTTTTAAAAACCAA
GATACAATCAAAATCCAGATGCTGGTCTCTATTCCCATGAAAAAGTGCTCATGACATATTGAGA
AGACCTACTTACAAAGTGGCATATATTGCAATTTATTTTAATTAAAAAGATACCTATTTATATATT
TCTTTATAGAAAAAGTCTGGAAGAGTTTACTTCAATTGTAGCAATGTCAGGGTGGTGGCAGTAT
AGGTGATTTTCTTTTAATTCTGTAAATTTATCTGTATTTTCTAATTTTTCTACAATGAAGATGA
ATTCTTGTATAAAAAATAAGAAAAGAAATTAATCTTGAGGTAAGCAGAGCAGACATCATCTCTGA
TTGTCTCAGCCTCCACTTCCCCAGAGTAAATTCAAATTGAATCGAGCTCTGCTGCTCTGGTTGG
TTGTAGTAGTGATCAGGAAACAGATCTCAGCAAAGCCACTGAGGAGGAGGCTGTGCTGAGTTTGT
GTGGCTGGAATCTCTGGGTAAGGAACCTAAAGAACA AAAATCATCTGGTAATTTCTTCTAGAA
GATCACAGCCCCCTGGGATTCCAAGGCATTGGATCCAGTCTCTAAGAAGGCTGCTGTACTGGTTGA
ATTGTGTCCCCCTCAAATTCACATCCTTCTTGGAATCTCAGTCTGTGAGTTTATTTGGAGATAAG
GTCTCTGCAGATGTAGTTAGTTAAGACAAGGTCATGCTGGATGAAGGTAGACCTAAATTCAATAT
GACTGGTTTTCTTGTATGAAAAGGAGAGGACACAGAGACAGAGGAGACGCGGGGAAGACTATGTA
AAGATGAAGGCAGAGATCGGAGTTTTGCAGCCACAAGCTAAGAAACACCAAGGATTGTGGCAACC
ATCAGAAGCTTGGAAGAGGCAAAGAAGAAATTTCTTCCCTAGAGGCTTTAGAGGGATAACGGCTCTG
CTGAAACCTTAATCTCAGACTTCCAGCCTCCTGAACGAAGAAAGAATAAATTTTCGGCTGTTTTAA
GCCACCAAGGATAATTGGTTACAGCAGCTCTAGGAACTAATACAGCTGCTAAAATGATCCCTGT
CTCCTCGTGTTTACATTCTGTGTGTGTCCCCCTCCACAATGTACCAAAGTTGTCTTTGTGACCAA
TAGAATATGGCAGAAGTGATGGCATGCCACTTCCAAGATTAGGTTATAAAAAGACACTGCAGCTTC
TACTTGAGCCCTCTCTCTCTGCCACCCACGCCCCCAATCTATCTTGGCTCACTCGCTCTGGGGG
AAGCTAGCTGCCATGCTATGAGCAGGCCATATAAAGAGACTTACGTGGTAAAAAATGAAGTCTCCT
GCCCACAGCCACATTAGTGAACCTAGAAGCAGAGACTCTGTGAGATAATCGATGTTTGTGTTTT
AAGTTGCTCAGTTTTGGTCTAACTTGTTATGCAGCAATAGATAAATAATATGCAGAGAAAGAG

FIGURE 152

MVLSGALCFRMKDSALKVLYLHNNQLLAGGLHAGKVIKGEESISVVPNRWLDASLSPVILGVQGG
QCLSCGVGQEPTLTLEPVNIMELYLGAKESKSFTFYRRDMGLTSSFESAAYPGWFLCTVPEADQP
VRLTQLPENGGWNAIPITDFYFQQCD

N-myristoylation sites.

amino acids 29-34, 30-35, 60-65, 63-68, 73-78, 91-96, 106-111

Interleukin-1 signature.

amino acids 111-131

Interleukin-1 proteins.

amino acids 8-29, 83-120, 95-134, 64-103

FIGURE 153

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCAATTGGCCGC
CCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTCTCTTGG
CCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGACAAGTCCAAC
TTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAGCTTGGCTGATAA
CAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTATGAGTGAGCGCTGCT
ATCTGATGAAGCAGGTGCTGAACTTCACCCTTGAAGAAGTGCTGTTCCCTCAATCTGATAGGTTC
CAGCCTTATATGCAGGAGGTGGTGCCCTTCCCTGGCCAGGCTCAGCAACAGGCTAAGCACATGTCA
TATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAAAGCTGAAGGACACAGTGAAAAAGC
TTGGAGAGAGTGGAGAGATCAAAGCAATTGGAGAACTGGATTTGCTGTTTATGTCTCTGAGAAAT
GCCTGCATTTGACCAGAGCAAAGCTGAAAAATGAATAACTAACCCCTTTCCCTGCTAGAAATAA
CAATTAGATGCCCCAAAGCGATTTTTTTTAACCAAAAGGAAGATGGGAAGCCAAACTCCATCATG
ATGGGTGGATTCCAAATGAACCCCTGCGTTAGTTACAAAGGAAACCAATGCCACTTTTGTTTATA
AGACCAGAAGGTAGACTTTCTAAGCATAGATATTTATTGATAACATTTTCATTGTAACCTGGTGTTT
TATACACAGAAAACAATTTATTTTTTAAATAATTGTCTTTTTCCATAAAAAAGATTACTTTCCAT
TCCTTTAGGGGAAAAAACCCCTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAA
TGATTTTATTATTATTATAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTAT
AGAAACATCATTCGATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTAT
AGAGCTATAACATGTTTATTTGACCTCAATAAACACTTGGATATCCC

FIGURE 154

MAALQKSVSSFLMGTLATSCLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKEASL
ADNNTDVRLIGEKLFGVSMSERCYLMKQVLNFTLEEVLFPQSDRFQPYMQEVVPFLARLSNRLS
TCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

FIGURE 155

GGCTTGCTGAAAATAAAATCAGGACTCCTAACCTGCTCCAGTCAGCCTGCTTCCACGAGGCCTGT
CAGTCAGTGCCCGACTTGTGACTGAGTGTGCAGTGCCAGCATGTACCAGGTCAGTGCAGAGGGC
TGCTTGAGGGCTGTGCTGAGAGGGAGAGGAGCAGAGATGCTGCTGAGGGTGGAGGGAGGCCAAGC
TGCCAGGTTTGGGGCTGGGGGCCAAGTGGAGTGAGAACTGGGATCCCAGGGGGAGGGTGCAGAT
GAGGGAGCGACCCAGATTAGGTGAGGACAGTTCTCTCATTAGCCTTTTCCTACAGGTGGTTGCAT
TCTTGGCAATGGTCATGGGAACCCACACCTACAGCCACTGGCCCAGCTGCTGCCCCAGCAAAGGG
CAGGACACCTCTGAGGAGCTGCTGAGGTGGAGCACTGTGCCTGTGCCTCCCCTAGAGCCTGCTAG
GCCCCAACCGCCACCCAGAGTCCTGTAGGGCCAGTGAAGATGGACCCCTCAACAGCAGGGCCATCT
CCCCCTGGAGATATGAGTTGGACAGAGACTTGAACCGGCTCCCCCAGGACCTGTACCACGCCCGT
TGCTGTGCCCCGCACTGCGTCAGCCTACAGACAGGCTCCCACATGGACCCCCGGGGCAACTCGGA
GCTGCTCTACCACAACCAGACTGTCTTCTACAGGCGGCCATGCCATGGCGAGAAGGGCAGCCACA
AGGGCTACTGCCTGGAGCGCAGGCTGTACCGTGTTCCTTAGCTTGTGTGTGTGTGCGGCCCCGT
GTGATGGGCTAGCCGGACCTGCTGGAGGCTGGTCCCTTTTGGGAAACCTGGAGCCAGGTGTACA
ACCACTTGCCATGAAGGGCCAGGATGCCAGATGCTTGGCCCCCTGTGAAGTGCTGTCTGGAGCAG
CAGGATCCCGGGACAGGATGGGGGGCTTTGGGGAAAACCTGCACTTCTGCACATTTTGAAAAGAG
CAGCTGCTGCTTAGGGCCGCCGGAAGCTGGTGTCTGTCAATTTCTCTCAGGAAAGGTTTTCAA
GTTCTGCCATTTCTGGAGGCCACCACTCCTGTCTCTTCCTCTTTTCCCATCCCCTGCTACCCTG
GCCCAGCACAGGCACTTTCTAGATATTTCCCCCTTGCTGGAGAAGAAAGAGCCCCTGGTTTTATT
TGTTTGTTTACTCATCACTCAGTGAGCATCTACTTTGGGTGCATTCTAGTGTAGTTACTAGTCTT
TTGACATGGATGATTCTGAGGAGGAAGCTGTTATTGAATGTATAGAGATTTATCCAAATAAATAT
CTTTATTTAAAAATGAAAAA

FIGURE 156

MRERPRLGEDSSLISLFLQVVAFLAMVMGHTYSHWPSCCPSKGQDTSEELLRWSTVPVPPLEPA
RPNRHPESCRASEDGPLNSRAISPWRYELDRDLNRLPQDLYHARCLCPHCVSLQTGSHMDPRGNS
ELLYHNQTVFYRRPCHGEKGTHKGYCLERRLYRVSLACVCVRPRVMG

Important features of the protein:

Signal peptide:

amino acids 1-32

N-glycosylation site.

amino acids 136-140

Tyrosine kinase phosphorylation site.

amino acids 127-135

N-myristoylation sites.

amino acids 44-50, 150-156

FIGURE 157

CCGGCGATGTCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAGCC
GACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTAATCC
CCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAACACTAGTGTTGCAACAGGGGACTATTCA
ATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAGGCCACCAA
GATTTGTGTGACGGGCAAAAGCAACTTCCAGTCTTACAGCTGTGTGAGGTGCAATTACACAGAGG
CCTTCCAGACTCAGACCAGACCCCTCTGGTGGTAAATGGACATTTTCTACATCGGCTTCCCTGTA
GAGCTGAACACAGTCTATTTTCATTGGGGCCATAATATTCCTAATGCAAATATGAATGAAGATGG
CCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATAATGAAATATAAAAAA
AGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGTAAGAAGAATGAGGAGACA
GTAGAAGTGAACCTTACAACCACTCCCCTGGGAAACAGATACATGGCTCTTATCCAACACAGCAC
TATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAACAAACGCGAGCTTCAGTGGTGA
TTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTGACTCCATATTTTCTACTTGTGGC
AGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGCCACAAACAGGCGTCCCTTTCCCTCT
GGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCTCTCCTCCTGCTGTCTCTGCTGGTGGCCA
CATGGGTGCTGGTGGCAGGGATCTATCTAATGTGGAGGCACGAAAGGATCAAGAAGACTTCCTTT
TCTACCACCACACTACTGCCCCCATTAAGGTTCTTGTGGTTTACCCATCTGAAATATGTTTCCA
TCACACAATTTGTTACTTCACTGAATTTCTTCAAACCATTCAGAGTGAAGTGCAGAGTGCCTTGAAA
AGTGGCAGAAAAAGAAAATAGCAGAGATGGGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCA
GCAGACAAAGTCGTCTTCTTCTTCCAATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAA
GAGCGAGGGCAGTCCAGTGAGAACTCTCAAGACCTCTTCCCCCTTGCCCTTAACCTTTTCTGCA
GTGATCTAAGAAGCCAGATTTCATCTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACA
AAAGACGATTACAATGCTCTCAGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTT
CTGTGCAGAACTTCTCCATGTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACG
ATGGCTGCTGCTCCTTGTAG

FIGURE 158

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLPGLRDLRVEPVTTSVATGDYSILMNVS
WVLRADASIRLLKATKICVTGKSNFQSYSCVRCNYTEAFQTQTRPSGGKWTF SYIGFPVELNTVYFIGAHNIP
NANMNEDGPSMSVNFTSPGCLDHIMKYKKKCKVAGSLWDPNITACKNEETVEVNFTTTPLGNRYMALIQH
STIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTPYFPTCGSDCIRHKGTVVLCPTGVPFPLDNNK
SKPGGWLPLLLLSLLVATWVLVAGIYLMWRHERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFL
QNHCRSEVILEKWQKKIAEMGPVQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGSPSENSQDLFPLA
FNLFCSDLRSQIHLHKYVVVYFREIDTKDDYNALSVC PKYHLMKDATAFCAELLHV KQQVSAGKRSQACHD
GCCSL

Important features of the protein:

Signal peptide:

amino acids 1-14

Transmembrane domain:

amino acids 290-309

N-glycosylation sites.

amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201 and 283
- 287

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 228 - 232 and 319 - 323

Casein kinase II phosphorylation sites.

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

N-myristoylation site.

amino acids 116-122

Amidation site.

amino acids 488-452

FIGURE 159

AGCCACCAGCGCAACATGACAGTGAAGACCCTGCATGGCCCAGCCATGGTCAAGTACTTGCTGCT
GTCGATATTGGGGCTTGCCTTTCTGAGTGAGGCGGCAGCTCGGAAAATCCCCAAAGTAGGACATA
CTTTTTTCCAAAAGCCTGAGAGTTGCCCGCCTGTGCCAGGAGGTAGTATGAAGCTTGACATTGGC
ATCATCAATGAAAACCAGCGCGTTTCCATGTCACGTAACATCGAGAGCCGCTCCACCTCCCCCTG
GAATTACACTGTCACTTGGGACCCCAACCGGTACCCCTCGGAAGTTGTACAGGCCCAGTGTAGGA
ACTTGGGCTGCATCAATGCTCAAGGAAAGGAAGACATCTCCATGAATTCGGTTCCCATCCAGCAA
GAGACCCTGGTCGTCCGGAGGAAGCACCAAGGCTGCTCTGTTTCTTTCCAGTTGGAGAAGGTGCT
GGTGACTGTTGGCTGCACCTGCGTCACCCCTGTCATCCACCATGTGCAGTTAAGAGGTGCATATCC
ACTCAGCTGAAGAAG

FIGURE 160

MTVKTLHGPA MVKYLLLSILGLAFLSEAAARKIPKVGHTFFQKPESCPPVPGGSMKLDIGIINEN
QRVSM SRNIESRSTSPWNYT VTWDPNRY PSEVVQAQCRNLGCINAQGKEDISMNSVPIQQETLVV
RRKHQGC SVSFQLEKVLVTVGCTCVTPVIHHVQ

Signal sequence:

amino acids 1-30

N-glycosylation site.

amino acids 83-87

N-myristoylation sites.

amino acids 106-111, 136-141

FIGURE 161

A C A C T G G C C A A A C A A A A A C G A A A G C A C T C C G T G C T G G A A G T A G G A G G A G A G T C A G G A C T C C C A G G
 A C A G A G A G T G C A C A A A C T A C C C A G C A C A G C C C C C T C C G C C C C T C T G G A G G C T G A A G A G G G A T T C
 C A G C C C C T G C C A C C C A C A G A C A C G G G C T G A C T G G G G T G T C T G C C C C C C T T G G G G G G G G C A G C A C
 A G G G C C T C A G G C C T G G G T G C C A C C T G G C A C C T A G A A G A T G C C T G T G C C C T G G T T C T T G C T G T C C T
 T G G C A C T G G G C C G A A G C C C A G T G G T C C T T T C T C T G G A G A G G C T T G T G G G C C T C A G G A C G C T A C C
 C A C T G C T C T C C G G G C C T C T C C T G C C G C C T C T G G G A C A G T G A C A T A C T C T G C C T G C C T G G G G A C A T
 C G T G C C T G C T C C G G G C C C C G T G C T G G C G C C T A C G C A C C T G C A G A C A G A G C T G G T G C T G A G G T G C C
 A G A A G G A G A C C G A C T G T G A C C T C T G T C T G C G T G T G G C T G T C C A C T T G G C C G T G C A T G G G C A C T G G
 G A A G A G C C T G A A G A T G A G G A A A G T T T G G A G G A G C A G C T G A C T C A G G G G T G G A G G A G C C T A G G A A
 T G C C T C T C C A G G C C C A A G T C G T G C T C T C C T T C C A G G C C T A C C C T A C T G C C C G C T G C G T C C T G C
 T G G A G G T G C A A G T G C C T G C C C T T G T G C A G T T T G G T C A G T C T G T G G G C T C T G T G G T A T A T G A C
 T G C T T C G A G G C T G C C C T A G G G A G T G A G G T A C G A A T C T G G T C C T A T A C T C A G C C C A G G T A C G A G A A
 G G A A C T C A A C C A C A C A C A G C A G C T G C C T G C C C T G C C C T G G C T C A A C G T G C A G C A G A T G G T G A C A
 A C G T G C A T C T G G T T C T G A A T G T C T C T G A G G A G C A G C A C T T C G G C C T C T C C C T G T A C T G G A A T C A G
 G T C C A G G G C C C C C A A A A C C C G G T G G C A C A A A A C C T G A C T G G A C C G C A G A T C A T T A C C T T G A A
 C C A C A C A G A C C T G G T T C C C T G C C T C T G T A T T C A G G T G T G G C C T C T G G A A C C T G A C T C C G T T A G G A
 C G A A C A T C T G C C C C T T C A G G G A G G A C C C C G C G C A C A C C A G A A C C T C T G G C A A G C C G C C C G A C T G
 C G A C T G C T G A C C C T G C A G A G C T G G C T G C T G G A C G C A C C G T G C T C G C T G C C C G C A G A A G C G G C A C T
 G T G C T G G C G G G C T C C G G G T G G G G A C C C C T G C C A G C C A C T G G T C C C A C C G C T T T C C T G G G A G A A C G
 T C A C T G T G G A C A A G G T T C T C G A G T T C C C A T T G C T G A A A G G C C A C C C T A A C C T C T G T G T T C A G G T G
 A A C A G C T C G G A G A A G C T G C A G C T G C A G G A G T G C T T G T G G G C T G A C T C C C T G G G G C C T C T C A A A G A
 C G A T G T G C T A C T G T T G G A G A C A C G A G G C C C C A G G A C A A C A G A T C C C C T C T G T G C C T T G G A A C C C A
 G T G G C T G T A C T T C A C T A C C C A G C A A A G C C T C C A C G A G G G C A G C T C G C C T T G G A G A G T A C T T A C T A
 C A A G A C C T G C A G T C A G G C C A G T G T C T G C A G C T A T G G G A C G A T G A C T T G G G A G C G C T A T G G G C C T G
 C C C C A T G G A C A A A T A C A T C C A C A A G C G C T G G G C C C T C G T G T G G C T G G C C T G C C T A C T C T T T G C C G
 C T G C G C T T T C C C T C A T C C T C C T T C T C A A A A A G G A T C A C G C G A A A G G G T G G C T G A G G C T C T T G A A A
 C A G G A C G T C C G C T C G G G G G C G G C C G C A G G G G C C G C G C G G C T C T G C T C C T C A C T C A G C C G A T G A
 C T C G G G T T T C G A G C G C C T G G T G G G C G C C C T G G C G T C G G C C C T G T G C C A G C T G C C G C T G C G C T G G
 C C G T A G A C C T G T G G A G C C G T C G T G A A C T G A G C G C G C A G G G G C C C G T G G C T T G G T T T C A C G C G C A G
 C G G C G C C A G A C C C T G C A G G A G G G C G G C G T G G T G G T C T T G C T C T T C T C C C G G T G C G G T G G C G C T
 G T G C A G C G A G T G G C T A C A G G A T G G G G T G T C C G G G C C G G G G C G C A C G C C C G C A C G A C G C C T T C C
 G C G C C T C G C T C A G C T G C G T G C T G C C C G A C T T C T T G C A G G G C C G G G C G C C C G G C A G C T A C G T G G G G
 G C C T G C T T C G A C A G G C T G C T C C A C C C G A C G C C G T A C C C G C C T T T T C C G A C C G T G C C C G T C T T
 C A C A C T G C C C T C C C A A C T G C C A G A C T T C C T G G G G G C C C T G C A G C A G C C T C G C G C C C G C G T T C C G
 G G C G G C T C C A A G A G A G A G C G G A G C A A G T G T C C C G G G C C C T T C A G C C A G C C C T G G A T A G C T A C T T C
 C A T C C C C C G G G G A C T C C C G C G C C G G G A C G C G G G T G G G A C C A G G G G C G G G A C C T G G G G C G G G G A
 C G G G A C T T A A A T A A A G G C A G A C G C T G T T T T T C T A A A A A A A

FIGURE 162

MPVPWFLLSLALGRSPVVLSELRLVGPQDATHCSPGLSCRLWDS DILCLPGDIVPAPGPVLAPTHLQTELV
LRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRNASLQAQVVL SFQAYPTARCVLLEV
QVPAALVQFGQSVG SVVYDCFEAALGSEVRIWSYTQPRYEKELNHTQQLPALPWLNV SADGDNVHLVLNV
EEQHFGLSLYWNQVQGPPKPRWHKNLTGPQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQ
LWQAARLRLTLQSWLLDAPCSLPAAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCVQ
VNSSEKLQEQECLWADSLGPLKDDVLLLETRGPQDNRS LCALEPSGCTSLPSKASTRAARLGEYLLQDLQS
GQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILLKKDHAKGWLRLKQDVRSGAAARG
RAALLLYSADDSGFERLVGALASALCQLPLRVAVDLWSRRELSAQGPVAFWFAQRQTLOEGGVVLLFSP
GAVALCSEWLQDGVSGPGAHGPHDAFRASLSCVLPDFLQGRAPGSYVGACFDRLLHPDAVPALFRTVPVET
LPSQLPDFLGALQQPRAPRSGRLOQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT

Signal sequence:

amino acids 1-20

Transmembrane domain.

amino acids 453-475

N-glycosylation sites.

amino acids 118-121, 186-189, 198-201, 211-214, 238-241, 248-251,
334-337, 357-360, 391-394

Glycosaminoglycan attachment site.

amino acids 583-586

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 552-555

N-myristoylation sites.

amino acids 107-112, 152-157, 319-324, 438-443, 516-521, 612-617,
692-697, 696-701, 700-705

FIGURE 163

GGGAGGGCTCTGTGCCAGCCCCGATGAGGACGCTGCTGACCATCTTGACTGTGGGATCCCTGGCT
GCTCACGCCCCCTGAGGACCCCTCGGATCTGCTCCAGCACGTGAAATTCAGTCCAGCAACTTTGA
AAACATCCTGACGTGGGACAGCGGGCCAGAGGGCACCCAGACACGGTCTACAGCATCGAGTATA
AGACGTACGGAGAGAGGGGACTGGGTGGCAAAGAAGGGCTGTCAGCGGATCACCCGGAAGTCCTGC
AACCTGACGGTGGAGACGGGCAACCTCACGGAGCTCTACTATGCCAGGGTCACCGCT
GTCAGTGCGGGAGGCCGGTCAGCCACCAAGATGACTGACAGGTTTCAGCTCTCTGCAGCACACTAC
CCTCAAGCCACCTGATGTGACCTGTATCTCCAAAGTGAGATCGATTTCAGATGATTGTTTCATCCTA
CCCCACGCCAATCCGTGCAGGCGATGGCCACCGGCTAACCTGGAAGACATCTTCCATGACCTG
TTCTACCACTTAGAGCTCCAGGTCAACCGCACCTACCAATGCACCTTGGAGGGAAGCAGAGAGA
ATATGAGTTCTTCGGCCTGACCCCTGACACAGAGTTCTTGGCACCATCATGATTTGCGTTCCCA
CCTGGGCCAAGGAGAGTGGCCCCCTACATGTGCCGAGTGAAGACACTGCCAGACCGGACATGGACC
TACTCCTTCTCCGGAGCCTTCTGTTCTCCATGGGCTTCTCTCGTCGCAGTACTCTGCTACCTGAG
CTACAGATATGTACCAAGCCGCTGCACCTCCCAACTCCCTGAACGTCCAGCGAGTCTGACTT
TCCAGCCGCTGCGCTTCATCCAGGAGCACGTCTGATCCCTGTCTTTGACCTCAGCGGCCCCAGC
AGTCTGGCCCAGCCTGTCCAGTACTCCAGATCAGGGTGTCTGGACCCAGGGAGCCCGCAGGAGC
TCCACAGCGGCATAGCCTGTCCGAGATCACCTACTTAGGGCAGCCAGACATCTCCATCTCCAGC
CCTCCAACGTGCCACCTCCCCAGATCCTCTCCCCACTGTCTATGCCCCAAACGCTGCCCCCTGAG
GTCGGGCCCCCATCCTATGCACCTCAGGTGACCCCCGAAGCTCAATTCCTATTCTACGCCCCACA
GGCCATCTCTAAGGTCCAGCCTTCTCTCTATGCCCCCAAGCCACTCCGGACAGCTGGCCTCCCT
CCTATGGGGTATGCATGGAAGGTTCTGGCAAAGACTCCCCCACTGGGACACTTTCTAGTCCTAAA
CACCTTAGGCCTAAAGGTGAGCTTCAGAAAGAGCCACCAGCTGGAAGCTGCATGTTAGGTGGCCT
TTCTCTGCAGGAGGTGACCTCCTTGGCTATGGAGGAATCCCAAGAAGCAAAATCATTGCACCAGC
CCCTGGGGATTTGCACAGACAGAACATCTGACCCAAATGTGCTACACAGTGGGGAGGAAGGGACA
CCACAGTACCTAAAGGGCCAGCTCCCCCTCCTCTCCTCAGTCCAGATCGAGGGCCACCCCATGTC
CCTCCCTTTGCAACCTCCTTCCGGTCCATGTTCCCCCTCGGACCAAGGTCCAAGTCCCTGGGGCC
TGCTGGAGTCCCTTGTGTGTCCCAAGGATGAAGCCAAGAGCCCAGCCCTGAGACCTCAGACCTG
GAGCAGCCCACAGAACTGGATTCTCTTTTCAGAGGCTGGCCCTGACTGTGCAGTGGGAGTCTTG
AGGGGAATGGGAAAGGCTTGGTGCTTCTCCTGTCCCTACCCAGTGTACATCCTTGGCTGTCA
ATCCCATGCCCTGCCCATGCCACACACTCTGCGATCTGGCCTCAGACGGGTGCCCTTGAGAGAAGC
AGAGGGAGTGGCATGCAGGGCCCCCTGCCATGGGTGCGCTCCTCACCGGAACAAAGCAGCATGATA
AGGACTGCAGCGGGGGAGCTCTGGGGAGCAGCTTGTGTAGACAAGCGCGTGTCTCGCTGAGCCCTG
CAAGGCAGAAATGACAGTGCAAGGAGGAAATGCAGGGGAACTCCCGAGGTCCAGAGCCCCACCTC
CTAACACCATGGATTCAAAGTGCTCAGGGAATTTGCCTCTCCTTGCCCCATTCTTGCCAGTTTC
ACAATCTAGCTCGACAGAGCATGAGGCCCCCTGCCTCTTCTGTCAATTGTTCAAAGGTGGGAAGAGA
GCCTGGAAAAGAACCAGGCCTGGAAAAGAACCAGAAGGAGGCTGGGCAGAACCAGAACAACCTGC
ACTTCTGCCAAGGCCAGGGCCAGCAGGACGGCAGGACTCTAGGGAGGGGTGTGGCCTGCAGCTCA
TTCCAGCCAGGGCAACTGCCTGACGTTGCACGATTTTCAGCTTCATTCTCTGATAGAACAAAGC
GAAATGCAGGTCCACCAGGGAGGGAGACACACAAGCCTTTTCTGCAGGCAGGAGTTTCAGACCCT
ATCCTGAGAATGGGGTTTGAAAGGAAGGTGAGGGCTGTGGCCCCCTGGACGGGTACAATAACACAC
TGTAATGATGTCACAACTTTGCAAGCTCTGCCTTGGGTTTCAGCCCATCTGGGCTCAAATTCCAGC
CTCACCCTCACAAGCTGTGTGACTTCAAACAAATGAAATCAGTGCCAGAACCTCGGTTTCTCTC
ATCTGTAATGTGGGGATCATAACACCTACCTCATGGAGTTGTGGTGAAGATGAAATGAAGTCATG
TCTTTAAAGTGCTTAATAGTGCCTGGTACATGGGCAGTGCCCAATAAACGGTAGCTATTTAAAAA
AAAAAAA

FIGURE 164

MRTLLTILTVGSLAAHAPEDPSDLLQHVKFQSSNFENILTWDSGPEGTPDTVYSIEYKTYGERDW
VAKKGCQRITRKSCNLTVETGNLTLEYARVTAVSAGGRSATKMTDRFSSLQHTTLKPPDVTCIS
KVRSIQMIVHPTPTPIRAGDGHRLTLEDIFHDLFYHLELQVNRTYQMH LGGKQREYEFFGLTPDT
EFLGTIMICVPTWAKESAPYMCRVKTL PDRTWTYSFSGAFLFSMGFLVAVLCYLSYRYVTKPPAP
PNSLVQVRVLT FQPLRFIQEHVLIPVFDLSGPSSLAQPVOYSQIRVSGPREPAGAPQRHSLSEIT
YLGQPDISILQPSNVPPPQILSPLSYAPNAAPEVGPPSYAPQVTPEAQFPFYAPQAISKVQPSSY
APQATPDSWPSPSYGVCMEGSGKDSPTGTLSSPKHLRPKGQLQKEPPAGSCMLGGLSLQEVTSIAM
EESQEAQSLHQPLGICTDRSDPNVLHSGEETGPQY LKGQLPLLSSVQIEGHPMSLPLQPPSGPC
SPSDQGPSPWG LLES LVC PKDEAKSPAPETSDLEQPTELDSLFRGLALT VQWES

Signal sequence.

amino acids 1-17

Transmembrane domain.

amino acids 233-250

N-glycosylation sites.

amino acids 80-83, 87-90, 172-175

N-myristoylation sites.

amino acids 11-16, 47-52, 102-107, 531-536, 565-570

FIGURE 165

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCGCGGTGGCCACAACATGG
CTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCC GG GGCAG
TCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTGCAGCAT
GTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGGCCCTGATTGTCTGTTTTGTGAATTTTA
AAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAAC TTTGGGCTGGA
AGTGTGTAACACAGTTTTTGGATATTTTCCAAAAGATTTGATCAAGGTACTTCATAAATACACGGA
AGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAGGAGGAAGAGATGATT
TTAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGA ACTGGAGGACTCTGTACCTGAAGAG
TCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGAGGAGTCTCGGGGGCGTGA
ACTTGACCCCTGTGCCTGAGCCCAGGCATTTCAGAGCTGATTCAGAGGATGGAGAAGGTGCTTTCT
CAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGAGAGCCACCCTCACACCAGCGGT
CCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACACTTTTGAAGAAATTCTGCACGATAA
ATTGAAAGTGCCGGGAAGCGAAAGCAGAACTGGCAATAGTTCTCCTGCCTCGGTGGAGCGGGAGA
AGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAGTCAGAGAGGAAGTGGACAGTGC GTTATT
CATTACAGCAAAGGATTTTCGTTGGCATCAAAAATCTAAGTTTGTTTTACAAAGATTGTTTTTTAGTA
CTAAGCTGCCTTGGCAGTTTGCATTTTTTGAGCCAAACAAAAATATATTATTTTCCCTTCTAAGTA
AAAAAAAAAAAAAAAAAAAAA

FIGURE 166

MAAAPGLLEFWLFLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPDCRFVN
FKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPAETDFVCFEGGRD
DFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPEESRGRELDVPEPEAFRADSEDGEGA
FSESTEG LQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDKLKVP GSESRTGNSSPASVER
EKTDAYKVLKTEMSQRGSGQCVIHYSKGFRWHQNL SLFYKDCF

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 294-298

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-34

Tyrosine kinase phosphorylation site.

amino acids 67-76

N-myristoylation sites.

amino acids 205-211, 225-231, 277-283

Amidation site.

amino acids 28-32

FIGURE 167

CCAGGACCAGGGCGCACCGGCTCAGCCTCTCACTTGTCTCAGAGGCCGGGGAAGAGAAGCAAAGCGC
AACGGTGTGGTCCAAGCCGGGGCTTCTGCTTCGCCTCTAGGACATACACGGGACCCCTAACTTC
AGTCCCCCAAACGCGCACCTTCGAAGTCTTGAATCCAGCCCCGCACATCCACGCGCGGCACAGG
CGCGGCAGGCGGCAGGTCCCGGCCGAAGGCGATGCGCGCAGGGGGTGGGGCAGCTGGGCTCGGGC
GGCGGGAGTAGGGCCCGGCAGGGAGGCAGGGAGGCTGCATATTCAGAGTCGCGGGCTGCGCCCTG
GGCAGAGGCCGCCCTCGCTCCACGCAACACCTGCTGCTGCCACCGCGCCGCGCATGAGCCGCGTGG
TCTCGCTGCTGCTGGGCGCCGCGCTGCTCTGCGGCCACGGAGCCTTCTGCCGCCGCGTGGTCTCAGC
GGCCAAAAGGTGTGTTTTGCTGACTTCAAGCATCCCTGCTACAAAATGGCCTACTTCCATGAAT
GTCCAGCCGAGTGAGCTTTCAGGAGGCACGCCCTGGCTTGTGAGAGTGAGGGAGGAGTCTCTCTCA
GCCTTGAGAATGAAGCAGAACAGAAGTTAATAGAGAGCATGTTGCAAAACCTGACAAAACCCGGG
ACAGGGATTTCTGATGGTGATTTCTGGATAGGGCTTTGGAGGAATGGAGATGGGCAAACATCTGG
TGCCTGCCCAGATCTCTACAGTGGTCTGATGGAAGCAATTCACAGTACCGAACTGGTACACAG
ATGAACCTTCTGCGGAAGTGAAAAGTGTTGTGATGTATCACCAACCACTGCCAATCCTGGC
CTTGGGGGTCCCTACCTTTACAGTGGAAATGATGACAGGTGTAACATGAAGCACAATTATATTTG
CAAGTATGAACCAGAGATTAATCCAACAGCCCCTGTAGAAAAGCCTTATCTTACAAATCAACCAG
GAGACACCCATCAGAATGTGGTTGTTACTGAAGCAGGTATAATTCCCAATCTAATTTATGTTGTT
ATACCAACAATACCCCTGCTCTTACTGATACTGGTTGCTTTTGGAACTGTTGTTTCCAGATGCT
GCATAAAAGTAAAGGAAGAACAAAAGTCTAGTCCAAACAGTCTACACTGTGGATTTCAAAGAGTA
CCAGAAAAGAAAGTGGCATGGAAGTATAAATACTCATTGACTTGGTCCAGAATTTTGTAATCT
GGATCTGTATAAGGAATGGCATCAGAACAAATAGCTTGGAAATGGCTTGAATCACAAGGATCTGC
AAGATGAACTGTAAGCTCCCCCTTGAGGCAAATATTAAAGTAATTTTATATGTCTATTATTTCA
TTTAAAGAATATGCTGTGCTAATAATGGAGTGAGACATGCTTATTTTGCTAAAGGATGCACCCAA
ACTTCAAACCTCAAGCAAATGAAATGGACAATGCAGATAAAGTTGTTATCAACACGTCGGGAGTA
TGTGTGTTAGAAGCAATTCCTTTTATTTCTTTACCTTTTCATAAGTTGTTATCTAGTCAATGTAA
TGTATATTGTATTGAAATTTACAGTGTGCAAAAGTATTTTACCTTTGCATAAGTGTTTGATAAAA
ATGAACCTGTTCTAATATTTATTTTATGGCATCTCATTTTTCAATACATGCTCTTTTGATTAAAG
AACTTATTACTGTTGTCAACTGAATTCACACACACACAAATATAGTACCATAGAAAAAGTTTGT
TTTCTCGAAATAATTCATCTTTCAGCTTCTCTGCTTTTGGTCAATGTCTAGGAAATCTCTTCAGA
AATAAGAAGCTATTTTATTAAAGTGATATAAACCTCCTCAAACATTTTACTTAGAGGCAAGGAT
TGTCTAATTTCAATTGTGCAAGACATGTGCCTTATAATTATTTTAGCTTAAATTAACAGATT
TTGTAATAATGTAACCTTTGTTAATAGGTGCATAAACACTAATGCAGTCAATTTGAACAAAAGAG
TGACATACACAATATAAATCATATGTCTTCACACGTTGCCTATATAATGAGAAGCAGCTCTCTGA
GGGTTCTGAAATCAATGTGGTCCCTCTCTTGCCCACTAAACAAAGATGGTTGTTTCGGGGTTTGGG
ATTGACACTGGAGGCAGATAGTTGCAAAGTTAGTCTAAGGTTTCCCTAGCTGTATTTAGCCTCTG
ACTATATTAGTATACAAAGAGGTCATGTGGTTGAGACCAGGTGAATAGTCACTATCAGTGTGGAG
ACAAGCACAGCACACAGACATTTTAGGAAGGAAAGGAACACGAAATCGTGTGAAAATGGGTTGG
AACCCTCAGTGATCGCATATTCATTGATGAGGGTTTGTCTGAGATAGAAAATGGTGGCTCCTTT
CTGTCTTATCTCCTAGTTTCTTCAATGCTTACGCCTTGTCTTCTCAAGAGAAAGTTGTAACCT
CTGGTCTTCATATGTCCCTGTGCTCCTTTTAAACCAATAAAGAGTTCTTGTCTTCTGGGGGAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 168

MSRVVSLLLGAALLCGHGAFRRVVGQKVCFADFKHPCYKMAYFHELSSRVSFQEARLACESE
GGVLLSLENEAEQKLIESMLQNLTKPGTGISDGDFWIGLWRNGDGQTSACPDLYQWSDGSNSQ
YRNWYTDEPSCGSEKCVVMYHQPTANPGLGGPYLYQWNDDRCNMKKNYICKYEPEINPTAPVEK
PYLTNQPGDTHQNVVVTEAGIIPNLIYVVIPTIPLLLLILVAFGTCCFQMLHKSNGRTKTSPNQ
STLWISKSTRKESGMEV

Important features of the protein:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 214-235

N-glycosylation sites.

amino acids 86-89, 255-258

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 266-269

N-myristoylation sites.

amino acids 27-32, 66-71, 91-96, 93-98, 102-107, 109-114, 140-
145, 212-217

**SECRETED AND TRANSMEMBRANE
POLYPEPTIDES AND NUCLEIC ACIDS
ENCODING THE SAME**

**CROSS REFERENCE TO RELATED
APPLICATIONS**

[0001] This is a continuation application claiming priority under 35 USC § 120 to U.S. Ser. No. 10/006867 Filed Dec. 6, 2001, and which claims priority under 35 USC § 119 to U.S. provisional serial Nos. 60/063,435 Filed Oct. 29, 1997; 60/064,215 Filed Oct. 29, 1997; 60/082,797 Filed Apr. 22, 1998; 60/083,495 Filed Apr. 29, 1998; 60/085,579 Filed May 15, 1998; 60/087,759 Filed Jun. 2, 1998; 60/088,021 Filed Jun. 4, 1998; 60/088,029 Filed Jun. 4, 1998; 60/088,030 Filed Jun. 4, 1998; 60/088,734 Filed Jun. 10, 1998; 60/088,740 Filed Jun. 10, 1998; 60/088,811 Filed Jun. 10, 1998; 60/088,824 Filed Jun. 10, 1998; 60/088,825 Filed Jun. 10, 1998; 60/088,863 Filed Jun. 11, 1998; 60/089,105 Filed Jun. 12, 1998; 60/089,514 Filed Jun. 16, 1998; 60/089,653 Filed Jun. 17, 1998; 60/089,952 Filed Jun. 19, 1998; 60/090,246 Filed Jun. 22, 1998; 60/090,444 Filed Jun. 24, 1998; 60/090,688 Filed Jun. 25, 1998; 60/090,696 Filed Jun. 25, 1998; 60/090,862 Filed Jun. 26, 1998; 60/091,628 Filed Jul. 2, 1998; 60/096,012 Filed Aug. 10, 1998; 60/096,757 Filed Aug. 17, 1998; 60/096,949 Filed Aug. 18, 1998; 60/096,959 Filed Aug. 18, 1998; 60/097,954 Filed Aug. 26, 1998; 60/097,971 Filed Aug. 26, 1998; 60/097,979 Filed Aug. 26, 1998; 60/098,749 Filed Sep. 1, 1998; 60/099,741 Filed Sep. 10, 1998; 60/099,763 Filed Sep. 10, 1998; 60/099,792 Filed Sep. 10, 1998; 60/099,812 Filed Sep. 10, 1998; 60/099,815 Filed Sep. 10, 1998; 60/100,627 Filed Sep. 16, 1998; 60/100,662 Filed Sep. 16, 1998; 60/100,683 Filed Sep. 17, 1998; 60/100,684 Filed Sep. 17, 1998; 60/100,930 Filed Sep. 17, 1998; 60/101,279 Filed Sep. 22, 1998; 60/101,475 Filed Sep. 23, 1998; 60/101,738 Filed Sep. 24, 1998; 60/101,743 Filed Sep. 24, 1998; 60/101,916 Filed Sep. 24, 1998; 60/102,570 Filed Sep. 30, 1998; 60/103,449 Filed Oct. 6, 1998; 60/103,678 Filed Oct. 8, 1998; 60/103,679 Filed Oct. 8, 1998; 60/103,711 Filed Oct. 8, 1998; 60/105,000 Filed Oct. 20, 1998; 60/105,002 Filed Oct. 20, 1998; 60/105,881 Filed Oct. 27, 1998; 60/106,030 Filed Oct. 28, 1998; 60/106,464 Filed Oct. 30, 1998; 60/106,856 Filed Nov. 3, 1998; 60/108,807 Filed Nov. 17, 1998; 60/112,419 Filed Dec. 15, 1998; 60/112,422 Filed Dec. 15, 1998; 60/112,853 Filed Dec. 16, 1998; 60/113,011 Filed Dec. 16, 1998; 60/112,854 Filed Dec. 16, 1998; 60/113,300 Filed Dec. 22, 1998; 60/113,408 Filed Dec. 22, 1998; 60/113,430 Filed Dec. 23, 1998; 60/113,621 Filed Dec. 23, 1998; 60/114,223 Filed Dec. 30, 1998; 60/115,614 Filed Jan. 12, 1999; 60/116,527 Filed Jan. 20, 1999; 60/116,843 Filed Jan. 22, 1999; 60/119,285 Filed Feb. 9, 1999; 60/119,287 Filed Feb. 9, 1999; 60/119,525 Filed Feb. 10, 1999; 60/119,549 Filed Feb. 10, 1999; 60/120,014 Filed Feb. 11, 1999; 60/129,122 Filed Apr. 13, 1999; 60/129,674 Filed Apr. 16, 1999; 60/131,291 Filed Apr. 27, 1999; 60/138,387 Filed Jun. 9, 1999; 60/144,791 Filed Jul. 20, 1999; 60/169,495 Filed Dec. 7, 1999; 60/175,481 Filed Jan. 11, 2000; 60/191,007 Filed Mar. 21, 2000; 60/199,397 Filed Apr. 25, 2000 and which claims priority under 35 USC § 120 to U.S. patent applications and PCT International Patent Application Serial Numbers 09/380,139 Filed Aug. 25, 1999; 09/311,832 Filed May 14, 1999; 09/380,137 Filed Aug. 25, 1999, now abandoned; 09/380,138 Filed Aug. 25, 1999, now abandoned; 09/380,142 Filed Aug. 25, 1999, now abandoned; 09/397,342 Filed

Sep. 15, 1999; 09/403,297 Filed Oct. 18, 1999, now abandoned; 09/423,844 Filed Nov. 12, 1999, now abandoned; 09/644,848 Filed Aug. 22, 2000; 09/665,350 Filed Sep. 18, 2000; 09/664,610 Filed Sep. 18, 2000, now abandoned; 09/709,238 Filed Nov. 8, 2000; 09/747,259 Filed Dec. 20, 2000; 09/816,744 Filed Mar. 22, 2001; 09/854,208 Filed May 10, 2001; 09/854,280 Filed May 10, 2001; 09/870,574 Filed May 30, 2001; 09/874,503 Filed Jun. 5, 2001; 09/908,827 Filed Jul. 18, 2001; 09/869,566 Filed Feb. 19, 2002; PCT/US98/19330 Filed Sep. 16, 1998; PCT/US99/05028 Filed Mar. 8, 1999; PCT/US99/10733 Filed May 14, 1999; PCT/US99/12252 Filed Jun. 2, 1999; PCT/US99/20111 Filed Sep. 1, 1999; PCT/US99/21090 Filed Sep. 15, 1999; PCT/US99/21194 Filed Sep. 15, 1999; PCT/US99/30720 Filed Dec. 22, 1999; PCT/US00/04341 Filed Feb. 18, 2000; PCT/US00/04342 Filed Feb. 18, 2000; PCT/US00/04414 Filed Feb. 22, 2000; PCT/US00/05601 Filed Mar. 1, 2000; PCT/US00/08439 Filed Mar. 30, 2000; PCT/US00/14042 Filed May 22, 2000; PCT/US00/15264 Filed Jun. 2, 2000; PCT/US00/23522 Filed Aug. 23, 2000; PCT/US00/23328 Filed Aug. 24, 2000; PCT/US00/30873 Filed Nov. 10, 2000; PCT/US00/32678 Filed Dec. 1, 2000; PCT/US00/34956 Filed Dec. 20, 2000; PCT/US01/06520 Filed Feb. 28, 2001; PCT/US01/06666 Filed Mar. 1, 2001; PCT/US01/17443 Filed May 30, 2001; PCT/US01/17800 Filed Jun. 1, 2001; PCT/US01/19692 Filed Jun. 20, 2001; PCT/US01/21066 Filed Jun. 29, 2001; PCT/US01/21735 Filed Jul. 9, 2001, the entire disclosures of which are hereby incorporated by reference.

BACKGROUND OF INVENTION

[0002] The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

[0003] Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of action in the extracellular environment.

[0004] Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., Proc. Natl. Acad. Sci. 93:7108-7113 (1996); U.S. Pat. No. 5,536,637].

[0005] Membrane-bound proteins and receptors can play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and nerve growth factor receptor.

[0006] Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

[0007] Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

SUMMARY OF INVENTION

[0008] In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

[0009] In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic

acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0010] In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0011] In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively

at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

[0012] Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

[0013] Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule frag-

ments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

[0014] In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

[0015] In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

[0016] In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

[0017] In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

[0018] Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

[0019] In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

[0020] In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

[0021] In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

[0022] Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

[0023] In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

[0024] In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

[0025] In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the

above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

[0026] In yet other embodiments, the invention provides oligonucleotide probes useful for isolating genomic and cDNA nucleotide sequences or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences.

[0027] In yet other embodiments, the present invention is directed to methods of using the PRO polypeptides of the present invention for a variety of uses based upon the functional biological assay data presented in the Examples below.

BRIEF DESCRIPTION OF DRAWINGS

[0028] FIG. 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO180 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA26843-1389".

[0029] FIG. 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in FIG. 1.

[0030] FIG. 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO218 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA30867-1335".

[0031] FIG. 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in FIG. 3.

[0032] FIG. 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO263 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA34431-1177".

[0033] FIG. 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in FIG. 5.

[0034] FIG. 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO295 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA38268-1188".

[0035] FIG. 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in FIG. 7.

[0036] FIG. 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO874 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA40621-1440".

[0037] FIG. 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in FIG. 9.

[0038] FIG. 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO300 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA40625-1189".

[0039] FIG. 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in FIG. 11.

[0040] FIG. 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA45409-2511".

[0041] FIG. 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in FIG. 13.

[0042] FIG. 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO1282 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA45495-1550".

[0043] FIG. 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in FIG. 15.

[0044] FIG. 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO1063 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA49820-1427".

[0045] FIG. 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in FIG. 17.

[0046] FIG. 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO1773 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA56406-1704".

[0047] FIG. 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in FIG. 19.

[0048] FIG. 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO1013 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA56410-1414".

[0049] FIG. 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in FIG. 21.

[0050] FIG. 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO937 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA56436-1448".

[0051] FIG. 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in FIG. 23.

[0052] FIG. 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO842 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA56855-1447".

[0053] FIG. 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in FIG. 25.

[0054] FIG. 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO1180 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA56860-1510".

[0055] FIG. 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in FIG. 27.

[0056] FIG. 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO831 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA56862-1343".

[0057] FIG. 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in FIG. 29.

[0058] FIG. 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO1115 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA56868-1478".

[0059] FIG. 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in FIG. 31.

[0060] FIG. 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO1277 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA56869-1545".

[0061] FIG. 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in FIG. 33.

[0062] FIG. 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO1074 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA57704-1452".

[0063] FIG. 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in FIG. 35.

[0064] FIG. 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO1344 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA58723-1588".

[0065] FIG. 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in FIG. 37.

[0066] FIG. 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO1136 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA57827-1493".

[0067] FIG. 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in FIG. 39.

[0068] FIG. 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO1109 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA58737-1473".

[0069] FIG. 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in FIG. 41.

[0070] FIG. 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO1003 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA58846-1409".

[0071] FIG. 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in FIG. 43.

[0072] FIG. 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO1138 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA58850-1495".

[0073] FIG. 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in FIG. 45.

[0074] FIG. 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO994 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA58855-1422".

[0075] FIG. 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in FIG. 47.

[0076] FIG. 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO1069 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA59211-1450".

[0077] FIG. 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in FIG. 49.

[0078] FIG. 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA59212-1627".

[0079] FIG. 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in FIG. 51.

[0080] FIG. 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO1129 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA59213-1487".

[0081] FIG. 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in FIG. 53.

[0082] FIG. 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO1027 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA59605-1418".

[0083] FIG. 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in FIG. 55.

[0084] FIG. 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO1106 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA59609-1470".

[0085] FIG. 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in FIG. 57.

[0086] FIG. 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO1291 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA59610-1556".

[0087] FIG. 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in FIG. 59.

[0088] FIG. 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO3573 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA59837-2545".

[0089] FIG. 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in FIG. 61.

[0090] FIG. 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO3566 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA59844-2542".

[0091] FIG. 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in FIG. 63.

[0092] FIG. 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO1098 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA59854-1459".

[0093] FIG. 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in FIG. 65.

[0094] FIG. 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO1158 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA60625-1507".

[0095] FIG. 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in FIG. 67.

[0096] FIG. 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO1124 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA60629-1481".

[0097] FIG. 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in FIG. 69.

[0098] FIG. 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO1287 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA61755-1554".

[0099] FIG. 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in FIG. 71.

[0100] FIG. 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO1335 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA62812-1594".

[0101] FIG. 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in FIG. 73.

[0102] FIG. 75 shows a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO1315 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA62815-1576".

[0103] FIG. 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in FIG. 75.

[0104] FIG. 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO1357 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA64881-1602".

[0105] FIG. 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in FIG. 77.

[0106] FIG. 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO1356 cDNA, wherein SEQ ID NO:79 is a clone designated herein as "DNA64886-1601".

[0107] FIG. 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in FIG. 79.

[0108] FIG. 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO1557 cDNA, wherein SEQ ID NO:81 is a clone designated herein as "DNA64902-1667".

[0109] FIG. 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in FIG. 81.

[0110] FIG. 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO1347 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA64950-1590".

[0111] FIG. 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in FIG. 83.

[0112] FIG. 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO1302 cDNA, wherein SEQ ID NO:85 is a clone designated herein as "DNA65403-1565".

[0113] FIG. 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in FIG. 85.

[0114] FIG. 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO1270 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA66308-1537".

[0115] FIG. 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in FIG. 87.

[0116] FIG. 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO1268 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA66519-1535".

[0117] FIG. 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in FIG. 89.

[0118] FIG. 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO1327 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA66521-1583".

[0119] FIG. 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in FIG. 91.

[0120] FIG. 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO1328 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA66658-1584".

[0121] FIG. 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in FIG. 93.

[0122] FIG. 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA66660-1585".

[0123] FIG. 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in FIG. 95.

[0124] FIG. 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO1340 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA66663-1598".

[0125] FIG. 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in FIG. 97.

[0126] FIG. 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO1342 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA66674-1599".

[0127] FIG. 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in FIG. 99.

[0128] FIG. 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO3579 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA68862-2546".

[0129] FIG. 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in FIG. 101.

[0130] FIG. 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO1472 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA68866-1644".

[0131] FIG. 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in FIG. 103.

[0132] FIG. 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO1461 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA68871-1638".

[0133] FIG. 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in FIG. 105.

[0134] FIG. 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO1568 cDNA, wherein SEQ ID NO:107 is a clone designated herein as "DNA68880-1676".

[0135] FIG. 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:107 shown in FIG. 107.

[0136] FIG. 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO1753 cDNA, wherein SEQ ID NO:109 is a clone designated herein as "DNA68883-1691".

[0137] FIG. 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in FIG. 109.

[0138] FIG. 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO1570 cDNA, wherein SEQ ID NO:111 is a clone designated herein as "DNA68885-1678".

[0139] FIG. 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in FIG. 111.

[0140] FIG. 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO1446 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA71277-1636".

[0141] FIG. 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in FIG. 113.

[0142] FIG. 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO1565 cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA73727-1673".

[0143] FIG. 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in FIG. 115.

[0144] FIG. 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO1572 cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA73734-1680".

[0145] FIG. 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in FIG. 117.

[0146] FIG. 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO1573 cDNA, wherein SEQ ID NO:119 is a clone designated herein as "DNA73735-1681".

[0147] FIG. 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in FIG. 119.

[0148] FIG. 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO1550 cDNA, wherein SEQ ID NO:121 is a clone designated herein as "DNA76393-1664".

[0149] FIG. 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in FIG. 121.

[0150] FIG. 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO1693 cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA77301-1708".

[0151] FIG. 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in FIG. 123.

[0152] FIG. 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO1566 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA77568-1626".

[0153] FIG. 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in FIG. 125.

[0154] FIG. 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO1774 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA77626-1705".

[0155] FIG. 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in FIG. 127.

[0156] FIG. 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA81754-2532".

[0157] FIG. 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in FIG. 129.

[0158] FIG. 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO1865 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA81757-2512".

[0159] FIG. 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in FIG. 131.

[0160] FIG. 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO1925 cDNA, wherein SEQ ID NO:133 is a clone designated herein as "DNA82302-2529".

[0161] FIG. 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in FIG. 133.

[0162] FIG. 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO1926 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA82340-2530".

[0163] FIG. 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in FIG. 135.

[0164] FIG. 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA83500-2506".

[0165] FIG. 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in FIG. 137.

[0166] FIG. 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO4405 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA84920-2614".

[0167] FIG. 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in FIG. 139.

[0168] FIG. 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO3435 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA85066-2534".

[0169] FIG. 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in FIG. 141.

[0170] FIG. 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO3543 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA86571-2551".

[0171] FIG. 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in FIG. 143.

[0172] FIG. 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO3443 cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA87991-2540".

[0173] FIG. 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in FIG. 145.

[0174] FIG. 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO3442 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA92238-2539".

[0175] FIG. 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in FIG. 147.

[0176] FIG. 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO5990 cDNA, wherein SEQ ID NO:149 is a clone designated herein as "DNA96042-2682".

[0177] FIG. 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in FIG. 149.

[0178] FIG. 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO4342 cDNA, wherein SEQ ID NO:151 is a clone designated herein as "DNA96787-2534".

[0179] FIG. 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in FIG. 151.

[0180] FIG. 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO10096 cDNA, wherein SEQ ID NO:153 is a clone designated herein as "DNA125185-2806".

[0181] FIG. 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in FIG. 153.

[0182] FIG. 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO10272 cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA147531-2821".

[0183] FIG. 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in FIG. 155.

[0184] FIG. 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO5801 cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA115291-2681".

[0185] FIG. 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in FIG. 157.

[0186] FIG. 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO20110 cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA166819".

[0187] FIG. 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in FIG. 159.

[0188] FIG. 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO20040 cDNA, wherein SEQ ID NO:161 is a clone designated herein as "DNA164625-2890".

[0189] FIG. 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in FIG. 161.

[0190] FIG. 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO20233 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA165608".

[0191] FIG. 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in FIG. 163.

[0192] FIG. 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO19670 cDNA, wherein SEQ ID NO:165 is a clone designated herein as "DNA131639-2874".

[0193] FIG. 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in FIG. 165.

[0194] FIG. 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1890 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA79230-2525".

[0195] FIG. 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in FIG. 167.

DETAILED DESCRIPTION

I. Definitions

[0196] The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the prepa-

ration of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

[0197] A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (e.g., an extracellular domain sequence), naturally-occurring variant forms (e.g., alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

[0198] The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

[0199] The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (e.g., Nielsen et al., *Prot. Eng.* 10:1-6 (1997) and von Heinje et al., *Nucl. Acids. Res.* 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal

sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

[0200] "PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

[0201] "Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a

candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, Calif. or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0202] In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction $\frac{X}{Y}$

[0203] where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X", "Y" and "Z" each represent different hypothetical amino acid residues.

[0204] Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the

ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=11, and scoring matrix=BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

[0205] Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, Md. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask=yes, strand=all, expected occurrences=10, minimum low complexity length=15/5, multi-pass e-value=0.01, constant for multi-pass=25, dropoff for final gapped alignment=25 and scoring matrix=BLOSUM62.

[0206] In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

100 times the fraction $\frac{X}{Y}$

[0207] where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

[0208] "PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length

native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.

[0209] Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

[0210] "Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein

the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, Calif. or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

[0211] In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction $\frac{W}{Z}$

[0212] where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent different hypothetical nucleotides.

[0213] Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span=1, overlap fraction=0.125, word threshold (T)=11, and scoring matrix=BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by

WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement “an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B”, the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

[0214] Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., *Nucleic Acids Res.* 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, Md. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask=yes, strand=all, expected occurrences=10, minimum low complexity length=15/5, multi-pass e-value=0.01, constant for multi-pass=25, dropoff for final gapped alignment=25 and scoring matrix=BLOSUM62.

[0215] In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction W/Z

[0216] where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

[0217] In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions, to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

[0218] “Isolated,” when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes

polypeptide in situ within recombinant cells, since at least one component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

[0219] An “isolated” PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of natural cells.

[0220] The term “control sequences” refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

[0221] Nucleic acid is “operably linked” when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, “operably linked” means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

[0222] The term “antibody” is used in the broadest sense and specifically covers, for example, single anti-PRO monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polypeptidic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term “monoclonal antibody” as used herein refers to an antibody obtained from a population of substantially homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

[0223] “Stringency” of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands

are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., *Current Protocols in Molecular Biology*, Wiley Interscience Publishers, (1995).

[0224] “Stringent conditions” or “high stringency conditions”, as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50° C.; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50 mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42° C.; or (3) employ 50% formamide, 5×SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5×Denhardt’s solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42° C., with washes at 42° C. in 0.2×SSC (sodium chloride/sodium citrate) and 50% formamide at 55° C., followed by a high-stringency wash consisting of 0.1×SSC containing EDTA at 55° C.

[0225] “Moderately stringent conditions” may be identified as described by Sambrook et al., *Molecular Cloning: A Laboratory Manual*, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and % SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37° C. in a solution comprising: 20% formamide, 5×SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5×Denhardt’s solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1×SSC at about 37-50° C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

[0226] The term “epitope tagged” when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a “tag polypeptide”. The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

[0227] As used herein, the term “immunoadhesin” designates antibody-like molecules which combine the binding specificity of a heterologous protein (an “adhesin”) with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding specificity which is other than the antigen recognition and binding site

of an antibody (i.e., is “heterologous”), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

[0228] “Active” or “activity” for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein “biological” activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an “immunological” activity refers to the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

[0229] The term “antagonist” is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term “agonist” is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

[0230] “Treatment” refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

[0231] “Chronic” administration refers to administration of the agent(s) in a continuous mode as opposed to an acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. “Intermittent” administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

[0232] “Mammal” for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

[0233] Administration “in combination with” one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

[0234] “Carriers” as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are non-toxic to the cell or mammal being exposed thereto at the

dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrins; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEENTM, polyethylene glycol (PEG), and PLURONICSTM.

[0235] "Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')₂, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062[1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

[0236] Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')₂ fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

[0237] "Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V_H-V_L dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

[0238] The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')₂ antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

[0239] The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

[0240] Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM,

and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

[0241] "Single-chain Fv" or "sFv" antibody fragments comprise the V_H and V_L domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V_H and V_L domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

[0242] The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) in the same polypeptide chain (V_H-V_L). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

[0243] An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

[0244] An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

[0245] The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical alteration of a substrate compound or composition which is detectable.

[0246] By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can

comprise the well of an assay plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Pat. No. 4,275,149.

[0247] A “liposome” is a small vesicle composed of various types of lipids, phospholipids and/or surfactant

which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of biological membranes.

[0248] A “small molecule” is defined herein to have a molecular weight below about 500 Daltons

TABLE 1

```

/*
 *
 * C—C increased from 12 to 15
 * Z is average of EQ
 * B is average of ND
 * match with stop is _M; stop—stop = 0; J (joker) match = 0
 */
#define _M -8 /* value of a match with a stop */
int _day[26][26] = {
/* A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ {2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ {0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
/* D */ {0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ {0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ {1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
/* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
/* N */ {0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ {_M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ {1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ {0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
/* S */ {1, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ {1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ {0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
/* X */ {0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ {0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, -2, -6, 0, -4, 4},
};
/*
 */
#include <stdio.h>
#include <ctype.h>
#define MAXJMP 16 /* max jumps in a diag */
#define MAXGAP 24 /* don't continue to penalize gaps larger than this */
#define JMPS 1024 /* max jmps in an path */
#define MX 4 /* save if there's at least MX-1 bases since last jmp */
#define DMAT 3 /* value of matching bases */
#define DMIS 0 /* penalty for mismatched bases */
#define DINS0 8 /* penalty for a gap */
#define DINS1 1 /* penalty per base */
#define PINS0 8 /* penalty for a gap */
#define PINS1 4 /* penalty per residue */
struct jmp {
    short n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short x[MAXJMP]; /* base no. of jmp in seq x */
    /* limits seq to 2 16 -1 */
};
struct diag {
    int score; /* score at last jmp */
    long offset; /* offset of prev block */
    short ijmp; /* current jmp index */
    struct jmp jp; /* list of jmps */
};
struct path {
    int spc; /* number of leading spaces */
    short n[JMPS]; /* size of jmp (gap) */
    int x[JMPS]; /* loc of jmp (last elem before gap) */
};

```

TABLE 1-continued

```

};
char      *ofile;           /* output file name */
char      *namex[2];        /* seq names: getseqs() */
char      *prog;            /* prog name for err msgs */
char      *seqx[2];         /* seqs: getseqs() */
int       dmax;             /* best diag: nw() */
int       dmax0;            /* final diag */
int       dna;              /* set if dna: main() */
int       endgaps;          /* set if penalizing end gaps */
int       gapx, gapy;        /* total gaps in seqs */
int       len0, len1;        /* seq lens */
int       ngapx, ngapy;      /* total size of gaps */
int       smax;             /* max score: nw() */
int       *xbm;             /* bitmap for matching */
long      offset;           /* current offset in jmp file */
struct    diag *dx;         /* holds diagonals */
struct    path pp[2];       /* holds path for seqs */
char      *calloc(), *malloc(), *index(), *strcpy();
char      *getseq(), *g_calloc();
/* Needleman-Wunsch alignment program
 *
 * usage: progs file1 file2
 * where file1 and file2 are two dna or two protein sequences.
 * The sequences can be in upper- or lower-case and may contain ambiguity
 * Any lines beginning with ';', '>' or '<' are ignored
 * Max file length is 65535 (limited by unsigned short x in the jmp struct)
 * A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
 * Output is in the file "align.out"
 *
 * The program may create a tmp file in /tmp to hold info about traceback.
 * Original version developed under BSD 4.3 on a vax 8650
 */
#include "nw.h"
#include "day.h"
static __dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
};
static __pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
    1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};
main(ac, av)
int      ac;
char     *av[];
{
    prog = av[0];
    if(ac != 3) {
        fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';', '>' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
    }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
    xbm = (dna)? __dbval : __pbval;
    endgaps = 0;           /* 1 to penalize endgaps */
    ofile = "align.out";   /* output file */
    nw();                 /* fill in the matrix, get the possible jmps */
    readjumps();          /* get the actual jmps */
    print();              /* print stats, alignment */
    cleanup(0);           /* unlink any tmp files */
}
/* do the alignment, return best score: main()
 * dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
 * pro: PAM 250 values
 * When scores are equal, we prefer mismatches to any gap, prefer
 * a new gap to extending an ongoing gap, and prefer a gap in seqx
 * to a gap in seq y.
 */
nw()

```

main

nw

TABLE 1-continued

```

{
    char    *px, *py;        /* seqs and ptrs */
    int     *ndely, *dely;    /* keep track of dely */
    int     ndelx, delx;      /* keep track of delx */
    int     *tmp;            /* for swapping row0, row1 */
    int     mis;             /* score for each type */
    int     ins0, ins1;       /* insertion penalties */
    register id;             /* diagonal index */
    register ij;            /* jmp index */
    register *col0, *col1;    /* score for curr, last row */
    register xx, yy;         /* index into seqs */
    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));
    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
    col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;
    smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
        }
        col0[0] = 0;        /* Waterman Bull Math Biol 84 */
    }
    else
        for (yy = 1; yy <= len1; yy++)
            dely[yy] = -ins0;
    /* fill in match matrix
    */
    for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
            else
                col1[0] = delx = col0[0]-ins1;
            ndelx = xx;
        }
        else {
            col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
    }

    for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
            mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
            mis += _day[*px-'A'][*py-'A'];
        /* update penalty for del in x seq;
        * favor new del over ongong del
        * ignore MAXGAP if weighting endgaps
        */
        if (endgaps || ndely[yy] < MAXGAP) {
            if (col0[yy] - ins0 >= dely[yy]) {
                dely[yy] = col0[yy] - (ins0+ins1);
                ndely[yy] = 1;
            } else {
                dely[yy] -= ins1;
                ndely[yy]++;
            }
        } else {
            if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                dely[yy] = col0[yy] - (ins0+ins1);
                ndely[yy] = 1;
            } else
                ndely[yy]++;
        }
        /* update penalty for del in y seq;
        * favor new del over ongong del
        */
        if (endgaps || ndelx < MAXGAP) {

```

TABLE 1-continued

```

        if(col1[yy-1] - ins0 >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else {
            delx -= ins1;
            ndelx++;
        }
    } else {
        if (col1[yy-1] - (ins0+ins1) >= delx) {
            delx = col1[yy-1] - (ins0+ins1);
            ndelx = 1;
        } else
            ndelx++;
    }
    /* pick the maximum score; we're favoring
    * mis over any del and delx over dely
    */

    id = xx - yy + len1 - 1;
    if (mis >= delx && mis >= dely[yy])
        col1[yy] = mis;
    else if (delx >= dely[yy]) {
        col1[yy] = delx;
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writeimps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
        }
        dx[id].jp.n[ij] = ndelx;
        dx[id].jp.x[ij] = xx;
        dx[id].score = delx;
    }
    else {
        col1[yy] = dely[yy];
        ij = dx[id].ijmp;
        if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
        && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
            dx[id].ijmp++;
            if (++ij >= MAXJMP) {
                writeimps(id);
                ij = dx[id].ijmp = 0;
                dx[id].offset = offset;
                offset += sizeof(struct jmp) + sizeof(offset);
            }
        }
        dx[id].jp.n[ij] = ndely[yy];
        dx[id].jp.x[ij] = xx;
        dx[id].score = dely[yy];
    }
    if (xx == len0 && yy < len1) {
        /* last col
        */
        if (endgaps)
            col1[yy] -= ins0+ins1*(len1-yy);
        if (col1[yy] > smax) {
            smax = col1[yy];
            dmax = id;
        }
    }
}
if (endgaps && xx < len0)
    col1[yy-1] -= ins0+ins1*(len0-xx);
if (col1[yy-1] > smax) {
    smax = col1[yy-1];
    dmax = id;
}
tmp = col0; col0 = col1; col1 = tmp;
}
(void) free((char *)ndely);
(void) free((char *)dely);

```

...nw

TABLE 1-continued

```

(void) free((char *)col0);
(void) free((char *)col1);
    }
/*
 *
 * print() -- only routine visible outside this module
 *
 * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */
#include "nw.h"
#define SPC 3
#define P_LINE 256 /* maximum output line */
#define P_SPC 3 /* space between name or num and seq */
extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */
print()
{
    int lx, ly, firstgap, lastgap; /* overlap */
    if ((fx = fopen(ofile, "w")) == 0) {
        fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "<first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "<second sequence: %s (length = %d)\n", namex[1], len1);
    olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
        pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
        pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
        lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
        lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}
/*
 * trace back the best path, count matches
 */
static
getmat(lx, ly, firstgap, lastgap)
{
    int lx, ly; /* "core" (minus endgaps) */
    int firstgap, lastgap; /* leading trailing overlap */
    {
        int nm, i0, i1, siz0, siz1;
        char outx[32];
        double pct;
        register n0, n1;
        register char *p0, *p1;
        /* get total matches, score
         */
        i0 = i1 = siz0 = siz1 = 0;
        p0 = seqx[0] + pp[1].spc;
        p1 = seqx[1] + pp[0].spc;
        n0 = pp[1].spc + 1;
        n1 = pp[0].spc + 1;
        nm = 0;

```

TABLE 1-continued

```

while ( *p0 && *p1 ) {
    if (siz0) {
        p1++;
        n1++;
        siz0--;
    }
    else if (siz1) {
        p0++;
        n0++;
        siz1--;
    }
    else {
        if (xbm[*p0-'A']&xbm[*p1-'A'])
            nm++;
        if (n0++ == pp[0].x[i0])
            siz0 = pp[0].n[i0++];
        if (n1++ == pp[1].x[i1])
            siz1 = pp[1].n[i1++];
        p0++;
        p1++;
    }
}
/* pct homology:
 * if penalizing endgaps, base is the shorter seq
 * else, knock off overhangs and take shorter core
 */
if (endgaps)
    lx = (len0 < len1)? len0 : len1;
else
    lx = (lx < ly)? lx : ly;
pct = 100.*(double)nm/(double)lx;
fprintf(fx, "\n");
fprintf(fx, "<end match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
fprintf(fx, "<gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, "(%d %s%s)",
        ngapx, (dna)? "base": "residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);
}
fprintf(fx, ", gaps in second sequence: %d", gapy);
if (gapy) {
    (void) sprintf(outx, "(%d %s%s)",
        ngapy, (dna)? "base": "residue", (ngapy == 1)? "" : "s");
    fprintf(fx, "%s", outx);
}
if (dna)
    fprintf(fx,
        "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
        smax, DMAT, DMIS, DINS0, DINS1);
else
    fprintf(fx,
        "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
        smax, PINS0, PINS1);
if (endgaps)
    fprintf(fx,
        "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
        firstgap, (dna)? "base": "residue", (firstgap == 1)? "" : "s",
        lastgap, (dna)? "base": "residue", (lastgap == 1)? "" : "s");
else
    fprintf(fx, "<endgaps not penalized\n");
}
static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char *ps[2];      /* ptr to current element */
static char *po[2];      /* ptr to next output char slot */
static char out[2][P_LINE]; /* output line */
static char star[P_LINE]; /* set by stars() */
/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()

```

...getmat

pr_align

TABLE 1-continued

```

{
    int      nn;          /* char count */
    int      more;
    register i;
    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(name[x[i]]);
        if (nn > lmax)
            lmax = nn;
        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
    for (nn = nm = 0, more = 1; more;) {
        for (i = more = 0; i < 2; i++) {
            /*
             * do we have more of this sequence?
             */
            if (!*ps[i])
                continue;
            more++;
            if (pp[i].spc) {
                /* leading space */
                *po[i]++ = ' ';
                pp[i].spc--;
            }
            else if (siz[i]) {
                /* in a gap */
                *po[i]++ = '-';
                siz[i]--;
            }
            else {
                /* we're putting a seq element
                 */
                *po[i] = *ps[i];
                if (islower(*ps[i]))
                    *ps[i] = toupper(*ps[i]);
                po[i]++;
                ps[i]++;
            }
            /*
             * are we at next gap for this seq?
             */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                 * we need to merge all gaps
                 * at this location
                 */
                siz[i] == pp[i].n[ij[i]++];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i]++];
            }
            ni[i]++;
        }
        if (++nn == olen || !more && nn) {
            dumpblock();
            for (i = 0; i < 2; i++)
                po[i] = out[i];
            nn = 0;
        }
    }
}
/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
static
dumpblock()
{
    register i;
    for (i = 0; i < 2; i++)
        *po[i]-- = '\0';

    (void) puts("\n", fx);
    for (i = 0; i < 2; i++) {
        if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
            if (i == 0)
                nums(i);
            if (i == 0 && *out[1])

```

TABLE 1-continued

<pre> stars(); putline(i); if (i == 0 && *out[1]) fprintf(fx, star); if (i == 1) nums(i); } } } /* * put out a number line: dumpblock() */ static nums(ix) int ix; /* index in out[] holding seq line */ { char nline[P_LINE]; register i, j; register char *pn, *px, *py; for(pn = nline, i = 0; i < lmax+P_SPC; i++, pn++) *pn = ' '; for (i = nc[ix], py = out[ix]; *py; py++, pn++) { if (*py == ' ' *py == '-') *pn = ' '; else { if (i%10 == 0 (i == 1 && nc[ix] != 1)) { j = (i < 0)? -i : i; for (px = pn; j; j/= 10, px--) *px = j%10 + '0'; if (i < 0) *px = '-'; } else *pn = ' '; i++; } } *pn = '\0'; nc[ix] = i; for (pn = nline; *pn; pn++) (void) putc(*pn, fx); (void) putc('\n', fx); } /* * put out a line (name, [num], seq. [num]): dumpblock() */ static putline(ix) int ix; { int i; register char *px; for (px = namex[ix], i = 0; *px && *px != '.'; px++, i++) (void) putc(*px, fx); for (i < lmax+P_SPC; i++) (void) putc(' ', fx); /* these count from 1: * ni[] is current element (from 1) * nc[] is number at start of current line */ for (px = out[ix]; *px; px++) (void) putc(*px&0x7F, fx); (void) putc('\n', fx); } /* * put a line of stars (seqs always in out[0], out[1]): dumpblock() */ static stars() { int i; register char *p0, *p1, cx, *px; if (!(*out[0] (*out[0] == ' ' && *(p0[0]) == ' ') !*out[1] (*out[1] == ' ' && *(p0[1]) == ' ')) return; px = star;</pre>		nums
		putline
		...putline
		stars

TABLE 1-continued

<pre>for (i = lmax+P__SPC; i; i--) *px++ = ' '; for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) { if (isalpha(*p0) && isalpha(*p1)) { if (xbm[*p0-'A']&xbm[*p1-'A']) { cx = '*'; nm++; } else if (!dna && __day[*p0-'A'][*p1-'A'] > 0) cx = '.'; else cx = ' '; } else cx = ' '; *px++ = cx; } *px++ = '\n'; *px = '\0'; } /* * strip path or prefix from pn, return len: pr__align() */ static stripname(pn) char *pn; /* file name (may be path) */ { register char *px, *py; py = 0; for (px = pn; *px; px++) if (*px == '/') py = px + 1; if (py) (void) strcpy(pn, py); return(strlen(pn)); } /* * cleanup() -- cleanup any tmp file * getseq() -- read in seq, set dna, len, maxlen * g__calloc() -- calloc() with error checkin * readjumps() -- get the good jumps, from tmp file if necessary * writejumps() -- write a filled array of jumps to a tmp file: nw() */ #include "nw.h" #include <sys/file.h> char *jname = "/tmp/homgXXXXXX"; /* tmp file for jumps */ FILE *fj; int cleanup(); /* cleanup tmp file */ long lseek(); /* * remove any tmp file if we blow */ cleanup(i) int i; { if (fj) (void) unlink(jname); exit(i); } /* * read, return ptr to seq, set dna, len, maxlen * skip lines starting with ';', '<', or '>' * seq in upper or lower case */ char * getseq(file, len) char *file; /* file name */ int *len; /* seq len */ { char line[1024], *pseq; register char *px, *py; int natgc, tlen; FILE *fp; if ((fp = fopen(file, "r")) == 0) { fprintf(stderr, "%s: can't read %s\n", prog, file); exit(1); } </pre>		
		stripname
		cleanup
		getseq

TABLE 1-continued

<pre> } tlen = natgc = 0; while (fgets(line, 1024, fp)) { if (*line == ';' *line == '<' *line == '>') continue; for (px = line; *px != '\n'; px++) if (isupper(*px) islower(*px)) tlen++; } if ((pseq = malloc((unsigned)(tlen+6))) == 0) { fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file); exit(1); } pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';</pre>		...getseq
<pre> py = pseq + 4; *len = tlen; rewind(fp); while (fgets(line, 1024, fp)) { if (*line == ';' *line == '<' *line == '>') continue; for (px = line; *px != '\n'; px++) { if (isupper(*px)) *py++ = *px; else if (islower(*px)) *py++ = toupper(*px); if (index("ATGCU", *(py-1))) natgc++; } } *py++ = '\0'; *py = '\0'; (void) fclose(fp); dna = natgc > (tlen/3); return(pseq+4); }</pre>		
<pre>char * g__calloc(msg, nx, sz) char *msg; /* program, calling routine */ int nx, sz; /* number and size of elements */ { char *px, *calloc(); if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) { if (*msg) { fprintf(stderr, "%s: g__calloc() failed %s (n= %d, sz= %d)\n", prog, msg, nx, sz); exit(1); } } return(px); }</pre>		g__calloc
<pre>/* * get final jmps from dx[] or tmp file, set pp[], reset dmax: main() */ readjmps()</pre>		readjmps
<pre>{ int fd = -1; int siz, i0, i1; register i, j, xx; if (fj) { (void) fclose(fj); if ((fd = open(jname, O_RDONLY, 0)) < 0) { fprintf(stderr, "%s: can't open() %s\n", prog, jname); cleanup(1); } } for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) { while (1) { for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--) ; if (j < 0 && dx[dmax].offset && fj) { (void) lseek(fd, dx[dmax].offset, 0); (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp)); (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset)); dx[dmax].ijmp = MAXJMP-1; } } }</pre>		...readjmps

TABLE 1-continued

```

        else
            break;
    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
            */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
/* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            i1++;
        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
/* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            i0++;
        }
    }
    else
        break;
}
/* reverse the order of jmps
*/
for (j = 0, i0--; j < i0; j++, i0--) {
    i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
    i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
}
for (j = 0, i1--; j < i1; j++, i1--) {
    i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
    i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
}
if (fd >= 0)
    (void) close(fd);
if (fj) {
    (void) unlink(jname);
    fj = 0;
    offset = 0;
}
}

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
writejumps(ix)
    int ix;
{
    char *mktemp();
    if (!fj) {
        if (mktemp(jname) < 0) {
            fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
            cleanup(1);
        }
        if ((fj = fopen(jname, "w")) == 0) {
            fprintf(stderr, "%s: can't write %s\n", prog, jname);
            exit(1);
        }
    }
    (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
    (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
}

```

writejumps

[0249]

TABLE 2

PRO	XXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXYYYYYYY	(Length = 12 amino acids)

% amino acid sequence identity = (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) = 5 divided by 15 = 33.3%

[0250]

TABLE 3

PRO	XXXXXXXXXX	(Length = 10 amino acids)
Comparison Protein	XXXXXXXXYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity = (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) = 5 divided by 10 = 50%

[0251]

TABLE 4

PRO-DNA	NNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity = (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) = 6 divided by 14 = 42.9%

[0252]

TABLE 5

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLVV	(Length = 9 nucleotides)

% nucleic acid sequence identity = (the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) = 4 divided by 12 = 33.3%

II. Compositions and Methods of the Invention

[0253] A. A Full-Length PRO Polypeptides

[0254] The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

[0255] As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual

nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

[0256] B. PRO Polypeptide Variants

[0257] In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

[0258] Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative mutations set forth, for instance, in U.S. Pat. No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

[0259] PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO polypeptide.

[0260] PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by poly-

merase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

[0261] In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

TABLE 6

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro; ala	ala
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe; norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala; tyr	leu
Pro (P)	ala	ala
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr; phe	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

[0262] Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- [0263] (1) hydrophobic: norleucine, met, ala, val, leu, ile;
- [0264] (2) neutral hydrophilic: cys, ser, thr;
- [0265] (3) acidic: asp, glu;
- [0266] (4) basic: asn, gln, his, lys, arg;
- [0267] (5) residues that influence chain orientation: gly, pro; and
- [0268] (6) aromatic: trp, tyr, phe.

[0269] Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

[0270] The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al., Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

[0271] Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W. H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

[0272] C. Modifications of PRO

[0273] Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C-terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis (diazocetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

[0274] Other modifications include deamidation of glutamyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the α -amino groups of lysine, arginine, and histidine side chains [T. E. Creighton, Proteins: Structure and Molecular Properties, W. H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

[0275] Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means),

and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

[0276] Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

[0277] Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 Sep. 1987, and in Aplin and Wriston, *CRC Crit. Rev. Biochem.*, pp. 259-306 (1981).

[0278] Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., *Arch. Biochem. Biophys.*, 259:52 (1987) and by Edge et al., *Anal. Biochem.*, 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., *Meth. Enzymol.*, 138:350 (1987).

[0279] Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Pat. Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

[0280] The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

[0281] In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl-terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., *Mol. Cell. Biol.*, 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., *Molecular and Cellular Biology*, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., *Protein Engineering*, 3(6):547-553

(1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., *BioTechnology*, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., *Science*, 255:192-194 (1992)]; an α -tubulin epitope peptide [Skinner et al., *J. Biol. Chem.*, 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., *Proc. Natl. Acad. Sci. USA*, 87:6393-6397 (1990)].

[0282] In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also U.S. Pat. No. 5,428,130 issued Jun. 27, 1995.

[0283] D. Preparation of PRO

[0284] The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., *Solid-Phase Peptide Synthesis*, W. H. Freeman Co., San Francisco, Calif. (1969); Merrifield, *J. Am. Chem. Soc.*, 85:2149-2154 (1963)]. In vitro protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, Calif.) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

[0285] 1. Isolation of DNA Encoding PRO

[0286] DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

[0287] Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual* (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., *supra*; Dieffenbach et al., *PCR Primer: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1995)].

[0288] The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like ^{32}P -labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

[0289] Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

[0290] Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

[0291] 2. Selection and Transformation of Host Cells

[0292] Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

[0293] Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example, CaCl_2 , CaPO_4 , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 Jun. 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Pat. No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene,

polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

[0294] Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 Apr. 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype tonA; *E. coli* W3110 strain 9E4, which has the complete genotype tonA ptr3; *E. coli* W3110 strain 27C7 (ATCC 55,244), which has the complete genotype tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan^r; *E. coli* W3110 strain 37D6, which has the complete genotype tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan^r; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant degP deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Pat. No. 4,946,783 issued 7 Aug. 1990. Alternatively, in vitro methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

[0295] In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Pat. No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al., J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424) *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilum* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*, *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces* such as *Schwanniomyces occidentalis* (EP 394,538 published 31 Oct. 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 Jan. 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983];

Tilburn et al., *Gene*, 26:205-221 [1983]; Yelton et al., *Proc. Natl. Acad. Sci. USA*, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, *EMBO J.*, 4:475-479 [1985]). Methylo-tropic yeasts are suitable herein and include, but are not limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, *The Biochemistry of Methylo-trophs*, 269 (1982).

[0296] Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., *J. Gen. Virol.*, 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, *Biol. Reprod.*, 23:243-251 (1980)); human lung cells (W138, ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

[0297] 3. Selection and Use of a Replicable Vector

[0298] The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

[0299] The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces* α -factor leaders, the latter described in U.S. Pat. No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 Apr. 1990), or the signal described in WO 90/13646 published 15 Nov. 1990. In mammalian cell expression, mammalian signal sequences

may be used to direct secretion of the protein, such as signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

[0300] Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2μ plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for cloning vectors in mammalian cells.

[0301] Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

[0302] An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., *Proc. Natl. Acad. Sci. USA*, 77:4216 (1980). A suitable selection gene for use in yeast is the trp 1 gene present in the yeast plasmid YRp7 [Stinchcomb et al., *Nature*, 282:39 (1979); Kingsman et al., *Gene*, 7:141 (1979); Tschemper et al., *Gene*, 10:157 (1980)]. The trp 1 gene provides a selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, *Genetics*, 85:12 (1977)].

[0303] Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well known. Promoters suitable for use with prokaryotic hosts include the β -lactamase and lactose promoter systems [Chang et al., *Nature*, 275:615 (1978); Goeddel et al., *Nature*, 281:544 (1979)], alkaline phosphatase, a tryptophan (trp) promoter system [Goeddel, *Nucleic Acids Res.*, 8:4057 (1980); EP 36,776], and hybrid promoters such as the tac promoter [deBoer et al., *Proc. Natl. Acad. Sci. USA*, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

[0304] Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., *J. Biol. Chem.*, 255:2073 (1980)] or other glycolytic enzymes [Hess et al., *J. Adv. Enzyme Reg.*, 7:149 (1968); Holland, *Biochemistry*, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

[0305] Other yeast promoters, which are inducible promoters having the additional advantage of transcription

controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

[0306] PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 Jul. 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

[0307] Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin, α -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

[0308] Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

[0309] Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., *Nature*, 293:620-625 (1981); Mantei et al., *Nature*, 281:40-46 (1979); EP 117,060; and EP 117,058.

[0310] 4. Detecting Gene Amplification/Expression

[0311] Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, *Proc. Natl. Acad. Sci. USA*, 77:5201(1980)], dot blotting (DNA analysis), or in situ hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

[0312] Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

[0313] 5. Purification of Polypeptide

[0314] Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

[0315] It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, *Methods in Enzymology*, 182 (1990); Scopes, *Protein Purification: Principles and Practice*, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

[0316] E. Uses for PRO

[0317] Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

[0318] The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases.

Hybridization probes may be labeled by a variety of labels, including radionucleotides such as ^{32}P or ^{35}S , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

[0319] Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

[0320] Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

[0321] Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable in vivo (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

[0322] Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

[0323] Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example, CaPO_4 -mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either in vivo or ex vivo. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retro-

virus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

[0324] Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

[0325] Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

[0326] Antisense or sense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

[0327] The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

[0328] Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome using known techniques, such as in situ hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

[0329] When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified. Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein

binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

[0330] Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or “knock out” animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Pat. Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

[0331] Alternatively, non-human homologues of PRO can be used to construct a PRO “knock out” animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, *Cell*, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., *Cell*, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a “knock out” animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knock-

out animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

[0332] Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve in vivo synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. “Gene therapy” includes both conventional gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes in vivo. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane. (Zamecnik et al., *Proc. Natl. Acad. Sci. USA* 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

[0333] There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells in vitro, or in vivo in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells in vitro include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation method, etc. The currently preferred in vivo gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau et al., *Trends in Biotechnology* 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu et al., *J. Biol. Chem.* 262, 4429-4432 (1987); and Wagner et al., *Proc. Natl. Acad. Sci. USA* 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson et al., *Science* 256, 808-813 (1992).

[0334] The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

[0335] The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

[0336] The PRO polypeptides and nucleic acid molecules of the present invention may also be used diagnostically for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

[0337] The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEENTM, PLURONICSTM or PEG.

[0338] The formulations to be used for in vivo administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

[0339] Therapeutic compositions herein 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.

[0340] The route of administration is in accord with known methods, e.g. injection or infusion by intravenous, intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

[0341] Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In *Toxicokinetics and New Drug Development*, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

[0342] When in vivo administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal dosage

amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 μ g/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may necessitate delivery in a manner different from that to another organ or tissue.

[0343] Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon-(rhIFN-), interleukin-2, and MN rgp120. Johnson et al., *Nat. Med.*, 2:795-799 (1996); Yasuda, *Biomed. Ther.*, 27:1221-1223 (1993); Hora et al., *Bio/Technology*, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in *Vaccine Design: The Subunit and Adjuvant Approach*, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

[0344] The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), *Biodegradable Polymers as Drug Delivery Systems* (Marcel Dekker: New York, 1990), pp. 1-41.

[0345] This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

[0346] The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

[0347] All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

[0348] In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction

mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

[0349] If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, *Nature* (London), 340:245-246 (1989); Chien et al., *Proc. Natl. Acad. Sci. USA*, 88:9578-9582 (1991)) as disclosed by Chevray and Nathans, *Proc. Natl. Acad. Sci. USA*, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-lac Z reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for β -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

[0350] Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the

reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

[0351] To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor 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). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide 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 PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single clone that encodes the putative receptor.

[0352] As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

[0353] In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

[0354] More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain anti-

bodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

[0355] Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is 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 (triple helix—see Lee et al., *Nucl. Acids Res.*, 6:3073 (1979); Cooney et al., *Science*, 241: 456 (1988); Dervan et al., *Science*, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA in vivo and blocks translation of the mRNA molecule into the PRO polypeptide (antisense—Okano, *Neurochem.*, 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, Fla., 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed in vivo to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

[0356] Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

[0357] Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA. Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, *Current Biology*, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published Sep. 18, 1997).

[0358] Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, supra.

[0359] These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

[0360] Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

[0361] F. Anti-PRO Antibodies

[0362] The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

[0363] 1. Polyclonal Antibodies

[0364] The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate). The immunization protocol may be selected by one skilled in the art without undue experimentation.

[0365] 2. Monoclonal Antibodies

[0366] The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, *Nature*, 256:495 (1975). In a hybridoma method, a mouse, 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.

[0367] The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. 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.

[0368] 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. 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].

[0369] The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. 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 immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

[0370] After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown in vivo as ascites in a mammal.

[0371] 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, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

[0372] The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Pat. No. 4,816,567. 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). The hybridoma 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 transfected 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.

[0373] 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.

[0374] 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.

[0375] 3. Human and Humanized Antibodies

[0376] The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')₂ or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. 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); and Presta, Curr. Op. Struct. Biol., 2:593-596 (1992)].

[0377] Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which 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 method of Winter and co-workers [Jones et al., Nature, 321:522-525 (1986); Riechmann et al., Nature, 332:323-327 (1988); Verhoeven 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 corre-

sponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

[0378] Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of 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 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; 5,661,016, and in the following scientific publications: Marks et al., *Bio/Technology* 10, 779-783 (1992); Lonberg et al., *Nature* 368 856-859 (1994); Morrison, *Nature* 368, 812-13 (1994); Fishwild et al., *Nature Biotechnology* 14, 845-51 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.* 13 65-93 (1995).

[0379] The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

[0380] 4. Bispecific Antibodies

[0381] Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

[0382] 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 13 May 1993, and in Traunecker et al., *EMBO J.*, 10:3655-3659 (1991).

[0383] 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-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., *Methods in Enzymology*, 121:210 (1986).

[0384] According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

[0385] Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')₂ bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical linkage. Brennan et al., *Science* 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')₂ fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

[0386] Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., *J. Exp. Med.* 175:217(1992) describe the production of a fully humanized bispecific antibody F(ab')₂ molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to directed chemical coupling in vitro to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

[0387] Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers. Kostelny et al., *J. Immunol.* 148 (5):1547-1553(1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the

hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448(1993) has provided an alternative mechanism for making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V_H) connected to a light-chain variable domain (V_L) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V_H and V_L domains of one fragment are forced to pair with the complementary V_L and V_H domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber et al., J. Immunol. 152:5368 (1994). Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt et al. J., Immunol. 147:60(1991).

[0388] Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (Fc γ R), such as Fc γ RI (CD64), Fc γ RII (CD32) and Fc γ RIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

[0389] 5. Heteroconjugate Antibodies

[0390] Heteroconjugate antibodies are also within the scope of 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 treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. 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 thioether 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.

[0391] 6. Effector Function Engineering

[0392] It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, e.g., the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron et al., J. Exp Med., 176: 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with

enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff et al. Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson et al., Anti-Cancer Drug Design, 3: 219-230 (1989).

[0393] 7. Immunoconjugates

[0394] The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent such as a chemotherapeutic agent, toxin (e.g., an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (i.e., a radioconjugate).

[0395] Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolacca americana* proteins (PAPI, PAPII, and PAP-S), *momordica charantia* inhibitor, curcin, crotin, *sapaonaria officinalis* inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include ^{212}Bi , ^{131}I , ^{131}In , ^{90}Y , and ^{186}Re . Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimide HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta et al., Science, 238: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionuclide to the antibody. See WO 94/11026.

[0396] In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (e.g., avidin) that is conjugated to a cytotoxic agent (e.g., a radionuclide).

[0397] 8. Immunoliposomes

[0398] The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein et al., Proc. Natl. Acad. Sci. USA, 82: 3688 (1985); Hwang et al., Proc. Natl. Acad. Sci. USA, 77: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Pat. No. 5,013,556.

[0399] Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid compo-

sition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin et al., J. Biol. Chem., 257: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon et al., J. National Cancer Inst., 81(19): 1484 (1989).

[0400] 9. Pharmaceutical Compositions of Antibodies

[0401] Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

[0402] If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, e.g., Marasco et al., Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

[0403] The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nano-capsules) or in macroemulsions. Such techniques are disclosed in Remington's *Pharmaceutical Sciences*, supra.

[0404] The formulations to be used for in vivo administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

[0405] Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and γ ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOTTM

(injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37° C., resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S—S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

[0406] G. Uses for Anti-PRO Antibodies

[0407] The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, e.g., detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. 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 ³H, ¹⁴C, ³²P, ³⁵S, or ¹²⁵I, a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase 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); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

[0408] Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such as Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

[0409] The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

[0410] All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

[0411] Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, Va.

Example 1

[0412] Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

[0413] The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (e.g., Dayhoff, GenBank), and proprietary databases (e.g., LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, Calif.). The search was performed using the computer program BLAST or BLAST-2 (Altschul et al., *Methods in Enzymology* 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, Wash.).

[0414] Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

[0415] Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5 kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel et al., *Current Protocols in Molecular Biology*, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

[0416] The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, Calif. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; see, Holmes et al., *Science*, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

Example 2

[0417] Isolation of cDNA Clones by Amylase Screening

[0418] 1. Preparation of Oligo dT Primed cDNA Library

[0419] mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, Calif. (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, Md. (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linked cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

[0420] 2. Preparation of Random Primed cDNA Library

[0421] A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

[0422] 3. Transformation and Detection

[0423] DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37° C. for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37° C.). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, e.g. CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

[0424] The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

[0425] The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT^{alpha}, ura3-52, leu2-3, leu2-112, his3-11, his3-15, MAL⁺, SUC⁺, GAL⁺. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in sec 71, sec 72, sec 62, with truncated sec 71 being most preferred. Alternatively,

antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (e.g., SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

[0426] Transformation was performed based on the protocol outlined by Gietz et al., Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30° C. The YEPD broth was prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., p. 207 (1994). The overnight culture was then diluted to about 2×10^6 cells/ml (approx. OD₆₀₀=0.1) into fresh YEPD broth (500 ml) and regrown to 1×10^7 cells/ml (approx. OD₆₀₀=0.4-0.5).

[0427] The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li₂OOCCH₃), and resuspended into LiAc/TE (2.5 ml).

[0428] Transformation took place by mixing the prepared cells (100 μ l) with freshly denatured single stranded salmon testes DNA (Lofstrand Labs, Gaithersburg, Md.) and transforming DNA (1 μ g, vol.<10 μ l) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600 μ l, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li₂OOCCH₃, pH 7.5) was added. This mixture was gently mixed and incubated at 30° C. while agitating for 30 minutes. The cells were then heat shocked at 42° C. for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500 μ l, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200 μ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

[0429] Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

[0430] The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser et al., Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, N.Y., p. 208-210 (1994). Transformants were grown at 30° C. for 2-3 days.

[0431] The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely et al., Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

[0432] The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

[0433] 4. Isolation of DNA by PCR Amplification

[0434] When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30 μ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5 μ l) was used as a template for the PCR reaction in a 25 μ l volume containing: 0.5 μ l KlenTaq (Clontech, Palo Alto, Calif.); 4.0 μ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5 μ l Kentaq buffer (Clontech); 0.25 μ l forward oligo 1; 0.25 μ l reverse oligo 2; 12.5 μ l distilled water. The sequence of the forward oligonucleotide 1 was:

[0435] 5'-TGTAACGACGGCCAGTTAAATA-GACCTGCAATTATTAATCT-3' (SEQ ID NO:169)

[0436] The sequence of reverse oligonucleotide 2 was:

[0437] 5'-CAGGAAACAGCTATGACCACCTG-CACACCTGCAAATCCATT-3' (SEQ ID NO:170)

[0438] PCR was then performed as follows:

a.		Denature	92° C.,	5 minutes
b.	3 cycles of:	Denature	92° C.,	30 seconds
		Anneal	59° C.,	30 seconds
		Extend	72° C.,	60 seconds
c.	3 cycles of:	Denature	92° C.,	30 seconds
		Anneal	57° C.,	30 seconds
		Extend	72° C.,	60 seconds
d.	25 cycles of:	Denature	92° C.,	30 seconds
		Anneal	55° C.,	30 seconds
		Extend	72° C.,	60 seconds
e.		Hold	4° C.	

[0439] The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

[0440] Following the PCR, an aliquot of the reaction (5 μ l) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook et al., supra. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, Calif.).

Example 3

[0441] Isolation of cDNA Clones Using Signal Algorithm Analysis

[0442] Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, Calif.) upon ESTs as well as clustered and assembled EST fragments from public (e.g., GenBank) and/or private (LIFESeq®, Incyte Pharmaceuticals, Inc., Palo Alto, Calif.) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

Example 4

[0443] Isolation of cDNA Clones Encoding Human PRO Polypeptides

[0444] Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, Va. 20110-2209, USA (ATCC) as shown in Tabl 7 below.

TABLE 7

Material	ATCC Dep. No.	Deposit Date
DNA26843-1389	203099	Aug. 4, 1998
DNA30867-1335	209807	Apr. 28, 1998
DNA34431-1177	209399	Oct. 17, 1997
DNA38268-1188	209421	Oct. 28, 1997
DNA40621-1440	209922	Jun. 2, 1998
DNA40625-1189	209788	Apr. 21, 1998
DNA45409-2511	203579	Jan. 12, 1999
DNA45495-1550	203156	Aug. 25, 1998
DNA49820-1427	209932	Jun. 2, 1998
DNA56406-1704	203478	Nov. 17, 1998
DNA56410-1414	209923	Jun. 2, 1998
DNA56436-1448	209902	May 27, 1998
DNA56855-1447	203004	Jun. 23, 1998
DNA56860-1510	209952	Jun. 9, 1998
DNA56862-1343	203174	Sep. 1, 1998
DNA56868-1478	203024	Jun. 23, 1998
DNA56869-1545	203161	Aug. 25, 1998
DNA57704-1452	209953	Jun. 9, 1998
DNA58723-1588	203133	Aug. 18, 1998
DNA57827-1493	203045	Jul. 1, 1998
DNA58737-1473	203136	Aug. 18, 1998
DNA58846-1409	209957	Jun. 9, 1998
DNA58850-1495	209956	Jun. 9, 1998
DNA58855-1422	203018	Jun. 23, 1998
DNA59211-1450	209960	Jun. 9, 1998

TABLE 7-continued

Material	ATCC Dep. No.	Deposit Date
DNA59212-1627	203245	Sep. 9, 1998
DNA59213-1487	209959	Jun. 9, 1998
DNA59605-1418	203005	Jun. 23, 1998
DNA59609-1470	209963	Jun. 9, 1998
DNA59610-1556	209990	Jun. 16, 1998
DNA59837-2545	203658	Feb. 9, 1999
DNA59844-2542	203650	Feb. 9, 1999
DNA59854-1459	209974	Jun. 16, 1998
DNA60625-1507	209975	Jun. 16, 1998
DNA60629-1481	209979	Jun. 16, 1998
DNA61755-1554	203112	Aug. 11, 1998
DNA62812-1594	203248	Sep. 9, 1998
DNA62815-1576	203247	Sep. 9, 1998
DNA64881-1602	203240	Sep. 9, 1998
DNA64886-1601	203241	Sep. 9, 1998
DNA64902-1667	203317	Oct. 6, 1998
DNA64950-1590	203224	Sep. 15, 1998
DNA65403-1565	203230	Sep. 15, 1998
DNA66308-1537	203159	Aug. 25, 1998
DNA66519-1535	203236	Sep. 15, 1998
DNA66521-1583	203225	Sep. 15, 1998
DNA66658-1584	203229	Sep. 15, 1998
DNA66660-1585	203279	Sep. 22, 1998
DNA66663-1598	203268	Sep. 22, 1998
DNA66674-1599	203281	Sep. 22, 1998
DNA68862-2546	203652	Feb. 9, 1999
DNA68866-1644	203283	Sep. 22, 1998
DNA68871-1638	203280	Sep. 22, 1998
DNA68880-1676	203319	Oct. 6, 1998
DNA68883-1691	203535	Dec. 15, 1998
DNA68885-1678	203311	Oct. 6, 1998
DNA71277-1636	203285	Sep. 22, 1998
DNA73727-1673	203459	Nov. 3, 1998
DNA73734-1680	203363	Oct. 20, 1998
DNA73735-1681	203356	Oct. 20, 1998
DNA76393-1664	203323	Oct. 6, 1998
DNA77301-1708	203407	Oct. 27, 1998
DNA77568-1626	203134	Aug. 18, 1998
DNA77626-1705	203536	Dec. 15, 1998
DNA81754-2532	203542	Dec. 15, 1998
DNA81757-2512	203543	Dec. 15, 1998
DNA82302-2529	203534	Dec. 15, 1998
DNA82340-2530	203547	Dec. 22, 1998
DNA83500-2506	203391	Oct. 29, 1998
DNA84920-2614	203966	Apr. 27, 1999
DNA85066-2534	203588	Jan. 12, 1999
DNA86571-2551	203660	Feb. 9, 1999
DNA87991-2540	203656	Feb. 9, 1999
DNA92238-2539	203602	Jan. 20, 1999
DNA96042-2682	PTA-382	Jul. 20, 1999
DNA96787-2534	203589	Jan. 12, 1999
DNA125185-2806	PTA-1031	Dec. 7, 1999
DNA147531-2821	PTA-1185	Jan. 11, 2000
DNA115291-2681	PTA-202	Jun. 8, 1999
DNA164625-28890	PTA-1535	Mar. 21, 2000
DNA131639-2874	PTA-1784	Apr. 25, 2000
DNA79230-2525	203549	Dec. 22, 1998

[0445] These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent

application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

[0446] The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

Example 5

[0447] Use of PRO as a Hybridization Probe

[0448] The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

[0449] DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

[0450] Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5×SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH 6.8, 2×Denhardt's solution, and 10% dextran sulfate at 42° C. for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1×SSC and 0.1% SDS at 42° C.

[0451] DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

Example 6

[0452] Expression of PRO in *E. coli*

[0453] This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

[0454] The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., *Gene*, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

[0455] The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., *supra*. Transformants are identified by their

ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

[0456] Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

[0457] After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

[0458] PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) Ion galE rpoHts (htpRts) clpP(lacIq)). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30° C. with shaking until an O.D.600 of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH₄)₂SO₄, 0.71 g sodium citrate-2H₂O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO₄) and grown for approximately 20-30 hours at 30° C. with shaking. Samples are removed to verify expression by SDS-page analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

[0459] *E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4° C. This step results in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Ultrapure grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4° C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

[0460] The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM

cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4° C. for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

[0461] Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

[0462] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

Example 7

[0463] Expression of PRO in Mammalian Cells

[0464] This example illustrates preparation of a potentially glycosylated form of PRO by recombinant expression in mammalian cells.

[0465] The vector, pRK5 (see EP 307,247, published Mar. 15, 1989), is employed as the expression vector. Optionally, the PRO DNA is ligated into pRK5 with selected restriction enzymes to allow insertion of the PRO DNA using ligation methods such as described in Sambrook et al., *supra*. The resulting vector is called pRK5-PRO.

[0466] In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10 μ g pRK5-PRO DNA is mixed with about 1 μ g DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500 μ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M CaCl₂. To this mixture is added, dropwise, 500 μ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM NaPO₄, and a precipitate is allowed to form for 10 minutes at 25° C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37° C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

[0467] Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture

medium (alone) or culture medium containing 200 μ Ci/ml ³⁵S-cysteine and 200 μ Ci/ml ³⁵S-methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

[0468] In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700 μ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5 μ g/ml bovine insulin and 0.1 μ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

[0469] In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as CaPO₄ or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as ³⁵S-methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

[0470] Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be concentrated and purified by any selected method, such as by Ni²⁺-chelate affinity chromatography.

[0471] PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

[0472] Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

[0473] Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using

standard techniques as described in Ausubel et al., *Current Protocols of Molecular Biology*, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., *Nucl. Acids Res.* 24:9 (1774-1779) (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

[0474] Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect® (Quiagen), Dosper® or Eugene® (Boehringer Mannheim). The cells are grown as described in Lucas et al., *supra*. Approximately 3×10^7 cells are frozen in an ampule for further growth and production as described below.

[0475] The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2 μ m filtered PS20 with 5% 0.2 μ m diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37° C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with 3×10^5 cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Pat. No. 5,122,469, issued Jun. 16, 1992 may actually be used. A 3L production spinner is seeded at 1.2×10^6 cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33° C., and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22 μ m filter. The filtrate was either stored at 4° C. or immediately loaded onto columns for purification.

[0476] For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 mL Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 mL/min. at 4° C. After loading, the column is washed with additional equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 mL G25 Superfine (Pharmacia) column and stored at -80° C.

[0477] Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The condi-

tioned medium is pumped onto a 5 mL Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 mL fractions into tubes containing 275 μ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

[0478] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

Example 8

[0479] Expression of PRO in Yeast

[0480] The following method describes recombinant expression of PRO in yeast.

[0481] First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

[0482] Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

[0483] Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

[0484] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

Example 9

[0485] Expression of PRO in Baculovirus-Infected Insect Cells

[0486] The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

[0487] The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if

the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then digested with those selected restriction enzymes and subcloned into the expression vector.

[0488] Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4-5 days of incubation at 28° C., the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilly et al., *Baculovirus expression vectors: A Laboratory Manual*, Oxford: Oxford University Press (1994).

[0489] Expressed poly-his tagged PRO can then be purified, for example, by Ni²⁺-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., *Nature*, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl₂; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 μ m filter. A Ni²⁺-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A₂₈₀ with loading buffer, at which point fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A₂₈₀ baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni²⁺-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His₁₀-tagged PRO are pooled and dialyzed against loading buffer.

[0490] Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

[0491] Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

Example 10

[0492] Preparation of Antibodies that Bind PRO

[0493] This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

[0494] Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, *supra*. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled artisan without undue experimentation.

[0495] Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, Mont.) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.

[0496] After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected with a final intravenous injection of PRO. Three to four days later, the mice are sacrificed and the spleen cells are harvested. The spleen cells are then fused (using 35% polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymidine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

[0497] The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill in the art.

[0498] The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in tissue culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

Example 11

[0499] Purification of PRO Polypeptides Using Specific Antibodies

[0500] Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide is purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

[0501] Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation or chromatography on immobilized Protein A. Partially purified immunoglobulin is covalently attached to a chromatographic resin such as CnBr-activated SEPHAROSE™ (Pharmacia LKB Biotech-

nology). The antibody is coupled to the resin, the resin is blocked, and the derivative resin is washed according to the manufacturer's instructions.

[0502] Such an immunoaffinity column is utilized in the purification of PRO polypeptide by preparing a fraction from cells containing PRO polypeptide in a soluble form. This preparation is derived by solubilization of the whole cell or of a subcellular fraction obtained via differential centrifugation by the addition of detergent or by other methods well known in the art. Alternatively, soluble PRO polypeptide containing a signal sequence may be secreted in useful quantity into the medium in which the cells are grown.

[0503] A soluble PRO polypeptide-containing preparation is passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRO polypeptide (e.g., high ionic strength buffers in the presence of detergent). Then, the column is eluted under conditions that disrupt antibody/PRO polypeptide binding (e.g., a low pH buffer such as approximately pH 2-3, or a high concentration of a chaotrope such as urea or thiocyanate ion), and PRO polypeptide is collected.

Example 12

[0504] Drug Screening

[0505] This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

[0506] Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with a PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

[0507] Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on Sep. 13, 1984. Briefly stated,

large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

[0508] This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

Example 13

[0509] Rational Drug Design

[0510] The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (i.e., a PRO polypeptide) or of small molecules with which they interact, e.g., agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide in vivo (c.f., Hodgson, *Bio/Technology*, 9: 19-21(1991)).

[0511] In one approach, the three-dimensional structure of the PRO polypeptide, or of a PRO polypeptide-inhibitor complex, is determined by X-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, *Biochemistry*, 31:7796-7801(1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda et al., *J. Biochem.*, 113:742-746(1993).

[0512] It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

[0513] By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to

perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to X-ray crystallography.

Example 14

[0514] Pericyte c-Fos Induction (Assay 93)

[0515] This assay shows that certain polypeptides of the invention act to induce the expression of c-fos in pericyte cells and, therefore, are useful not only as diagnostic markers for particular types of pericyte-associated tumors but also for giving rise to antagonists which would be expected to be useful for the therapeutic treatment of pericyte-associated tumors. Induction of c-fos expression in pericytes is also indicative of the induction of angiogenesis and, as such, PRO polypeptides capable of inducing the expression of c-fos would be expected to be useful for the treatment of conditions where induced angiogenesis would be beneficial including, for example, wound healing, and the like. Specifically, on day 1, pericytes are received from VEC Technologies and all but 5 ml of media is removed from flask. On day 2, the pericytes are trypsinized, washed, spun and then plated onto 96 well plates. On day 7, the media is removed and the pericytes are treated with 100 μ l of PRO polypeptide test samples and controls (positive control=DME+5% serum+/-PDGF at 500 ng/ml; negative control=protein 32). Replicates are averaged and SD/CV are determined. Fold increase over Protein 32 (buffer control) value indicated by chemiluminescence units (RLU) luminometer reading verses frequency is plotted on a histogram. Two-fold above Protein 32 value is considered positive for the assay. ASY Matrix: Growth media=low glucose DMEM=20% FBS+1 $\times$ pen strep+1 $\times$ fungizone. Assay Media=low glucose DMEM+5% FBS.

[0516] The following polypeptides tested positive in this assay: PRO1347 and PRO1340.

Example 15

[0517] Ability of PRO Polypeptides to Stimulate the Release of Proteoglycans from Cartilage (Assay 97)

[0518] The ability of various PRO polypeptides to stimulate the release of proteoglycans from cartilage tissue was tested as follows.

[0519] The metacarpophalangeal joint of 4-6 month old pigs was aseptically dissected, and articular cartilage was removed by free hand slicing being careful to avoid the underlying bone. The cartilage was minced and cultured in bulk for 24 hours in a humidified atmosphere of 95% air, 5% CO₂ in serum free (SF) media (DME/F12 1:1) with 0.1% BSA and 100U/ml penicillin and 100 μ g/ml streptomycin. After washing three times, approximately 100 mg of articular cartilage was aliquoted into micronics tubes and incubated for an additional 24 hours in the above SF media. PRO polypeptides were then added at 1% either alone or in combination with 18 ng/ml interleukin-1 α , a known stimulator of proteoglycan release from cartilage tissue. The supernatant was then harvested and assayed for the amount of proteoglycans using the 1,9-dimethyl-methylene blue (DMB) colorimetric assay (Farndale and Buttle, Biochem. Biophys. Acta 883:173-177 (1985)). A positive result in this

assay indicates that the test polypeptide will find use, for example, in the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis.

[0520] When various PRO polypeptides were tested in the above assay, the polypeptides demonstrated a marked ability to stimulate release of proteoglycans from cartilage tissue both basally and after stimulation with interleukin-1 α and at 24 and 72 hours after treatment, thereby indicating that these PRO polypeptides are useful for stimulating proteoglycan release from cartilage tissue. As such, these PRO polypeptides are useful for the treatment of sports-related joint problems, articular cartilage defects, osteoarthritis or rheumatoid arthritis. The polypeptides testing positive in this assay are: PRO1565, PRO1693, PRO1801 and PRO10096.

Example 16

[0521] Detection of Polypeptides that Affect Glucose or FFA Uptake in Skeletal Muscle (Assay 106)

[0522] This assay is designed to determine whether PRO polypeptides show the ability to affect glucose or FFA uptake by skeletal muscle cells. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of disorders where either the stimulation or inhibition of glucose uptake by skeletal muscle would be beneficial including, for example, diabetes or hyper- or hypo-insulinemia.

[0523] In a 96 well format, PRO polypeptides to be assayed are added to primary rat differentiated skeletal muscle, and allowed to incubate overnight. Then fresh media with the PRO polypeptide and +/- insulin are added to the wells. The sample media is then monitored to determine glucose and FFA uptake by the skeletal muscle cells. The insulin will stimulate glucose and FFA uptake by the skeletal muscle, and insulin in media without the PRO polypeptide is used as a positive control, and a limit for scoring. As the PRO polypeptide being tested may either stimulate or inhibit glucose and FFA uptake, results are scored as positive in the assay if greater than 1.5 times or less than 0.5 times the insulin control.

[0524] The following PRO polypeptides tested positive as either stimulators or inhibitors of glucose and/or FFA uptake in this assay: PRO4405.

Example 17

[0525] Identification of PRO Polypeptides that Stimulate TNF- α Release in Human Blood (Assay 128)

[0526] This assay shows that certain PRO polypeptides of the present invention act to stimulate the release of TNF- α in human blood. PRO polypeptides testing positive in this assay are useful for, among other things, research purposes where stimulation of the release of TNF- α would be desired and for the therapeutic treatment of conditions wherein enhanced TNF- α release would be beneficial. Specifically, 200 μ l of human blood supplemented with 50 mM Hepes buffer (pH 7.2) is aliquotted per well in a 96 well test plate. To each well is then added 300 μ l of either the test PRO polypeptide in 50 mM Hepes buffer (at various concentrations) or 50 mM Hepes buffer alone (negative control) and the plates are incubated at 37° C. for 6 hours. The samples

are then centrifuged and 50 μ l of plasma is collected from each well and tested for the presence of TNF- α by ELISA assay. A positive in the assay is a higher amount of TNF- α in the PRO polypeptide treated samples as compared to the negative control samples.

[0527] The following PRO polypeptides tested positive in this assay: PRO263, PRO295, PRO1282, PRO1063, PRO1356, PRO3543, and PRO5990.

Example 18

[0528] Tumor Versus Normal Differential Tissue Expression Distribution

[0529] Oligonucleotide probes were constructed from some of the PRO polypeptide-encoding nucleotide sequences shown in the accompanying figures for use in quantitative PCR amplification reactions. The oligonucleotide probes were chosen so as to give an approximately 200-600 base pair amplified fragment from the 3' end of its associated template in a standard PCR reaction. The oligonucleotide probes were employed in standard quantitative PCR amplification reactions with cDNA libraries isolated from different human tumor and normal human tissue samples and analyzed by agarose gel electrophoresis so as to obtain a quantitative determination of the level of expression of the PRO polypeptide-encoding nucleic acid in the various tumor and normal tissues tested. β -actin was used as a control to assure that equivalent amounts of nucleic acid was used in each reaction. Identification of the differential expression of the PRO polypeptide-encoding nucleic acid in one or more tumor tissues as compared to one or more normal tissues of the same tissue type renders the molecule useful diagnostically for the determination of the presence or absence of tumor in a subject suspected of possessing a tumor as well as therapeutically as a target for the treatment of a tumor in a subject possessing such a tumor. These assays provided the following results.

Molecule	is more highly expressed in:	as compared to:
DNA26843-1389	normal lung	lung tumor
	rectum tumor	normal rectum
DNA30867-1335	normal kidney	kidney tumor
DNA40621-1440	normal lung	lung tumor
DNA40625-1189	normal lung	lung tumor
DNA45409-2511	melanoma tumor	normal akin
DNA56406-1704	kidney tumor	normal kidney
	normal skin	melanoma tumor
DNA56410-1414	normal stomach	stomach tumor
DNA56436-1448	normal skin	melanoma tumor
DNA56855-1447	normal esophagus	esophageal tumor
	rectum tumor	normal rectum
DNA56860-1510	normal kidney	kidney tumor
	rectum tumor	normal rectum
DNA56862-1343	kidney tumor	normal kidney
	normal lung	lung tumor
DNA56868-1478	normal stomach	stomach tumor
	normal lung	lung tumor
DNA56869-1545	normal esophagus	esophageal tumor
	normal skin	melanoma tumor
DNA57704-1452	normal stomach	stomach tumor
	rectum tumor	normal rectum
DNA58723-1588	normal stomach	stomach tumor
	kidney tumor	normal kidney
	normal akin	melanoma tumor
DNA57827-1493	normal stomach	stomach tumor
	normal skin	melanoma tumor

-continued

Molecule	is more highly expressed in:	as compared to:
DNA58737-1473	esophageal tumor	normal esophagus
	normal stomach	stomach tumor
DNA58846-1409	lung tumor	normal lung
DNA58850-1495	esophageal tumor	normal esophagus
	kidney tumor	normal kidney
DNA58855-1422	normal stomach	stomach tumor
	rectum tumor	normal rectum
DNA59211-1450	normal kidney	kidney tumor
DNA59212-1627	normal skin	melanoma tumor
DNA59213-1487	normal stomach	stomach tumor
	normal skin	melanoma tumor
DNA59605-1418	melanoma tumor	normal skin
DNA59609-1470	esophageal tumor	normal esophagus
DNA59610-1556	esophageal tumor	normal esophagus
	lung tumor	normal lung
	normal skin	melanoma tumor
DNA59837-2545	normal skin	melanoma tumor
DNA59844-2542	normal skin	melanoma tumor
	esophageal tumor	normal esophagus
DNA59854-1459	normal esophagus	esophageal tumor
	stomach tumor	normal stomach
	normal lung	lung tumor
DNA60625-1507	normal lung	lung tumor
DNA60629-1481	normal esophagus	esophageal tumor
	normal rectum	rectum tumor
DNA61755-1554	normal stomach	stomach tumor
	kidney tumor	normal kidney
DNA62812-1594	normal stomach	stomach tumor
	normal lung	lung tumor
	normal rectum	rectum tumor
	normal skin	melanoma tumor
DNA62815-1576	esophageal tumor	normal esophagus
DNA64881-1602	normal stomach	stomach tumor
	normal lung	lung tumor
DNA64902-1667	esophageal tumor	normal esophagus
	kidney tumor	normal kidney
DNA65403-1565	normal esophagus	esophageal tumor
DNA66308-1537	normal lung	lung tumor
DNA66519-1535	kidney tumor	normal kidney
DNA66521-1583	normal esophagus	esophageal tumor
	normal stomach	stomach tumor
	normal lung	lung tumor
	normal rectum	rectum tumor
	normal skin	melanoma tumor
DNA66658-1584	normal lung	lung tumor
	melanoma tumor	normal skin
DNA66660-1585	lung tumor	normal lung
DNA66674-1599	kidney tumor	normal kidney
	normal lung	lung tumor
DNA68862-2546	melanoma tumor	normal skin
DNA68866-1644	normal stomach	stomach tumor
DNA68871-1638	lung tumor	normal lung
	normal skin	melanoma tumor
DNA68880-1676	normal lung	lung tumor
	normal skin	melanoma tumor
DNA68883-1691	esophageal tumor	normal esophagus
DNA68885-1678	lung tumor	normal lung
DNA71277-1636	normal stomach	stomach tumor
DNA73734-1680	normal lung	lung tumor
DNA73735-1681	esophageal tumor	normal esophagus
	normal kidney	kidney tumor
	lung tumor	normal lung
	normal skin	melanoma tumor
DNA76393-1664	esophageal tumor	normal esophagus
	stomach tumor	normal stomach
	lung tumor	normal lung
	rectum tumor	normal rectum
DNA77568-1626	normal stomach	stomach tumor
	lung tumor	normal lung

-continued

Molecule	is more highly expressed in:	as compared to:
DNA77626-1705	normal rectum	rectum tumor
DNA81754-2532	normal skin	melanoma tumor
DNA81757-2512	esophageal tumor	normal esophagus
	normal stomach	stomach tumor
	melanoma tumor	normal skin
DNA82302-2529	normal stomach	stomach tumor
	normal lung	lung tumor
DNA82340-2530	normal esophagus	esophageal tumor
DNA85066-2534	lung tumor	normal lung
	normal skin	melanoma tumor
DNA87991-2540	esophageal tumor	normal esophagus
DNA92238-2539	normal skin	melanoma tumor
DNA96787-2534	normal kidney	kidney tumor

Example 19

[0530] Identification of Receptor/Ligand Interactions

[0531] In this assay, various PRO polypeptides are tested for ability to bind to a panel of potential receptor or ligand molecules for the purpose of identifying receptor/ligand interactions. The identification of a ligand for a known receptor, a receptor for a known ligand or a novel receptor/ligand pair is useful for a variety of indications including, for example, targeting bioactive molecules (linked to the ligand or receptor) to a cell known to express the receptor or ligand, use of the receptor or ligand as a reagent to detect the presence of the ligand or receptor in a composition suspected of containing the same, wherein the composition may comprise cells suspected of expressing the ligand or receptor, modulating the growth of or another biological or immunological activity of a cell known to express or respond to the receptor or ligand, modulating the immune response of cells or toward cells that express the receptor or ligand, allowing the preparation of agonists, antagonists and/or antibodies directed against the receptor or ligand which will modulate the growth of or a biological or immunological activity of a cell expressing the receptor or ligand, and various other indications which will be readily apparent to the ordinarily skilled artisan.

[0532] The assay is performed as follows. A PRO polypeptide of the present invention suspected of being a ligand for a receptor is expressed as a fusion protein containing the Fc domain of human IgG (an immunoadhesin). Receptor-ligand binding is detected by allowing interaction of the immunoadhesin polypeptide with cells (e.g. Cos cells) expressing candidate PRO polypeptide receptors and visualization of bound immunoadhesin with fluorescent reagents directed toward the Fc fusion domain and examination by microscope. Cells expressing candidate receptors are produced by transient transfection, in parallel, of defined subsets of a library of cDNA expression vectors encoding PRO polypeptides that may function as receptor molecules. Cells are then incubated for 1 hour in the presence of the PRO polypeptide immunoadhesin being tested for possible receptor binding. The cells are then washed and fixed with

paraformaldehyde. The cells are then incubated with fluorescent conjugated antibody directed against the Fc portion of the PRO polypeptide immunoadhesin (e.g. FITC conjugated goat anti-human-Fc antibody). The cells are then washed again and examined by microscope. A positive interaction is judged by the presence of fluorescent labeling of cells transfected with cDNA encoding a particular PRO polypeptide receptor or pool of receptors and an absence of similar fluorescent labeling of similarly prepared cells that have been transfected with other cDNA or pools of cDNA. If a defined pool of cDNA expression vectors is judged to be positive for interaction with a PRO polypeptide immunoadhesin, the individual cDNA species that comprise the pool are tested individually (the pool is “broken down”) to determine the specific cDNA that encodes a receptor able to interact with the PRO polypeptide immunoadhesin.

[0533] In another embodiment of this assay, an epitope-tagged potential ligand PRO polypeptide (e.g. 8 histidine “His” tag) is allowed to interact with a panel of potential receptor PRO polypeptide molecules that have been expressed as fusions with the Fc domain of human IgG (immunoadhesins). Following a 1 hour co-incubation with the epitope tagged PRO polypeptide, the candidate receptors are each immunoprecipitated with protein A beads and the beads are washed. Potential ligand interaction is determined by western blot analysis of the immunoprecipitated complexes with antibody directed towards the epitope tag. An interaction is judged to occur if a band of the anticipated molecular weight of the epitope tagged protein is observed in the western blot analysis with a candidate receptor, but is not observed to occur with the other members of the panel of potential receptors.

[0534] Using these assays, the following receptor/ligand interactions have been herein identified:

[0535] (1) PRO10272 binds to PRO5801.

[0536] (2) PRO20110 binds to the human IL-17 receptor (Yao et al., *Cytokine* 9(11):794-800 (1997); also herein designated as PRO1) and to PRO20040.

[0537] (3) PRO10096 binds to PRO20233.

[0538] (4) PRO19670 binds to PRO1890.

[0539] The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention. The present invention is not to be limited in scope by the construct deposited, since the deposited embodiment is intended as a single illustration of certain aspects of the invention and any constructs that are functionally equivalent are within the scope of this invention. The deposit of material herein does not constitute an admission that the written description herein contained is inadequate to enable the practice of any aspect of the invention, including the best mode thereof, nor is it to be construed as limiting the scope of the claims to the specific illustrations that it represents. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and fall within the scope of the appended claims.

SEQUENCE LISTING	
<160> NUMBER OF SEQ ID NOS: 170	
<210> SEQ ID NO 1	
<211> LENGTH: 1173	
<212> TYPE: DNA	
<213> ORGANISM: Homo Sapien	
<400> SEQUENCE: 1	
ggggcttcgg cgccagcggc cagcgctagt cggctctggta aggatttaca	50
aaaggtgcag gtatgagcag gtctgaagac taacattttg tgaagttgta	100
aaacagaaaa cctgttagaa atgtgggtgt ttcagcaagg cctcagtttc	150
cttccttcag cccttgtaat ttggacatct gctgctttca tattttcata	200
cattactgca gtaacactcc accatataga cccggcttta ccttatatca	250
gtgacactgg tacagtagct ccagaaaaat gcttatttgg ggcaatgcta	300
aattattgcg cagttttatg cattgctacc atttatgttc gttataagca	350
agttcatgct ctgagtcctg aagagaacgt tatcatcaaa ttaaacaagg	400
ctggccctgt acttgaata ctgagttgtt taggactttc tattgtggca	450
aactccaga aaacaaccct ttttgctgca catgtaagtg gagctgtgct	500
tacctttggt atgggctcat tatatatgtt tgttcagacc atcctttcct	550
accaaataca gcccaaatc catggcaaac aagtcttctg gatcagactg	600
ttgttggtta tctgggtgtg agtaagtgca cttagcatgc tgacttgctc	650
atcagttttg cacagtggca attttgggac tgatttagaa cagaaaactcc	700
attggaaccc cgaggacaaa ggttatgtgc ttcacatgat cactactgca	750
gcgaatggt ctatgtcatt ttccttcttt ggttttttcc tgacttacat	800
tcgtgatatt cagaaaattt ctttacgggt ggaagccaat ttacatggat	850
taacctctca tgacactgca ccttgcctta ttaacaatga acgaacacgg	900
ctactttcca gagatatttg atgaaaggat aaaatatttc tgtaatgatt	950
atgattctca gggattgggg aaagggtcac agaagttgct tattcttctc	1000
tgaaattttc aaccacttaa tcaaggctga cagtaacact gatgaatgct	1050
gataatcagg aaacatgaaa gaagccattt gatagattat tctaaaggat	1100
atcatcaaga agactattaa aaacacctat gcctatactt ttttatctca	1150
gaaaataaag tcaaaagact atg	1173
<210> SEQ ID NO 2	
<211> LENGTH: 266	
<212> TYPE: PRT	
<213> ORGANISM: Homo Sapien	
<400> SEQUENCE: 2	
Met Trp Trp Phe Gln Gln Gly Leu Ser Phe Leu Pro Ser Ala Leu	
1 5 10 15	
Val Ile Trp Thr Ser Ala Ala Phe Ile Phe Ser Tyr Ile Thr Ala	
20 25 30	
Val Thr Leu His His Ile Asp Pro Ala Leu Pro Tyr Ile Ser Asp	
35 40 45	

-continued

Thr	Gly	Thr	Val	Ala	Pro	Glu	Lys	Cys	Leu	Phe	Gly	Ala	Met	Leu	
				50					55					60	
Asn	Ile	Ala	Ala	Val	Leu	Cys	Ile	Ala	Thr	Ile	Tyr	Val	Arg	Tyr	
				65					70					75	
Lys	Gln	Val	His	Ala	Leu	Ser	Pro	Glu	Glu	Asn	Val	Ile	Ile	Lys	
				80					85					90	
Leu	Asn	Lys	Ala	Gly	Leu	Val	Leu	Gly	Ile	Leu	Ser	Cys	Leu	Gly	
				95					100					105	
Leu	Ser	Ile	Val	Ala	Asn	Phe	Gln	Lys	Thr	Thr	Leu	Phe	Ala	Ala	
				110					115					120	
His	Val	Ser	Gly	Ala	Val	Leu	Thr	Phe	Gly	Met	Gly	Ser	Leu	Tyr	
				125					130					135	
Met	Phe	Val	Gln	Thr	Ile	Leu	Ser	Tyr	Gln	Met	Gln	Pro	Lys	Ile	
				140					145					150	
His	Gly	Lys	Gln	Val	Phe	Trp	Ile	Arg	Leu	Leu	Leu	Val	Ile	Trp	
				155					160					165	
Cys	Gly	Val	Ser	Ala	Leu	Ser	Met	Leu	Thr	Cys	Ser	Ser	Val	Leu	
				170					175					180	
His	Ser	Gly	Asn	Phe	Gly	Thr	Asp	Leu	Glu	Gln	Lys	Leu	His	Trp	
				185					190					195	
Asn	Pro	Glu	Asp	Lys	Gly	Tyr	Val	Leu	His	Met	Ile	Thr	Thr	Ala	
				200					205					210	
Ala	Glu	Trp	Ser	Met	Ser	Phe	Ser	Phe	Phe	Gly	Phe	Phe	Leu	Thr	
				215					220					225	
Tyr	Ile	Arg	Asp	Phe	Gln	Lys	Ile	Ser	Leu	Arg	Val	Glu	Ala	Asn	
				230					235					240	
Leu	His	Gly	Leu	Thr	Leu	Tyr	Asp	Thr	Ala	Pro	Cys	Pro	Ile	Asn	
				245					250					255	
Asn	Glu	Arg	Thr	Arg	Leu	Leu	Ser	Arg	Asp	Ile					
				260					265						

<210> SEQ ID NO 3
<211> LENGTH: 2037
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 3

cggacgcgtg	ggcggacgcg	tgggggagag	ccgcagtc	ggctgcagca	50
cctggggagaa	ggcagaccgt	gtgagggggc	ctgtggcccc	agcgtgctgt	100
ggcctcgggg	agtgggaagt	ggaggcagga	gccttcctta	cacttcgcca	150
tgagtttcct	catcgactcc	agcatcatga	ttacctccca	gatactat	200
tttgga	tttg	gctttt	cttcatgcgc	caattgttta	250
aagactatga					
gatacgtcag	tatgtgttac	aggtgatctt	ctccgtgacg	tttgcatttt	300
cttgaccat	gtttgagctc	atcatctttg	aaatcttagg	agtattgaat	350
agcagctccc	gttattttca	ctggaaaatg	aacctgtgtg	taattctgct	400
gatcctgggt	ttcatggtgc	ctttttacat	tggctat	ttt	450
attgtgagca					
atatccgact	actgcataaa	caacgactgc	ttttttcctg	tctcttatgg	500
ctgaccttta	tgtatttctt	ctggaaacta	ggagatccct	ttccattct	550
cagcccaaaa	catgggatct	tatccataga	acagctcatc	agccgggttg	600

-continued

gtgtgattgg agtgactctc atggctcttc tttctggatt tgggtctgtc	650
aactgcccat acacttacat gtcttacttc ctcaggaatg tgactgacac	700
ggatattcta gccctggaac ggcgactgct gcaaaccatg gatatgatca	750
taagcaaaaa gaaaaggatg gcaatggcac ggagaacaat gttccagaag	800
ggggaagtgc ataacaaacc atcaggtttc tggggaatga taaaaagtgt	850
taccacttca gcatcaggaa gtgaaaatct tactcttatt caacaggaag	900
tggatgcttt ggaagaatta agcaggcagc tttttctgga aacagctgat	950
ctatatgcta ccaaggagag aatagaatac tccaaaacct tcaaggggaa	1000
atattttaat tttcttggtt actttttctc tatttactgt gtttgaaaa	1050
ttttcatggc taccatcaat attgtttttg atcgagttgg gaaaacggat	1100
cctgtcacaa gaggcattga gatcactgtg aattatctgg gaatccaatt	1150
tgatgtgaag ttttggtooc aacacatttc cttcattctt gttggaataa	1200
tcacgtctac atccatcaga ggattgctga tcactcttac caagttcttt	1250
tatgccatct ctacgagtaa gtctctcaat gtcattgtcc tgctattagc	1300
acagataatg ggcatgtact ttgtctcctc tgtgctgctg atccgaatga	1350
gtatgccttt agaataccgc accataatca ctgaagtcct tggagaactg	1400
cagttcaact tctatcaccg ttggtttgat gtgatcttcc tggtcagcgc	1450
tctctctagc atactcttcc tctatttggc tcacaaacag gcaccagaga	1500
agcaaatggc accttgaact taagcctact acagactggt agaggccagt	1550
ggtttcaaaa tttagatata agagggggga aaaatggaac cagggcctga	1600
cattttataa acaaacaaaa tgctatggta gcatttttca ccttcatagc	1650
atactccttc cccgtcaggt gatactatga ccatgagtag catcagccag	1700
aacatgagag ggagaactaa ctcaagacaa tactcagcag agagcatccc	1750
gtgtggatat gaggctgggt tagaggcgga gaggagcaa gaaactaaag	1800
gtgaaaaata cactggaact ctggggcaag acatgtctat ggtagctgag	1850
ccaaacacgt aggatttccg ttttaaggtt cacatggaaa aggttatagc	1900
tttgccctga gattgactca ttaaaatcag agactgtaac aaaaaaaaaa	1950
aaaaaaaaaa agggcgggcg cgactctaga gtcgacctgc agaagcttgg	2000
cgcacatggc ccaacttggt tattgcagct tataatg	2037

<210> SEQ ID NO 4
<211> LENGTH: 455
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 4

Met Ser Phe Leu Ile Asp Ser Ser Ile Met Ile Thr Ser Gln Ile
1 5 10 15

Leu Phe Phe Gly Phe Gly Trp Leu Phe Phe Met Arg Gln Leu Phe
20 25 30

Lys Asp Tyr Glu Ile Arg Gln Tyr Val Val Gln Val Ile Phe Ser
35 40 45

-continued

Val	Thr	Phe	Ala	Phe	Ser	Cys	Thr	Met	Phe	Glu	Leu	Ile	Ile	Phe	50	55	60
Glu	Ile	Leu	Gly	Val	Leu	Asn	Ser	Ser	Ser	Arg	Tyr	Phe	His	Trp	65	70	75
Lys	Met	Asn	Leu	Cys	Val	Ile	Leu	Leu	Ile	Leu	Val	Phe	Met	Val	80	85	90
Pro	Phe	Tyr	Ile	Gly	Tyr	Phe	Ile	Val	Ser	Asn	Ile	Arg	Leu	Leu	95	100	105
His	Lys	Gln	Arg	Leu	Leu	Phe	Ser	Cys	Leu	Leu	Trp	Leu	Thr	Phe	110	115	120
Met	Tyr	Phe	Phe	Trp	Lys	Leu	Gly	Asp	Pro	Phe	Pro	Ile	Leu	Ser	125	130	135
Pro	Lys	His	Gly	Ile	Leu	Ser	Ile	Glu	Gln	Leu	Ile	Ser	Arg	Val	140	145	150
Gly	Val	Ile	Gly	Val	Thr	Leu	Met	Ala	Leu	Leu	Ser	Gly	Phe	Gly	155	160	165
Ala	Val	Asn	Cys	Pro	Tyr	Thr	Tyr	Met	Ser	Tyr	Phe	Leu	Arg	Asn	170	175	180
Val	Thr	Asp	Thr	Asp	Ile	Leu	Ala	Leu	Glu	Arg	Arg	Leu	Leu	Gln	185	190	195
Thr	Met	Asp	Met	Ile	Ile	Ser	Lys	Lys	Lys	Arg	Met	Ala	Met	Ala	200	205	210
Arg	Arg	Thr	Met	Phe	Gln	Lys	Gly	Glu	Val	His	Asn	Lys	Pro	Ser	215	220	225
Gly	Phe	Trp	Gly	Met	Ile	Lys	Ser	Val	Thr	Thr	Ser	Ala	Ser	Gly	230	235	240
Ser	Glu	Asn	Leu	Thr	Leu	Ile	Gln	Gln	Glu	Val	Asp	Ala	Leu	Glu	245	250	255
Glu	Leu	Ser	Arg	Gln	Leu	Phe	Leu	Glu	Thr	Ala	Asp	Leu	Tyr	Ala	260	265	270
Thr	Lys	Glu	Arg	Ile	Glu	Tyr	Ser	Lys	Thr	Phe	Lys	Gly	Lys	Tyr	275	280	285
Phe	Asn	Phe	Leu	Gly	Tyr	Phe	Phe	Ser	Ile	Tyr	Cys	Val	Trp	Lys	290	295	300
Ile	Phe	Met	Ala	Thr	Ile	Asn	Ile	Val	Phe	Asp	Arg	Val	Gly	Lys	305	310	315
Thr	Asp	Pro	Val	Thr	Arg	Gly	Ile	Glu	Ile	Thr	Val	Asn	Tyr	Leu	320	325	330
Gly	Ile	Gln	Phe	Asp	Val	Lys	Phe	Trp	Ser	Gln	His	Ile	Ser	Phe	335	340	345
Ile	Leu	Val	Gly	Ile	Ile	Ile	Val	Thr	Ser	Ile	Arg	Gly	Leu	Leu	350	355	360
Ile	Thr	Leu	Thr	Lys	Phe	Phe	Tyr	Ala	Ile	Ser	Ser	Ser	Lys	Ser	365	370	375
Ser	Asn	Val	Ile	Val	Leu	Leu	Leu	Ala	Gln	Ile	Met	Gly	Met	Tyr	380	385	390
Phe	Val	Ser	Ser	Val	Leu	Leu	Ile	Arg	Met	Ser	Met	Pro	Leu	Glu	395	400	405
Tyr	Arg	Thr	Ile	Ile	Thr	Glu	Val	Leu	Gly	Glu	Leu	Gln	Phe	Asn	410	415	420
Phe	Tyr	His	Arg	Trp	Phe	Asp	Val	Ile	Phe	Leu	Val	Ser	Ala	Leu			

-continued

[illegible]

-continued

caaggagtca ctgagaccaa ggctttctct actgattccg cagctcagac	1550
cctttcttca gctctgaaag agaaacacgt atcccacctg acatgtcctt	1600
ctgagcccgg taagagcaaa agaatggcag aaaagtttag ccctgaaaag	1650
ccatggagat tctcataact tgagacctaa tctctgtaaa gctaaaataa	1700
agaaatagaa caaggctgag gatacgacag tacactgtca gcagggactg	1750
taaacacaga cagggtcaaa gtgttttctc tgaacacatt gagttggaat	1800
cactgtttag aacacacaca cttacttttt ctggtctcta ccactgctga	1850
tattttctct aggaaatata cttttacaag taacaaaaat aaaaactctt	1900
ataaatttct atttttatct gagttacaga aatgattact aaggaagatt	1950
actcagtaat ttgtttaaaa agtaataaaa ttcaacaaac atttgctgaa	2000
tagctactat atgtcaagtg ctgtgcaagg tattacactc tgtaattgaa	2050
tattattcct caaaaaattg cacatagtag aacgctatct gggaagctat	2100
ttttttcagt ttgatatatt ctagcttata tacttccaaa ctaattttta	2150
tttttgctga gactaatctt attcattttc tctaatatgg caaccattat	2200
aaccttaatt tattattaac atacctaaga agtacattgt tacctctata	2250
taccaaagca cattttaaaa gtgccattaa caaatgtatc actagccctc	2300
ctttttccaa caagaaggga ctgagagatg cagaaatatt tgtgacaaaa	2350
aattaaagca tttagaaaac tt	2372

<210> SEQ ID NO 6
<211> LENGTH: 322
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 6

Met Ala Arg Cys Phe Ser Leu Val Leu Leu Leu Thr Ser Ile Trp	
1 5 10 15	
Thr Thr Arg Leu Leu Val Gln Gly Ser Leu Arg Ala Glu Glu Leu	
20 25 30	
Ser Ile Gln Val Ser Cys Arg Ile Met Gly Ile Thr Leu Val Ser	
35 40 45	
Lys Lys Ala Asn Gln Gln Leu Asn Phe Thr Glu Ala Lys Glu Ala	
50 55 60	
Cys Arg Leu Leu Gly Leu Ser Leu Ala Gly Lys Asp Gln Val Glu	
65 70 75	
Thr Ala Leu Lys Ala Ser Phe Glu Thr Cys Ser Tyr Gly Trp Val	
80 85 90	
Gly Asp Gly Phe Val Val Ile Ser Arg Ile Ser Pro Asn Pro Lys	
95 100 105	
Cys Gly Lys Asn Gly Val Gly Val Leu Ile Trp Lys Val Pro Val	
110 115 120	
Ser Arg Gln Phe Ala Ala Tyr Cys Tyr Asn Ser Ser Asp Thr Trp	
125 130 135	
Thr Asn Ser Cys Ile Pro Glu Ile Ile Thr Thr Lys Asp Pro Ile	
140 145 150	
Phe Asn Thr Gln Thr Ala Thr Gln Thr Thr Glu Phe Ile Val Ser	
155 160 165	

-continued

Asp	Ser	Thr	Tyr	Ser	Val	Ala	Ser	Pro	Tyr	Ser	Thr	Ile	Pro	Ala	
				170					175					180	
Pro	Thr	Thr	Thr	Pro	Pro	Ala	Pro	Ala	Ser	Thr	Ser	Ile	Pro	Arg	
				185					190					195	
Arg	Lys	Lys	Leu	Ile	Cys	Val	Thr	Glu	Val	Phe	Met	Glu	Thr	Ser	
				200					205					210	
Thr	Met	Ser	Thr	Glu	Thr	Glu	Pro	Phe	Val	Glu	Asn	Lys	Ala	Ala	
				215					220					225	
Phe	Lys	Asn	Glu	Ala	Ala	Gly	Phe	Gly	Gly	Val	Pro	Thr	Ala	Leu	
				230					235					240	
Leu	Val	Leu	Ala	Leu	Leu	Phe	Phe	Gly	Ala	Ala	Ala	Gly	Leu	Gly	
				245					250					255	
Phe	Cys	Tyr	Val	Lys	Arg	Tyr	Val	Lys	Ala	Phe	Pro	Phe	Thr	Asn	
				260					265					270	
Lys	Asn	Gln	Gln	Lys	Glu	Met	Ile	Glu	Thr	Lys	Val	Val	Lys	Glu	
				275					280					285	
Glu	Lys	Ala	Asn	Asp	Ser	Asn	Pro	Asn	Glu	Glu	Ser	Lys	Lys	Thr	
				290					295					300	
Asp	Lys	Asn	Pro	Glu	Glu	Ser	Lys	Ser	Pro	Ser	Lys	Thr	Thr	Val	
				305					310					315	
Arg	Cys	Leu	Glu	Ala	Glu	Val									
				320											

<210> SEQ ID NO 7
<211> LENGTH: 2586
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 7

cgccgcgctc ccgcacccgc ggcccgccca ccgcgccgct ccgcacatctg	50
caccgcgcgc ccggcgccct cccggcgggg gcgagcagat ccagtcgggc	100
ccgcagcgca actcgggtcca gtcggggcgg cggctgcggg cgcagagcgg	150
agatgcagcg gcttggggcc accctgctgt gcctgctgct ggcggcgggc	200
gtccccacgg cccccgcgcc cgctccgacg gcgacctcgg ctccagtcac	250
gccccggccc gctctcagct acccgcagga ggaggccacc ctcaatgaga	300
tggtccgcga ggttgaggaa ctgatggagg acacgcagca caaattgcgc	350
agcgcggtgg aagagatgga ggcagaagaa gctgctgcta aagcatcatc	400
agaagtgaac ctggcaaaact tacctcccag ctatcacaaat gagaccaaca	450
cagacacgaa ggttggaat aataccatcc atgtgcaccg agaaattcac	500
aagataacca acaaccagac tggacaaatg gtcttttcag agacagttat	550
cacatctgtg ggagacgaag aaggcagaag gagccacgag tgcacatcgc	600
acgaggactg tgggcccgag atgtactgcc agtttgccag cttccagtac	650
acctgccagc catgccgggg ccagaggatg ctctgcaccc gggacagtga	700
gtgctgtgga gaccagctgt gtgtctgggg tcaactgcacc aaaatggcca	750
ccagggcgag caatgggacc atctgtgaca accagaggga ctgccagccg	800
gggctgtgct gtgccttcca gagaggcctg ctgttcctg tgtgcacacc	850

-continued

cctgcccgtg gagggcgagc ttgccaatga ccccgccagc cggcttcttg	900
acctcatcac ctgggagcta gagcctgatg gagccttgga ccgatgccct	950
tggtgccagt gcctcctctg ccagcccccac agccacagcc tgggtgatgt	1000
gtgcaagccg accttcgttg ggagccgtga ccaagatggg gagatcctgc	1050
tgcccagaga ggtcccgat gagtatgaag ttggcagctt catggaggag	1100
gtgcgccagg agctggagga cctggagagg agcctgactg aagagatggc	1150
gctgggggag cctgcggctg ccgccctgc actgctggga ggggaagaga	1200
tttagatctg gaccaggctg tgggtagatg tgcaatagaa atagctaatt	1250
tatttcccca ggtgtgtgct ttaggcgttg gctgaccagg cttcttctta	1300
catcttcttc ccagtaagtt tcccctctg cttgacagca tgagggtgtg	1350
tgcatattgt cagctcccc aggctgttct ccaggcttca cagtctggtg	1400
cttgggagag tcaggcaggg ttaaactgca ggagcagttt gccaccctg	1450
tccagattat tggctgcttt gcctctacca gttggcagac agcgtttgt	1500
tctacatggc ttgataaatt gtttgagggg aggagatgga aacaatgttg	1550
agtctccctc tgattggttt tggggaaatg tggagaagag tgccctgctt	1600
tgcaaacatc aacctggcaa aaatgcaaca aatgaatttt ccacgcagtt	1650
ctttccatgg gcataggtaa gctgtgcctt cagctgttgc agatgaaatg	1700
ttctgttca cctgcattac atgtgtttat tcatccagca gtgttgctca	1750
gctcctacct ctgtgccagg gcagcatttt catatccaag atcaattccc	1800
tctctcagca cagcctgggg agggggtcat tgttctcctc gtccatcagg	1850
gatctcagag gctcagagac tgcaagctgc ttgccaagt cacacagcta	1900
gtgaagacca gagcagtttc atctggttgt gactctaagc tcagtgtctt	1950
ctccactacc ccacaccagc cttggtgcca ccaaaagtgc tccccaaaag	2000
gaaggagaat gggatttttc ttgaggcatg cacatctgga attaaggtca	2050
aactaattct cacatccctc taaaagttaa ctactgttag gaacagcagt	2100
gttctcacag tgtggggcag ccgtccttct aatgaagaca atgatattga	2150
cactgtccct ctttggcagt tgcattagta actttgaaag gtatatgact	2200
gagcgtagca tacagggtta cctgcagaaa cagtacttag gtaattgtag	2250
ggcgaggatt ataaatgaaa ttgcaaaat cacttagcag caactgaaga	2300
caattatcaa ccacgtggag aaaatcaaac cgagcagggc tgtgtgaaac	2350
atggttgtaa tatgcgactg cgaacactga actctacgcc actccacaaa	2400
tgatgttttc aggtgtcatg gactgttgcc accatgtatt catccagagt	2450
tcttaaagtt taaagttgca catgattgta taagcatgct ttctttgagt	2500
tttaaatatg gtataaacat aagttgcatt tagaaatcaa gcataaatca	2550
cttcaactgc aaaaaaaaaa aaaaaaaaaa aaaaaa	2586

<210> SEQ ID NO 8

<211> LENGTH: 350

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 8

```

Met Gln Arg Leu Gly Ala Thr Leu Leu Cys Leu Leu Leu Ala Ala
 1          5          10          15

Ala Val Pro Thr Ala Pro Ala Pro Ala Pro Thr Ala Thr Ser Ala
20          25          30

Pro Val Lys Pro Gly Pro Ala Leu Ser Tyr Pro Gln Glu Glu Ala
35          40          45

Thr Leu Asn Glu Met Phe Arg Glu Val Glu Glu Leu Met Glu Asp
50          55          60

Thr Gln His Lys Leu Arg Ser Ala Val Glu Glu Met Glu Ala Glu
65          70          75

Glu Ala Ala Ala Lys Ala Ser Ser Glu Val Asn Leu Ala Asn Leu
80          85          90

Pro Pro Ser Tyr His Asn Glu Thr Asn Thr Asp Thr Lys Val Gly
95          100         105

Asn Asn Thr Ile His Val His Arg Glu Ile His Lys Ile Thr Asn
110         115         120

Asn Gln Thr Gly Gln Met Val Phe Ser Glu Thr Val Ile Thr Ser
125         130         135

Val Gly Asp Glu Glu Gly Arg Arg Ser His Glu Cys Ile Ile Asp
140         145         150

Glu Asp Cys Gly Pro Ser Met Tyr Cys Gln Phe Ala Ser Phe Gln
155         160         165

Tyr Thr Cys Gln Pro Cys Arg Gly Gln Arg Met Leu Cys Thr Arg
170         175         180

Asp Ser Glu Cys Cys Gly Asp Gln Leu Cys Val Trp Gly His Cys
185         190         195

Thr Lys Met Ala Thr Arg Gly Ser Asn Gly Thr Ile Cys Asp Asn
200         205         210

Gln Arg Asp Cys Gln Pro Gly Leu Cys Cys Ala Phe Gln Arg Gly
215         220         225

Leu Leu Phe Pro Val Cys Thr Pro Leu Pro Val Glu Gly Glu Leu
230         235         240

Cys His Asp Pro Ala Ser Arg Leu Leu Asp Leu Ile Thr Trp Glu
245         250         255

Leu Glu Pro Asp Gly Ala Leu Asp Arg Cys Pro Cys Ala Ser Gly
260         265         270

Leu Leu Cys Gln Pro His Ser His Ser Leu Val Tyr Val Cys Lys
275         280         285

Pro Thr Phe Val Gly Ser Arg Asp Gln Asp Gly Glu Ile Leu Leu
290         295         300

Pro Arg Glu Val Pro Asp Glu Tyr Glu Val Gly Ser Phe Met Glu
305         310         315

Glu Val Arg Gln Glu Leu Glu Asp Leu Glu Arg Ser Leu Thr Glu
320         325         330

Glu Met Ala Leu Gly Glu Pro Ala Ala Ala Ala Ala Ala Leu Leu
335         340         345

Gly Gly Glu Glu Ile
350

```

<210> SEQ ID NO 9

-continued

<211> LENGTH: 1395
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 9

cggacgcgtg ggcggacgcg tgggggctgt gagaaagtgc caataaatac 50
atcatgcaac cccacggccc accttgtgaa ctctctgtgc ccagggctga 100
tgtgcgtctt ccagggtac tcatccaaag gcctaatacca acgttctgtc 150
ttcaatctgc aaatctatgg ggtcctgggg ctcttctgga cccttaactg 200
ggtactggcc ctggggccaat gcgtcctcgc tggagccttt gcctccttct 250
actgggcctt ccacaagccc caggacatcc ctaccttccc cttaatctct 300
gccttcaccc gcacactccg ttaccacact gggtcattgg catttgagc 350
cctcatcctg acccttgtgc agatagcccc ggtcatcttg gagtatattg 400
accacaagct cagaggagtg cagaaccctg tagcccgctg catcatgtgc 450
tgtttcaagt gctgcctctg gtgtctggaa aaatttatca agttcctaaa 500
ccgcaatgca tacatcatga tcgccatcta cgggaagaat ttctgtgtct 550
cagccaaaaa tgcgttcatg ctactcatgc gaaacattgt cagggtggtc 600
gtcctggaca aagtcacaga cctgctgctg ttctttggga agctgctggt 650
ggtcggaggc gtgggggtcc tgtccttctt ttttttctcc ggtcgcatcc 700
cggggctggg taaagacttt aagagcccc acctcaacta ttactggctg 750
cccacatga cctccatcct gggggcctat gtcatcgcca gcggcttctt 800
cagcgttttc ggcatgtgtg tggacacgct ctctctctgc ttcttggaag 850
acctggagcg gaacaacggc tccctggacc ggccctaacta catgtccaag 900
agccttctaa agattctggg caagaagaac gaggcgcccc cggacaacaa 950
gaagagggaag aagtgcacgc tccggccctg atccaggact gcacccacc 1000
cccaccgtcc agccatccaa cctcacttgc ccttacaggt ctccattttg 1050
tggtaaaaaa aggttttagg ccaggcgccg tggctcacgc ctgtaatacca 1100
acactttgag aggctgaggc gggcgatca cctgagtcag gagttcgaga 1150
ccagcctggc caacatggtg aaacctccgt ctctattaaa aatacaaaaa 1200
ttagccgaga gtggtggcat gcacctgtca tcccagctac tcgggagggt 1250
gaggcaggag aatcgcttga acccgggagg cagaggttgc agtgagccga 1300
gatcgcgcca ctgcactcca acctgggtga cagactctgt ctccaaaaca 1350
aaacaaacaa acaaaaagat ttattataag atattttgtt aactc 1395

<210> SEQ ID NO 10
<211> LENGTH: 321
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 10

Arg Thr Arg Gly Arg Thr Arg Gly Gly Cys Glu Lys Val Pro Ile
1 5 10 15

Asn Thr Ser Cys Asn Pro Thr Ala His Leu Val Asn Ser Ser Cys
20 25 30

-continued

Pro	Gly	Leu	Met	Cys	Val	Phe	Gln	Gly	Tyr	Ser	Ser	Lys	Gly	Leu
			35						40					45
Ile	Gln	Arg	Ser	Val	Phe	Asn	Leu	Gln	Ile	Tyr	Gly	Val	Leu	Gly
			50						55					60
Leu	Phe	Trp	Thr	Leu	Asn	Trp	Val	Leu	Ala	Leu	Gly	Gln	Cys	Val
			65						70					75
Leu	Ala	Gly	Ala	Phe	Ala	Ser	Phe	Tyr	Trp	Ala	Phe	His	Lys	Pro
			80						85					90
Gln	Asp	Ile	Pro	Thr	Phe	Pro	Leu	Ile	Ser	Ala	Phe	Ile	Arg	Thr
			95						100					105
Leu	Arg	Tyr	His	Thr	Gly	Ser	Leu	Ala	Phe	Gly	Ala	Leu	Ile	Leu
			110						115					120
Thr	Leu	Val	Gln	Ile	Ala	Arg	Val	Ile	Leu	Glu	Tyr	Ile	Asp	His
			125						130					135
Lys	Leu	Arg	Gly	Val	Gln	Asn	Pro	Val	Ala	Arg	Cys	Ile	Met	Cys
			140						145					150
Cys	Phe	Lys	Cys	Cys	Leu	Trp	Cys	Leu	Glu	Lys	Phe	Ile	Lys	Phe
			155						160					165
Leu	Asn	Arg	Asn	Ala	Tyr	Ile	Met	Ile	Ala	Ile	Tyr	Gly	Lys	Asn
			170						175					180
Phe	Cys	Val	Ser	Ala	Lys	Asn	Ala	Phe	Met	Leu	Leu	Met	Arg	Asn
			185						190					195
Ile	Val	Arg	Val	Val	Val	Leu	Asp	Lys	Val	Thr	Asp	Leu	Leu	Leu
			200						205					210
Phe	Phe	Gly	Lys	Leu	Leu	Val	Val	Gly	Gly	Val	Gly	Val	Leu	Ser
			215						220					225
Phe	Phe	Phe	Phe	Ser	Gly	Arg	Ile	Pro	Gly	Leu	Gly	Lys	Asp	Phe
			230						235					240
Lys	Ser	Pro	His	Leu	Asn	Tyr	Tyr	Trp	Leu	Pro	Ile	Met	Thr	Ser
			245						250					255
Ile	Leu	Gly	Ala	Tyr	Val	Ile	Ala	Ser	Gly	Phe	Phe	Ser	Val	Phe
			260						265					270
Gly	Met	Cys	Val	Asp	Thr	Leu	Phe	Leu	Cys	Phe	Leu	Glu	Asp	Leu
			275						280					285
Glu	Arg	Asn	Asn	Gly	Ser	Leu	Asp	Arg	Pro	Tyr	Tyr	Met	Ser	Lys
			290						295					300
Ser	Leu	Leu	Lys	Ile	Leu	Gly	Lys	Lys	Asn	Glu	Ala	Pro	Pro	Asp
			305						310					315
Asn	Lys	Lys	Arg	Lys	Lys									
			320											

<210> SEQ ID NO 11
<211> LENGTH: 1901
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 11

gccccgcgcc	cgcgccggg	cgccgaagc	cgggagccac	cgccatggg	50
gcctgcctgg	gagcctgctc	cctgctcagc	tgcgcgtcct	gcctctgcgg	100
ctctgcccc	tgcatcctgt	gcagctgctg	ccccgccagc	cgcaactcca	150
cggtgagcgc	cctcatcttc	acgtttcttc	tcttctctgg	ggtgctgggtg	200

-continued

tccatcatta tgctgagccc gggcgtggag agtcagctct acaagctgcc	250
ctgggtgtgt gaggaggggg cgggatccc caccgtcctg cagggccaca	300
tcgactgtgg ctccctgctt ggctaccgag ctgtctaccg catgtgcttc	350
gccacggcgg ccttcttctt cttcttttct accctgctca tgctctgcgt	400
gagcagcagc cgggaccccc gggctgcat ccagaatggg ttttggttct	450
ttaagtctct gatcctgggt ggctcaccg tgggtgcctt ctacatccct	500
gacggctcct tcaccaacat ctggtttctac ttctggctcg tgggtcctt	550
cctcttctac ctcatccagc tgggtgctgt catcgacttt gcgcactcct	600
ggaaccacgg gtggctgggc aaggccgagg agtgcgattc ccgtgcctgg	650
tacgcaggcc tcttcttctt cactctcctc ttctacttgc tgcgatcgc	700
ggccgtggcg ctgatgttca tgtactacac tgagccagc ggcgtccacg	750
agggcaaggt cttcatcagc ctcaacctca cctctgtgt ctgcgtgtcc	800
atcgctgtgt tcctgcccac ggtccaggac gccagccca actcgggtct	850
gctgcaggcc tcggtcatca ccctctacac catgtttgtc acctggtcag	900
ccctatccag tatccctgaa cagaaatgca accccattt gccaacccag	950
ctgggcaacg agacagtgtt ggcaggcccc gagggctatg agaccagtg	1000
gtgggatgcc ccgagcattg tgggcctcat catcttctc ctgtgcaccc	1050
tcttcatcag tctgcgtccc tcagaccacc ggcaggtgaa cagcctgatg	1100
cagaccgagg agtgccacc tatgctagac gccacacagc agcagcagca	1150
gcaggtggca gcctgtgagg gccgggcctt tgacaacgag caggacggcg	1200
tcacctacag ctactccttc ttccacttct gcctgggtgt ggctcactg	1250
cacgtcatga tgacgtctac caactggtag aagcccggtg agaccggaa	1300
gatgatcagc acgtggaccg ccgtgtgggt gaagatctgt gccagctggg	1350
cagggctgct cctctacctg tggaccctgg tagccccact cctcctgcgc	1400
aaccgcgact tcagctgagg cagcctcaca gcctgccatc tggtgccctc	1450
tgccacctgg tgctctctcg ctcggtgaca gccaacctgc cccctcccca	1500
caccaatcag ccaggctgag cccccacccc tgccccagct ccaggacctg	1550
cccctgagcc gggccttcta gtctagtgc cttcagggtc cgaggagcat	1600
caggctcctg cagagcccca tccccccgcc acaccacac ggtggagctg	1650
cctcttctct cccctcctcc ctggtgccca tactcagcat ctcgatgaa	1700
agggctcctt tgcctcagg ctccacggga gcggggctgc tggagagagc	1750
ggggaactcc caccacagtg gggcatccgg cactgaagcc ctgggtgttc	1800
tggtcacgtc cccaggggga ccctgcccc ttctggact tcgtgcctta	1850
ctgagtctct aagacttttt ctaataaaca agccagtgcg tgtaaaaaaa	1900
a	1901

<210> SEQ ID NO 12

<211> LENGTH: 457

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 12

Met	Gly	Ala	Cys	Leu	Gly	Ala	Cys	Ser	Leu	Leu	Ser	Cys	Ala	Ser	1	5	10	15
Cys	Leu	Cys	Gly	Ser	Ala	Pro	Cys	Ile	Leu	Cys	Ser	Cys	Cys	Pro	20	25	30	
Ala	Ser	Arg	Asn	Ser	Thr	Val	Ser	Arg	Leu	Ile	Phe	Thr	Phe	Phe	35	40	45	
Leu	Phe	Leu	Gly	Val	Leu	Val	Ser	Ile	Ile	Met	Leu	Ser	Pro	Gly	50	55	60	
Val	Glu	Ser	Gln	Leu	Tyr	Lys	Leu	Pro	Trp	Val	Cys	Glu	Glu	Gly	65	70	75	
Ala	Gly	Ile	Pro	Thr	Val	Leu	Gln	Gly	His	Ile	Asp	Cys	Gly	Ser	80	85	90	
Leu	Leu	Gly	Tyr	Arg	Ala	Val	Tyr	Arg	Met	Cys	Phe	Ala	Thr	Ala	95	100	105	
Ala	Phe	Phe	Phe	Phe	Phe	Phe	Thr	Leu	Leu	Met	Leu	Cys	Val	Ser	110	115	120	
Ser	Ser	Arg	Asp	Pro	Arg	Ala	Ala	Ile	Gln	Asn	Gly	Phe	Trp	Phe	125	130	135	
Phe	Lys	Phe	Leu	Ile	Leu	Val	Gly	Leu	Thr	Val	Gly	Ala	Phe	Tyr	140	145	150	
Ile	Pro	Asp	Gly	Ser	Phe	Thr	Asn	Ile	Trp	Phe	Tyr	Phe	Gly	Val	155	160	165	
Val	Gly	Ser	Phe	Leu	Phe	Ile	Leu	Ile	Gln	Leu	Val	Leu	Leu	Ile	170	175	180	
Asp	Phe	Ala	His	Ser	Trp	Asn	Gln	Arg	Trp	Leu	Gly	Lys	Ala	Glu	185	190	195	
Glu	Cys	Asp	Ser	Arg	Ala	Trp	Tyr	Ala	Gly	Leu	Phe	Phe	Phe	Thr	200	205	210	
Leu	Leu	Phe	Tyr	Leu	Leu	Ser	Ile	Ala	Ala	Val	Ala	Leu	Met	Phe	215	220	225	
Met	Tyr	Tyr	Thr	Glu	Pro	Ser	Gly	Cys	His	Glu	Gly	Lys	Val	Phe	230	235	240	
Ile	Ser	Leu	Asn	Leu	Thr	Phe	Cys	Val	Cys	Val	Ser	Ile	Ala	Ala	245	250	255	
Val	Leu	Pro	Lys	Val	Gln	Asp	Ala	Gln	Pro	Asn	Ser	Gly	Leu	Leu	260	265	270	
Gln	Ala	Ser	Val	Ile	Thr	Leu	Tyr	Thr	Met	Phe	Val	Thr	Trp	Ser	275	280	285	
Ala	Leu	Ser	Ser	Ile	Pro	Glu	Gln	Lys	Cys	Asn	Pro	His	Leu	Pro	290	295	300	
Thr	Gln	Leu	Gly	Asn	Glu	Thr	Val	Val	Ala	Gly	Pro	Glu	Gly	Tyr	305	310	315	
Glu	Thr	Gln	Trp	Trp	Asp	Ala	Pro	Ser	Ile	Val	Gly	Leu	Ile	Ile	320	325	330	
Phe	Leu	Leu	Cys	Thr	Leu	Phe	Ile	Ser	Leu	Arg	Ser	Ser	Asp	His	335	340	345	
Arg	Gln	Val	Asn	Ser	Leu	Met	Gln	Thr	Glu	Glu	Cys	Pro	Pro	Met	350	355	360	
Leu	Asp	Ala	Thr	Gln	Gln	Gln	Gln	Gln	Gln	Val	Ala	Ala	Cys	Glu	365	370	375	

-continued

Gly	Arg	Ala	Phe	Asp	Asn	Glu	Gln	Asp	Gly	Val	Thr	Tyr	Ser	Tyr
380									385					390
Ser	Phe	Phe	His	Phe	Cys	Leu	Val	Leu	Ala	Ser	Leu	His	Val	Met
395									400					405
Met	Thr	Leu	Thr	Asn	Trp	Tyr	Lys	Pro	Gly	Glu	Thr	Arg	Lys	Met
410									415					420
Ile	Ser	Thr	Trp	Thr	Ala	Val	Trp	Val	Lys	Ile	Cys	Ala	Ser	Trp
425									430					435
Ala	Gly	Leu	Leu	Leu	Tyr	Leu	Trp	Thr	Leu	Val	Ala	Pro	Leu	Leu
440									445					450
Leu	Arg	Asn	Arg	Asp	Phe	Ser								
455														

<210> SEQ ID NO 13
<211> LENGTH: 1572
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 13

cgggccagcc	tggggcggcc	ggccaggaac	caccggttaa	ggtgtcttct	50
ctttagggat	ggtgaggttg	gaaaagact	cctgtaacc	tcctccagga	100
tgaaccacct	gccagaagac	atggagaacg	ctctcaccgg	gagccagagc	150
tcccatgctt	ctctgcgcaa	tatccattcc	atcaaccca	cacaactcat	200
ggccaggatt	gagtcctatg	aaggaaggga	aaagaaaggc	atatctgatg	250
tcaggaggac	tttctgtttg	tttgtcacct	ttgacctott	attcgtaaca	300
ttactgtgga	taatagagtt	aaatgtgaat	ggaggcattg	agaacacatt	350
agagaaggag	gtgatgcagt	atgactacta	ttcttcatat	ttgatatat	400
ttcttctggc	agtttttcga	tttaaagtgt	taatacttgc	atatgctgtg	450
tgcagactgc	gccattgggtg	ggcaatagcg	ttgacaacgg	cagtgaccag	500
tgccctttta	ctagcaaaaag	tgatcccttc	gaagcttttc	tctcaagggg	550
cttttggtta	tgtgctgccc	atcatttcat	tcattccttg	ctggattgag	600
acgtgggtcc	tggatttcaa	agtgttacct	caagaagcag	aagaagaaaa	650
cagactcctg	atagttcagg	atgcttcaga	gagggcagca	cttataacctg	700
gtggctcttc	tgatggtcag	ttttattccc	ctcctgaatc	cgaagcagga	750
tctgaagaag	ctgaagaaaa	acaggacagt	gagaaaccac	ttttagaact	800
atgagtacta	cttttggtta	atgtgaaaaa	ccctcacaga	aagtcatcga	850
ggcaaaaaga	ggcaggcagt	ggagtctccc	tgtcgacagt	aaagttgaaa	900
tggtgacgtc	cactgctggc	tttattgaac	agctaataaa	gattttattd	950
ttgtaatacc	tcacaaacgt	tgtaccatat	ccatgcacat	ttagttgcct	1000
gcctgtggct	ggtaaggtaa	tgtcatgatt	catcctctct	tcagtgagac	1050
tgagcctgat	gtgttaacaa	atagggtgaag	aaagtcttgt	gctgtattcc	1100
taatcaaaag	acttaatat	ttgaagtaac	acttttttag	taagcaagat	1150
acctttttat	ttcaattcac	agaatggaat	ttttttgttt	catgtctcag	1200
atattttttg	tatttctttt	ttaacactct	acatttcctt	tgttttttaa	1250

-continued

```
ctcatgcaca tgtgctcttt gtacagtttt aaaaagtgtg ataaaatctg      1300
acatgtcaat gtggctagtt ttatttttct tgttttgcac tatgtgtatg      1350
gcctgaagtg ttggacttgc aaaaggggaa gaaaggaatt gcgaatacat      1400
gtaaaatgtc accagacatt tgtattattt ttatcatgaa atcatgtttt      1450
tctctgattg ttctgaaatg ttctaaatac tcttattttg aatgcacaaa      1500
atgacttaaa ccattcatat catgtttcct ttgcgttcag ccaatttcaa      1550
ttaaaatgaa ctaaattaaa aa                                     1572
```

<210> SEQ ID NO 14
<211> LENGTH: 234
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 14

```
Met Asn His Leu Pro Glu Asp Met Glu Asn Ala Leu Thr Gly Ser
 1             5             10             15
Gln Ser Ser His Ala Ser Leu Arg Asn Ile His Ser Ile Asn Pro
          20             25             30
Thr Gln Leu Met Ala Arg Ile Glu Ser Tyr Glu Gly Arg Glu Lys
          35             40             45
Lys Gly Ile Ser Asp Val Arg Arg Thr Phe Cys Leu Phe Val Thr
          50             55             60
Phe Asp Leu Leu Phe Val Thr Leu Leu Trp Ile Ile Glu Leu Asn
          65             70             75
Val Asn Gly Gly Ile Glu Asn Thr Leu Glu Lys Glu Val Met Gln
          80             85             90
Tyr Asp Tyr Tyr Ser Ser Tyr Phe Asp Ile Phe Leu Leu Ala Val
          95             100            105
Phe Arg Phe Lys Val Leu Ile Leu Ala Tyr Ala Val Cys Arg Leu
          110            115            120
Arg His Trp Trp Ala Ile Ala Leu Thr Thr Ala Val Thr Ser Ala
          125            130            135
Phe Leu Leu Ala Lys Val Ile Leu Ser Lys Leu Phe Ser Gln Gly
          140            145            150
Ala Phe Gly Tyr Val Leu Pro Ile Ile Ser Phe Ile Leu Ala Trp
          155            160            165
Ile Glu Thr Trp Phe Leu Asp Phe Lys Val Leu Pro Gln Glu Ala
          170            175            180
Glu Glu Glu Asn Arg Leu Leu Ile Val Gln Asp Ala Ser Glu Arg
          185            190            195
Ala Ala Leu Ile Pro Gly Gly Leu Ser Asp Gly Gln Phe Tyr Ser
          200            205            210
Pro Pro Glu Ser Glu Ala Gly Ser Glu Glu Ala Glu Glu Lys Gln
          215            220            225
Asp Ser Glu Lys Pro Leu Leu Glu Leu
          230
```

<210> SEQ ID NO 15
<211> LENGTH: 2768
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 15

actcgaacgc agttgcttcg ggacccagga cccctcggg cccgaccgc	50
caggaaagac tgaggccgcg gcctgccccg cccggctccc tgcgcgcgcg	100
ccgcctcccg ggacagaaga tgtgtcccag ggtccctctg ctgctgccgc	150
tgtccttct actggccctg gggcctgggg tgcagggctg cccatccggc	200
tgccagtgc gccagccaca gacagtcttc tgcactgccc gccaggggac	250
cacggtgcc cgagacgtgc cccccgacac ggtggggctg tacgtctttg	300
agaacggcat caccatgctc gacgcaggca gctttgccgg cctgccgggc	350
ctgcagctcc tggacctgtc acagaaccag atcgccagcc tgcccagcgg	400
ggtcttcag cactcgcca acctcagcaa cctggacctg acggccaaca	450
ggctgcata aatcaccaat gagaccttc gtggcctgcg gcgcctcgag	500
cgctcttacc tgggcaagaa ccgcatccgc cacatccagc ctggtgcctt	550
cgacacgtc gaccgcctcc tggagctcaa gctgcaggac aacgagctgc	600
gggactgcc cccgctgcgc ctgccccgc tgtgtctgct ggacctcagc	650
cacaacagcc tcctggccct ggagcccgcc atcctggaca ctgccaacgt	700
ggagcgctg cggctggctg gtctggggct gcagcagctg gacgaggggc	750
tcctcagcgc cttgcgcaac ctccacgacc tggatgtgtc cgacaaccag	800
ctggagcgag tgccacctgt gatccgaggc ctccggggcc tgacgcgcct	850
gcggctggcc ggcaacaccc gattgccca gctgcggccc gaggacctgg	900
ccggcctggc tgccctgcag gagctggatg tagcaacct aagcctgcag	950
gccctgcctg gcgacctctc gggcctcttc cccgcctgc ggtgctggc	1000
agctgccccg aacccttca actgcgtgtg cccctgagc tggtttgccc	1050
cctgggtgcg cgagagccac gtcacactgg ccagccctga ggagacgcgc	1100
tgccacttcc cgccaagaa cgctggcccg ctgctcctgg agcttgacta	1150
cgccgacttt ggtgcccag ccaccaccac cacagccaca gtgccacca	1200
cgaggcccggt ggtgcgggag cccacagcct tgtcttctag cttggtcct	1250
acctggctta gcccacagc gccggccact gagccccca gccgccttc	1300
cactgcccga ccgactgtag ggctgtccc ccagccccag gactgcccac	1350
cgctccacct cctcaatggg ggcacatgcc acctggggac acggcaccac	1400
ctggcgctgt tgtgccccga aggccttcag ggctgtact gtgagagcca	1450
gatggggcag gggacacggc ccagccctac accagtacg ccgaggccac	1500
cacggtccct gaccctgggc atcgagccgg taggccccac ctccctgcgc	1550
gtggggctgc agcgctacct ccaggggagc tccgtgcagc tcaggagcct	1600
ccgtctcacc tatcgcaacc tatcgggccc tgataagcgg ctggtgacgc	1650
tgcgactgcc tgcctcgctc gctgagtaca cggtcaccca gctgcggccc	1700
aacgccactt actccgtctg tgtatgcct ttggggcccg ggcgggtgcc	1750
ggagggcgag gaggcctgcg gggaggccca tacaccccca gccgtccact	1800
ccaaccacgc ccagtcacc caggcccgcg agggcaacct gccgtcctc	1850

-continued

```
attgcgccccg ccctggccgc ggtgctcctg gccgcgctgg ctgcggtggg      1900
ggcagccctac tgtgtgcggc gggggcgggc catggcagca gcggctcagg      1950
acaaagggca ggtggggcca ggggctgggc ccctggaact ggagggagtg      2000
aaggccccct tggagccagg cccgaaggca acagaggcg gtggagaggc      2050
cctgcccagc gggctctgagt gtgaggtgcc actcatgggc tccccagggc      2100
ctggcctcca gtcaccctc caccgaaagc cctacatcta agccagagag      2150
agacagggca gctggggccg ggctctcagc cagtgagatg gccagccccc      2200
tcctctgtgc acaccacgta agttctcagt cccaacctcg gggatgtgtg      2250
cagacagggc tgtgtgacca cagctgggccc ctgttccctc tggacctcgg      2300
tctcctcatc tgtgagatgc tgtggcccag ctgacgagcc ctaacgtccc      2350
cagaaccgag tgcctatgag gacagtgtcc gccctgccct ccgcaacgtg      2400
cagtcctctg gcacggcggg ccctgccatg tgctggtaac gcatgcctgg      2450
gtcctgtctg gctctccac tccaggcgga ccctgggggc cagtgaaggga      2500
agctcccgga aagagcagag ggagagcggg taggcggctg tgtgactcta      2550
gtcttggccc caggaagcga aggaacaaaa gaaactggaa aggaagatgc      2600
tttaggaaca tgttttgctt ttttaaaata tataatatta taagagatcc      2650
tttcccattt attctgggaa gatgttttcc aaactcagag acaaggactt      2700
tggtttttgt aagacaaacg atgatatgaa ggccttttgt aagaaaaaat      2750
aaaagatgaa gtgtgaaa      2768
```

<210> SEQ ID NO 16
<211> LENGTH: 673
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 16

```
Met Cys Ser Arg Val Pro Leu Leu Leu Pro Leu Leu Leu Leu      15
 1          5          10          15
Ala Leu Gly Pro Gly Val Gln Gly Cys Pro Ser Gly Cys Gln Cys      30
          20          25          30
Ser Gln Pro Gln Thr Val Phe Cys Thr Ala Arg Gln Gly Thr Thr      45
          35          40          45
Val Pro Arg Asp Val Pro Pro Asp Thr Val Gly Leu Tyr Val Phe      60
          50          55          60
Glu Asn Gly Ile Thr Met Leu Asp Ala Gly Ser Phe Ala Gly Leu      75
          65          70          75
Pro Gly Leu Gln Leu Leu Asp Leu Ser Gln Asn Gln Ile Ala Ser      90
          80          85          90
Leu Pro Ser Gly Val Phe Gln Pro Leu Ala Asn Leu Ser Asn Leu     105
          95          100         105
Asp Leu Thr Ala Asn Arg Leu His Glu Ile Thr Asn Glu Thr Phe     120
          110         115         120
Arg Gly Leu Arg Arg Leu Glu Arg Leu Tyr Leu Gly Lys Asn Arg     135
          125         130         135
Ile Arg His Ile Gln Pro Gly Ala Phe Asp Thr Leu Asp Arg Leu     150
          140         145         150
```

-continued

Leu Glu Leu Lys	Leu Gln Asp Asn Glu	Leu Arg Ala Leu Pro Pro	155	160	165
Leu Arg Leu Pro	Arg Leu Leu Leu Leu	Asp Leu Ser His Asn Ser	170	175	180
Leu Leu Ala Leu	Glu Pro Gly Ile Leu	Asp Thr Ala Asn Val Glu	185	190	195
Ala Leu Arg Leu	Ala Gly Leu Gly Leu	Gln Gln Leu Asp Glu Gly	200	205	210
Leu Phe Ser Arg	Leu Arg Asn Leu His	Asp Leu Asp Val Ser Asp	215	220	225
Asn Gln Leu Glu	Arg Val Pro Pro Val	Ile Arg Gly Leu Arg Gly	230	235	240
Leu Thr Arg Leu	Arg Leu Ala Gly Asn	Thr Arg Ile Ala Gln Leu	245	250	255
Arg Pro Glu Asp	Leu Ala Gly Leu Ala	Ala Leu Gln Glu Leu Asp	260	265	270
Val Ser Asn Leu	Ser Leu Gln Ala Leu	Pro Gly Asp Leu Ser Gly	275	280	285
Leu Phe Pro Arg	Leu Arg Leu Leu Ala	Ala Ala Arg Asn Pro Phe	290	295	300
Asn Cys Val Cys	Pro Leu Ser Trp Phe	Gly Pro Trp Val Arg Glu	305	310	315
Ser His Val Thr	Leu Ala Ser Pro Glu	Glu Thr Arg Cys His Phe	320	325	330
Pro Pro Lys Asn	Ala Gly Arg Leu Leu	Leu Glu Leu Asp Tyr Ala	335	340	345
Asp Phe Gly Cys	Pro Ala Thr Thr Thr	Thr Ala Thr Val Pro Thr	350	355	360
Thr Arg Pro Val	Val Arg Glu Pro Thr	Ala Leu Ser Ser Ser Leu	365	370	375
Ala Pro Thr Trp	Leu Ser Pro Thr Ala	Pro Ala Thr Glu Ala Pro	380	385	390
Ser Pro Pro Ser	Thr Ala Pro Pro Thr	Val Gly Pro Val Pro Gln	395	400	405
Pro Gln Asp Cys	Pro Pro Ser Thr Cys	Leu Asn Gly Gly Thr Cys	410	415	420
His Leu Gly Thr	Arg His His Leu Ala	Cys Leu Cys Pro Glu Gly	425	430	435
Phe Thr Gly Leu	Tyr Cys Glu Ser Gln	Met Gly Gln Gly Thr Arg	440	445	450
Pro Ser Pro Thr	Pro Val Thr Pro Arg	Pro Pro Arg Ser Leu Thr	455	460	465
Leu Gly Ile Glu	Pro Val Ser Pro Thr	Ser Leu Arg Val Gly Leu	470	475	480
Gln Arg Tyr Leu	Gln Gly Ser Ser Val	Gln Leu Arg Ser Leu Arg	485	490	495
Leu Thr Tyr Arg	Asn Leu Ser Gly Pro	Asp Lys Arg Leu Val Thr	500	505	510
Leu Arg Leu Pro	Ala Ser Leu Ala Glu	Tyr Thr Val Thr Gln Leu	515	520	525

-continued

Arg	Pro	Asn	Ala	Thr	Tyr	Ser	Val	Cys	Val	Met	Pro	Leu	Gly	Pro	
				530					535					540	
Gly	Arg	Val	Pro	Glu	Gly	Glu	Glu	Ala	Cys	Gly	Glu	Ala	His	Thr	
				545					550					555	
Pro	Pro	Ala	Val	His	Ser	Asn	His	Ala	Pro	Val	Thr	Gln	Ala	Arg	
				560					565					570	
Glu	Gly	Asn	Leu	Pro	Leu	Leu	Ile	Ala	Pro	Ala	Leu	Ala	Ala	Val	
				575					580					585	
Leu	Leu	Ala	Ala	Leu	Ala	Ala	Val	Gly	Ala	Ala	Tyr	Cys	Val	Arg	
				590					595					600	
Arg	Gly	Arg	Ala	Met	Ala	Ala	Ala	Ala	Gln	Asp	Lys	Gly	Gln	Val	
				605					610					615	
Gly	Pro	Gly	Ala	Gly	Pro	Leu	Glu	Leu	Glu	Gly	Val	Lys	Val	Pro	
				620					625					630	
Leu	Glu	Pro	Gly	Pro	Lys	Ala	Thr	Glu	Gly	Gly	Gly	Glu	Ala	Leu	
				635					640					645	
Pro	Ser	Gly	Ser	Glu	Cys	Glu	Val	Pro	Leu	Met	Gly	Phe	Pro	Gly	
				650					655					660	
Pro	Gly	Leu	Gln	Ser	Pro	Leu	His	Ala	Lys	Pro	Tyr	Ile			
				665					670						

<210> SEQ ID NO 17
<211> LENGTH: 1672
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 17

gcagcggcga ggcggcggtg gtggctgagt ccgtgggtggc agaggcgaag	50
gcgacagctc atgcgggtcc ggatagggct gacgctgctg ctgtgtgcgg	100
tgctgctgag cttggcctcg gcgtcctcgg atgaagaagg cagccaggat	150
gaatccttag attccaagac tactttgaca tcagatgagt cagtaaagga	200
ccatactact gcaggcagag tagttgctgg tcaaatatth cttgattcag	250
aagaatctga attagaatcc tctattcaag aagaggaaga cagcctcaag	300
agccaagagg gggaaagtgt cacagaagat atcagcttcc tagagtctcc	350
aaatccagaa aacaaggact atgaagagcc aaagaaagta cggaaaccag	400
ctttgaccgc cattgaaggc acagcacatg gggagccctg ccacttcctc	450
tttcttttcc tagataagga gtatgatgaa tgtacatcag atgggaggga	500
agatggcaga ctgtggtgtg ctacaacctg tgactacaaa gcagatgaaa	550
agtggggctt ttgtgaaact gaagaagagg ctgctaagag acggcagatg	600
caggaagcag aaatgatgta tcaaactgga atgaaaatcc ttaatggaag	650
caataagaaa agccaaaaaa gagaagcata tcggtatctc caaaaggcag	700
caagcatgaa ccataccaaa gccctggaga gagtgtcata tgctctttta	750
tttggtgatt acttgccaca gaatatccag gcagcgagag agatgtttga	800
gaagctgact gaggaaggct ctccaaggg acagactgct cttggctttc	850
tgtagcctc ttgacttggt gttaattcaa gtcaggcaaa ggctcttgta	900
tattatacat ttggagctct tgggggcaat ctaatagccc acatggtttt	950

-continued

ggtaagtaga ctttagtgga aggctaataa tattaacatc agaagaattt	1000
gtggtttata gcggccacaa ctttttcagc tttcatgac cagatttgct	1050
tgtattaaga ccaaatattc agttgaactt ccttcaaatt cttgttaatg	1100
gatataacac atggaatcta catgtaaatg aaagttggtg gagtcacaaa	1150
tttttcttta aaatgattag ttggctgat tgcccctaaa aagagagatc	1200
tgataaatgg ctcttttttaa attttctctg agttggaatt gtcagaatca	1250
ttttttacat tagattatca taattttaaa aatttttctt tagtttttca	1300
aaattttgta aatggtggct atagaaaaac aacatgaaat attatacaat	1350
attttgcaac aatgccctaa gaattgttaa aattcatgga gttatttgctg	1400
cagaatgact ccagagagct ctactttctg ttttttactt ttcatgattg	1450
gctgtcttcc catttattct ggtcatttat tgctagtgc actgtgcctg	1500
cttcagtag tctcattttc cctattttgc taatttgta ctttttcttt	1550
gctaatttgg aagattaact catttttaat aaaattatgt ctaagattaa	1600
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1650
aaaaaaaaa aaaaaaaaaa aa	1672

<210> SEQ ID NO 18
<211> LENGTH: 301
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 18

Met Arg Val Arg Ile Gly Leu Thr Leu Leu Leu Cys Ala Val Leu	1	5	10	15
Leu Ser Leu Ala Ser Ala Ser Ser Asp Glu Glu Gly Ser Gln Asp	20	25	30	
Glu Ser Leu Asp Ser Lys Thr Thr Leu Thr Ser Asp Glu Ser Val	35	40	45	
Lys Asp His Thr Thr Ala Gly Arg Val Val Ala Gly Gln Ile Phe	50	55	60	
Leu Asp Ser Glu Glu Ser Glu Leu Glu Ser Ser Ile Gln Glu Glu	65	70	75	
Glu Asp Ser Leu Lys Ser Gln Glu Gly Glu Ser Val Thr Glu Asp	80	85	90	
Ile Ser Phe Leu Glu Ser Pro Asn Pro Glu Asn Lys Asp Tyr Glu	95	100	105	
Glu Pro Lys Lys Val Arg Lys Pro Ala Leu Thr Ala Ile Glu Gly	110	115	120	
Thr Ala His Gly Glu Pro Cys His Phe Pro Phe Leu Phe Leu Asp	125	130	135	
Lys Glu Tyr Asp Glu Cys Thr Ser Asp Gly Arg Glu Asp Gly Arg	140	145	150	
Leu Trp Cys Ala Thr Thr Tyr Asp Tyr Lys Ala Asp Glu Lys Trp	155	160	165	
Gly Phe Cys Glu Thr Glu Glu Glu Ala Ala Lys Arg Arg Gln Met	170	175	180	
Gln Glu Ala Glu Met Met Tyr Gln Thr Gly Met Lys Ile Leu Asn	185	190	195	

-continued

Gly	Ser	Asn	Lys	Lys	Ser	Gln	Lys	Arg	Glu	Ala	Tyr	Arg	Tyr	Leu
			200						205					210
Gln	Lys	Ala	Ala	Ser	Met	Asn	His	Thr	Lys	Ala	Leu	Glu	Arg	Val
			215						220					225
Ser	Tyr	Ala	Leu	Leu	Phe	Gly	Asp	Tyr	Leu	Pro	Gln	Asn	Ile	Gln
			230						235					240
Ala	Ala	Arg	Glu	Met	Phe	Glu	Lys	Leu	Thr	Glu	Glu	Gly	Ser	Pro
			245						250					255
Lys	Gly	Gln	Thr	Ala	Leu	Gly	Phe	Leu	Tyr	Ala	Ser	Gly	Leu	Gly
			260						265					270
Val	Asn	Ser	Ser	Gln	Ala	Lys	Ala	Leu	Val	Tyr	Tyr	Thr	Phe	Gly
			275						280					285
Ala	Leu	Gly	Gly	Asn	Leu	Ile	Ala	His	Met	Val	Leu	Val	Ser	Arg
			290						295					300
Leu														

<210> SEQ ID NO 19
<211> LENGTH: 1508
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 19

aattcagatt ttaagcccat tctgcagtgg aatttcatga actagcaaga	50
ggacaccatc ttcttgtatt atacaagaaa ggagtgtacc tatcacacac	100
agggggaaaa atgctctttt ggggtgctagg cctcctaatac ctctgtggtt	150
ttctgtggag tcgtaaagga aaactaaaga ttgaagacat cactgataag	200
tacattttta tcaactggatg tgactcgggc tttggaaact tggcagccag	250
aacttttgat aaaaagggat ttcattgtaat cgctgcctgt ctgactgaat	300
caggatcaac agctttaaag gcagaaacct cagagagact tcgtactgtg	350
cttctggatg tgaccgaccc agagaatgtc aagaggactg cccagtgggt	400
gaagaaccaa gttggggaga aaggtctctg gggctctgac aataatgctg	450
gtgttcccgg cgtgctggct cccactgact ggctgacact agaggactac	500
agagaacctt ttgaagtga cctgtttgga ctcactcagt tgacactaaa	550
tatgcttcct ttggtcaaga aagctcaagg gagagttatt aatgtctcca	600
gtgttgaggg tcgccttgca atcgttgagg ggggctatac tccatccaaa	650
tatgcagtgg aaggtttcaa tgacagctta agacgggaca tgaaagcttt	700
tggtgtgcac gtctcatgca ttgaaccagg attgttcaa acaaacttgg	750
cagatccagt aaaggttaatt gaaaaaaaaa tcgccatttg ggagcagctg	800
tctccagaca tcaaacaaca atatggagaa gggttacattg aaaaaagtct	850
agacaaaactg aaaggcaata aatcctatgt gaacatggac ctctctccgg	900
tggtagagtg catggaccac gctctaacaa gtctcttccc taagactcat	950
tatgccgctg gaaaagatgc caaaattttc tggatacctc tgtctcacat	1000
gccagcagct ttgcaagact ttttattgtt gaaacagaaa gcagagctgg	1050
ctaattcccaa ggcagtgtga ctacagtaac cacaaatgtc tcctccaggc	1100

-continued

tatgaaattg gccgatttca agaacacatc tccttttcaa ccccatctct	1150
tatctgctcc aacctggact catttagatc gtgcttattt ggattgcaaa	1200
agggagtgccc accatcgctg gtggtatccc agggtcctcg ctcaagtttt	1250
ctttgaaaaa gagggctgga atggtacatc acataggcaa gtcctgccct	1300
gtatttaggc tttgcctgct tgggtgatg taagggaat tgaaagactt	1350
gcccattcaa aatgatcttt accgtggcct gcccctgct tatgggtccc	1400
agcattttaca gtaacttggt aatgttaagt atcatctctt atctaaatat	1450
taaaagataa gtcaacccaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1500
aaaaaaaa	1508

<210> SEQ ID NO 20

<211> LENGTH: 319

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 20

Met Leu Phe Trp Val Leu Gly Leu Leu Ile Leu Cys Gly Phe Leu	
1 5 10 15	
Trp Thr Arg Lys Gly Lys Leu Lys Ile Glu Asp Ile Thr Asp Lys	
20 25 30	
Tyr Ile Phe Ile Thr Gly Cys Asp Ser Gly Phe Gly Asn Leu Ala	
35 40 45	
Ala Arg Thr Phe Asp Lys Lys Gly Phe His Val Ile Ala Ala Cys	
50 55 60	
Leu Thr Glu Ser Gly Ser Thr Ala Leu Lys Ala Glu Thr Ser Glu	
65 70 75	
Arg Leu Arg Thr Val Leu Leu Asp Val Thr Asp Pro Glu Asn Val	
80 85 90	
Lys Arg Thr Ala Gln Trp Val Lys Asn Gln Val Gly Glu Lys Gly	
95 100 105	
Leu Trp Gly Leu Ile Asn Asn Ala Gly Val Pro Gly Val Leu Ala	
110 115 120	
Pro Thr Asp Trp Leu Thr Leu Glu Asp Tyr Arg Glu Pro Ile Glu	
125 130 135	
Val Asn Leu Phe Gly Leu Ile Ser Val Thr Leu Asn Met Leu Pro	
140 145 150	
Leu Val Lys Lys Ala Gln Gly Arg Val Ile Asn Val Ser Ser Val	
155 160 165	
Gly Gly Arg Leu Ala Ile Val Gly Gly Gly Tyr Thr Pro Ser Lys	
170 175 180	
Tyr Ala Val Glu Gly Phe Asn Asp Ser Leu Arg Arg Asp Met Lys	
185 190 195	
Ala Phe Gly Val His Val Ser Cys Ile Glu Pro Gly Leu Phe Lys	
200 205 210	
Thr Asn Leu Ala Asp Pro Val Lys Val Ile Glu Lys Lys Leu Ala	
215 220 225	
Ile Trp Glu Gln Leu Ser Pro Asp Ile Lys Gln Gln Tyr Gly Glu	
230 235 240	
Gly Tyr Ile Glu Lys Ser Leu Asp Lys Leu Lys Gly Asn Lys Ser	
245 250 255	

-continued

Tyr	Val	Asn	Met	Asp	Leu	Ser	Pro	Val	Val	Glu	Cys	Met	Asp	His
			260						265					270
Ala	Leu	Thr	Ser	Leu	Phe	Pro	Lys	Thr	His	Tyr	Ala	Ala	Gly	Lys
			275						280					285
Asp	Ala	Lys	Ile	Phe	Trp	Ile	Pro	Leu	Ser	His	Met	Pro	Ala	Ala
			290						295					300
Leu	Gln	Asp	Phe	Leu	Leu	Lys	Gln	Lys	Ala	Glu	Leu	Ala	Asn	
			305					310					315	
Pro	Lys	Ala	Val											

<210> SEQ ID NO 21
<211> LENGTH: 1849
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 21

ctgaggcggc ggtagcatgg agggggagag tacgtcggcg gtgctctcgg	50
gctttgtgct cggcgcactc gctttccagc acctcaacac ggactcggac	100
acggaaggtt ttcttcttgg ggaagtaaaa ggtgaagcca agaacagcat	150
tactgattcc caaatggatg atgttgaagt tgtttataca attgacattc	200
agaaatatat tccatgctat cagcttttta gcttttataa ttcttcaggc	250
gaagtaaatg agcaagcact gaagaaaata ttatcaaatg tcaaaaagaa	300
tgtggtaggt tggtaacaaat tccgtcgtca ttcagatcag atcatgacgt	350
ttagagagag gctgcttcac aaaaacttgc aggagcattt ttcaaaccaa	400
gaccttgttt ttctgctatt aacaccaagt ataataacag aaagctgctc	450
tactcatcga ctggaacatt ccttatataa acctcaaaaa ggactttttc	500
acaggggtacc tttagtgggt gccaatctgg gcatgtctga acaactgggt	550
tataaaactg tatcaggttc ctgtatgtcc actggtttta gccgagcagt	600
acaaacacac agctctaata tttttgaaga agatggatcc ttaaaggag	650
tacataagat aaatgaaatg tatgcttcat tacaagagga attaaagagt	700
atatgcaaaa aagtgaaga cagtgaacaa gcagtagata aactagtaaa	750
ggatgtaaac agattaaaac gagaaattga gaaaaggaga ggagcacaga	800
ttcaggcagc aagagagaag aacatccaaa aagaccctca ggagaacatt	850
tttctttgtc aggcattacg gacctttttt ccaaattctg aatttcttca	900
ttcatgtgtt atgtctttaa aaaatagaca tgtttctaaa agtagctgta	950
actacaacca ccatctcgat gtagtagaca atctgacctt aatggtagaa	1000
cacactgaca ttctgaagc tagtccagct agtacaccac aaatcattaa	1050
gcataaagcc ttagacttag atgacagatg gcaattcaag agatctcgg	1100
tgtagatata acaagacaaa cgatctaaag caaatactgg tagtagtaac	1150
caagataaag catccaaaat gagcagccca gaaacagatg aagaaattga	1200
aaagatgaag ggttttggtg aatattcacg gtctoctaca ttttgatcct	1250
tttaacctta caaggagatt tttttatttg gctgatgggt aaagccaaac	1300
atttctattg tttttactat gttgagctac ttgcagtaag ttcatttgtt	1350

-continued

tttactatgt tcacctgttt gcagtaatac acagataact cttagtgc	1400
ttacttcaca aagtactttt tcaaacaatca gatgctttta tttccaaacc	1450
ttttttttcac ctttcactaa gttgttgagg ggaaggctta cacagacaca	1500
ttcttttagaa ttggaaaagt gagaccaggc acagtggctc acacctgtaa	1550
ttccagcact tagggaagac aagtcaggag gattgattga agctaggagt	1600
tagagaccag cctgggcaac gtattgagac catgtctatt aaaaaataaa	1650
atggaaaagc aagaatagcc ttattttcaa aatatggaaa gaaatttata	1700
tgaaaattta tctgagtcac taaaattctc cttaagtgat acttttttag	1750
aagtacatta tggctagagt tgccagataa aatgctggat atcatgcaat	1800
aaatttgcaa aacatcatct aaaatttaaa aaaaaaaaaa aaaaaaaaaa	1849

<210> SEQ ID NO 22

<211> LENGTH: 409

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 22

Met Glu Gly Glu Ser Thr Ser Ala Val Leu Ser Gly Phe Val Leu	1	5	10	15
Gly Ala Leu Ala Phe Gln His Leu Asn Thr Asp Ser Asp Thr Glu	20	25	30	
Gly Phe Leu Leu Gly Glu Val Lys Gly Glu Ala Lys Asn Ser Ile	35	40	45	
Thr Asp Ser Gln Met Asp Asp Val Glu Val Val Tyr Thr Ile Asp	50	55	60	
Ile Gln Lys Tyr Ile Pro Cys Tyr Gln Leu Phe Ser Phe Tyr Asn	65	70	75	
Ser Ser Gly Glu Val Asn Glu Gln Ala Leu Lys Lys Ile Leu Ser	80	85	90	
Asn Val Lys Lys Asn Val Val Gly Trp Tyr Lys Phe Arg Arg His	95	100	105	
Ser Asp Gln Ile Met Thr Phe Arg Glu Arg Leu Leu His Lys Asn	110	115	120	
Leu Gln Glu His Phe Ser Asn Gln Asp Leu Val Phe Leu Leu Leu	125	130	135	
Thr Pro Ser Ile Ile Thr Glu Ser Cys Ser Thr His Arg Leu Glu	140	145	150	
His Ser Leu Tyr Lys Pro Gln Lys Gly Leu Phe His Arg Val Pro	155	160	165	
Leu Val Val Ala Asn Leu Gly Met Ser Glu Gln Leu Gly Tyr Lys	170	175	180	
Thr Val Ser Gly Ser Cys Met Ser Thr Gly Phe Ser Arg Ala Val	185	190	195	
Gln Thr His Ser Ser Lys Phe Phe Glu Glu Asp Gly Ser Leu Lys	200	205	210	
Glu Val His Lys Ile Asn Glu Met Tyr Ala Ser Leu Gln Glu Glu	215	220	225	
Leu Lys Ser Ile Cys Lys Lys Val Glu Asp Ser Glu Gln Ala Val	230	235	240	

-continued

Asp Lys Leu Val Lys Asp Val Asn Arg Leu Lys Arg Glu Ile Glu
245 250 255

Lys Arg Arg Gly Ala Gln Ile Gln Ala Ala Arg Glu Lys Asn Ile
260 265 270

Gln Lys Asp Pro Gln Glu Asn Ile Phe Leu Cys Gln Ala Leu Arg
275 280 285

Thr Phe Phe Pro Asn Ser Glu Phe Leu His Ser Cys Val Met Ser
290 295 300

Leu Lys Asn Arg His Val Ser Lys Ser Ser Cys Asn Tyr Asn His
305 310 315

His Leu Asp Val Val Asp Asn Leu Thr Leu Met Val Glu His Thr
320 325 330

Asp Ile Pro Glu Ala Ser Pro Ala Ser Thr Pro Gln Ile Ile Lys
335 340 345

His Lys Ala Leu Asp Leu Asp Asp Arg Trp Gln Phe Lys Arg Ser
350 355 360

Arg Leu Leu Asp Thr Gln Asp Lys Arg Ser Lys Ala Asn Thr Gly
365 370 375

Ser Ser Asn Gln Asp Lys Ala Ser Lys Met Ser Ser Pro Glu Thr
380 385 390

Asp Glu Glu Ile Glu Lys Met Lys Gly Phe Gly Glu Tyr Ser Arg
395 400 405

Ser Pro Thr Phe

<210> SEQ ID NO 23
<211> LENGTH: 2651
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 23

ggcacagccg cgcgggcgag ggcagagtca gccgagccga gtccagccgg 50

acgagcgggc cagcgcaggg cagcccaagc agcgcgagc gaacgcccgc 100

cgccgcccac accctctgcg gtccccgcgg cgcctgccac ccttccctcc 150

ttccccgcgt ccccgctcgc cgggccagtc agcttgccgg gttcgctgcc 200

ccgcgaaacc ccgaggtcac cagcccgcg cctctgcttc ctgggcccgcg 250

cgccgcctcc acgcccctct tctcccctgg cccggcgctt ggcaccgggg 300

accgttgctt gacgcgaggc ccagctctac ttttcgcccc gcgtctcctc 350

cgctgctcgc cctcttccac caactccaac tccttctccc tccagctcca 400

ctcgctagtc cccgactccg ccagccctcg gcccgctgcc gtagcgccgc 450

ttcccgctcg gtcccaaagg tgggaacgcg tccgccccgg cccgcacccat 500

ggcacgggtc ggcttgcccg cgcttctctg caccctggca gtgctcagcg 550

ccgcgctgct ggctgccgag ctcaagtcca aaagtgtctt ggaagtgcga 600

cgctctttacg tgtccaaagg cttcaacaag aacgatgccc cctccacga 650

gatcaacggt gatcatttga agatctgtcc ccagggttct acctgctgct 700

ctcaagagat ggaggagaag tacagcctgc aaagtaaaga tgatttcaaa 750

agtgtggtca gcgaacagtg caatcatttg caagctgtct ttgcttcacg 800

-continued

ttacaagaag tttgatgaat tcttcaaaga actacttgaa aatgcagaga	850
aatccctgaa tgatatgttt gtgaagacat atggccatth atacatgcaa	900
aattctgagc tatttaaaga tctcttcgta gagttgaaac gttactacgt	950
gggtgggaaat gtgaacctgg aagaaatgct aaatgacttc tgggctcgcc	1000
tcttgagcgc gatgttcgcg ctggtgaact ccagtagca ctttacagat	1050
gagtatctgg aatgtgtgag caagtatacg gagcagctga agcccttcgg	1100
agatgtccct cgcaaatga agctccaggt tactcgtgct tttgtagcag	1150
cccgtactth cgctcaaggc ttagcggttg cgggagatgt cgtgagcaag	1200
gtctccgtgg taaacccac agcccagtg acccatgccc tgttgaagat	1250
gatctactgc tcccactgcc ggggtctcgt gactgtgaag ccatgttaca	1300
actactgctc aaacatcatg agaggctgtt tggccaacca aggggatctc	1350
gattttgaat ggaacaatth catagatgct atgctgatgg tggcagagag	1400
gctagagggt cctttcaaca ttgaatcggt catggatccc atcgaagtga	1450
agatttctga tgctattatg aacatgcagg ataatagtgt tcaagtgtct	1500
cagaaggtht tccagggatg tggaccccc aagcccctcc cagctggacg	1550
aatttctcgt tccatctctg aaagtgcctt cagtgtctgc ttcagaccac	1600
atcaccccca ggaacgcca accacagcag ctggcactag tttggaccga	1650
ctggttactg atgtcaagga gaaactgaaa caggccaaga aattctggtc	1700
ctcccttcgc agcaacgtht gcaacgatga gaggatggct gcaggaaacg	1750
gcaatgagga tgactgttgg aatgggaaa gcaaaagcag gtacctgtth	1800
gcagtgcagc gaaatggatt agccaaccag ggcaacaacc cagagggtcca	1850
gggtgacacc agcaaacag acatactgat ccttcgtcaa atcatggctc	1900
ttcgaagtga gaccagcaag atgaagaatg catacaatgg gaacgacgtg	1950
gacttctttg atatcagtga tgaaagtatg ggagaaggaa gtggaagtgg	2000
ctgtgagtat cagcagtgcc cttcagagth tgactacaat gccactgacc	2050
atgctgggaa gagtgccaat gagaagccg acagtgtgg tgtccgtcct	2100
ggggcacagg cctacctcct cactgtcttc tgcactctgt tcttggttat	2150
gcagagagag tggagataat tctcaaactc tgagaaaaag tgttcatcaa	2200
aaagttaaaa ggcaccagth atcactthtc taccatccta gtgactttgc	2250
tttttaaatg aatggacaac aatgtacagt ttttactatg tggccactgg	2300
tttaagaagt gctgactttg ttttctcatt cagttttggg aggaaaaggg	2350
actgtgcatt gagtggthc ctgctcccc aaaccatgth aaacgtggct	2400
aacagtgtag gtacagaact atagttagth gtgcatttht gatthttatca	2450
ctctattatt tgthtgtatg ttttttctc atttcgtthg tgggtthttt	2500
tttccaactg tgatctcgcc ttgthtctta caagcaaac agggctccct	2550
cttggcacgt aacatgtacg tatttctgaa atattaaata gctgtacaga	2600
agcaggthtt atttatcatg ttatcttatt aaaagaaaa gcccaaaaag	2650
c	2651

-continued

<210> SEQ ID NO 24

<211> LENGTH: 556

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 24

```

Met Ala Arg Phe Gly Leu Pro Ala Leu Leu Cys Thr Leu Ala Val
 1              5              10              15

Leu Ser Ala Ala Leu Leu Ala Ala Glu Leu Lys Ser Lys Ser Cys
20              25              30

Ser Glu Val Arg Arg Leu Tyr Val Ser Lys Gly Phe Asn Lys Asn
35              40              45

Asp Ala Pro Leu His Glu Ile Asn Gly Asp His Leu Lys Ile Cys
50              55              60

Pro Gln Gly Ser Thr Cys Cys Ser Gln Glu Met Glu Glu Lys Tyr
65              70              75

Ser Leu Gln Ser Lys Asp Asp Phe Lys Ser Val Val Ser Glu Gln
80              85              90

Cys Asn His Leu Gln Ala Val Phe Ala Ser Arg Tyr Lys Lys Phe
95              100             105

Asp Glu Phe Phe Lys Glu Leu Leu Glu Asn Ala Glu Lys Ser Leu
110             115             120

Asn Asp Met Phe Val Lys Thr Tyr Gly His Leu Tyr Met Gln Asn
125             130             135

Ser Glu Leu Phe Lys Asp Leu Phe Val Glu Leu Lys Arg Tyr Tyr
140             145             150

Val Val Gly Asn Val Asn Leu Glu Glu Met Leu Asn Asp Phe Trp
155             160             165

Ala Arg Leu Leu Glu Arg Met Phe Arg Leu Val Asn Ser Gln Tyr
170             175             180

His Phe Thr Asp Glu Tyr Leu Glu Cys Val Ser Lys Tyr Thr Glu
185             190             195

Gln Leu Lys Pro Phe Gly Asp Val Pro Arg Lys Leu Lys Leu Gln
200             205             210

Val Thr Arg Ala Phe Val Ala Ala Arg Thr Phe Ala Gln Gly Leu
215             220             225

Ala Val Ala Gly Asp Val Val Ser Lys Val Ser Val Val Asn Pro
230             235             240

Thr Ala Gln Cys Thr His Ala Leu Leu Lys Met Ile Tyr Cys Ser
245             250             255

His Cys Arg Gly Leu Val Thr Val Lys Pro Cys Tyr Asn Tyr Cys
260             265             270

Ser Asn Ile Met Arg Gly Cys Leu Ala Asn Gln Gly Asp Leu Asp
275             280             285

Phe Glu Trp Asn Asn Phe Ile Asp Ala Met Leu Met Val Ala Glu
290             295             300

Arg Leu Glu Gly Pro Phe Asn Ile Glu Ser Val Met Asp Pro Ile
305             310             315

Asp Val Lys Ile Ser Asp Ala Ile Met Asn Met Gln Asp Asn Ser
320             325             330

Val Gln Val Ser Gln Lys Val Phe Gln Gly Cys Gly Pro Pro Lys
335             340             345

```

-continued

Pro Leu Pro Ala Gly Arg Ile Ser Arg Ser Ile Ser Glu Ser Ala
350 355 360
Phe Ser Ala Arg Phe Arg Pro His His Pro Glu Glu Arg Pro Thr
365 370 375
Thr Ala Ala Gly Thr Ser Leu Asp Arg Leu Val Thr Asp Val Lys
380 385 390
Glu Lys Leu Lys Gln Ala Lys Lys Phe Trp Ser Ser Leu Pro Ser
395 400 405
Asn Val Cys Asn Asp Glu Arg Met Ala Ala Gly Asn Gly Asn Glu
410 415 420
Asp Asp Cys Trp Asn Gly Lys Gly Lys Ser Arg Tyr Leu Phe Ala
425 430 435
Val Thr Gly Asn Gly Leu Ala Asn Gln Gly Asn Asn Pro Glu Val
440 445 450
Gln Val Asp Thr Ser Lys Pro Asp Ile Leu Ile Leu Arg Gln Ile
455 460 465
Met Ala Leu Arg Val Met Thr Ser Lys Met Lys Asn Ala Tyr Asn
470 475 480
Gly Asn Asp Val Asp Phe Phe Asp Ile Ser Asp Glu Ser Ser Gly
485 490 495
Glu Gly Ser Gly Ser Gly Cys Glu Tyr Gln Gln Cys Pro Ser Glu
500 505 510
Phe Asp Tyr Asn Ala Thr Asp His Ala Gly Lys Ser Ala Asn Glu
515 520 525
Lys Ala Asp Ser Ala Gly Val Arg Pro Gly Ala Gln Ala Tyr Leu
530 535 540
Leu Thr Val Phe Cys Ile Leu Phe Leu Val Met Gln Arg Glu Trp
545 550 555

Arg

<210> SEQ ID NO 25
<211> LENGTH: 870
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 25

ctcgccctca aatgggaacg ctggcctggg actaaagcat agaccaccag 50
gctgagtatc ctgacctgag tcatccccag ggatcaggag cctccagcag 100
ggaaccttcc attatattct tcaagcaact tacagctgca cgcacagttg 150
cgatgaaagt tctaattctt tccctcctcc tgttgctgcc actaatgctg 200
atgtccatgg tctctagcag cctgaatcca ggggtcgcca gaggccacag 250
ggaccgaggg caggcttcta ggagatggct ccaggaaggc ggccaagaat 300
gtgagtgcaa agattggttc ctgagagccc cgagaagaaa attcatgaca 350
tgtgtctggg tgccaaagaa gcagtgcccc tgtgatcatt tcaaggggaa 400
tgtgaagaaa acaagacacc aaaggcacca cagaaagcca aacaagcatt 450
ccagagcctg ccagcaattt ctcaacaat gtcagctaag aagctttgct 500
ctgcctttgt aggagctctg agcgcccact cttccaatta aacattctca 550
gccaagaaga cagtgagcac acctaccaga cactcttctt ctcccacctc 600

-continued

actctcccac tgtaccacc cctaaatcat tccagtgtc tcaaaaagca	650
tgtttttcaa gatcattttg ttgttgctc tctctagtgt cttcttctct	700
cgtcagtctt agcctgtgcc ctccccttac ccaggcttag gcttaattac	750
ctgaaagatt ccaggaaact gtagcttcct agctagtgtc atttaacctt	800
aaatgcaatc aggaaagtag caaacagaag tcaataaata tttttaaatg	850
tcaaaaaaaaa aaaaaaaaaa	870

<210> SEQ ID NO 26
<211> LENGTH: 119
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 26

Met Lys Val Leu Ile Ser Ser Leu Leu Leu Leu Pro Leu Met	
1 5 10 15	
Leu Met Ser Met Val Ser Ser Ser Leu Asn Pro Gly Val Ala Arg	
20 25 30	
Gly His Arg Asp Arg Gly Gln Ala Ser Arg Arg Trp Leu Gln Glu	
35 40 45	
Gly Gly Gln Glu Cys Glu Cys Lys Asp Trp Phe Leu Arg Ala Pro	
50 55 60	
Arg Arg Lys Phe Met Thr Val Ser Gly Leu Pro Lys Lys Gln Cys	
65 70 75	
Pro Cys Asp His Phe Lys Gly Asn Val Lys Lys Thr Arg His Gln	
80 85 90	
Arg His His Arg Lys Pro Asn Lys His Ser Arg Ala Cys Gln Gln	
95 100 105	
Phe Leu Lys Gln Cys Gln Leu Arg Ser Phe Ala Leu Pro Leu	
110 115	

<210> SEQ ID NO 27
<211> LENGTH: 1371
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 27

ggagccagc gcctgcagag gctgagcagg gaaaaagcca gtgcccagc	50
ggaagcacag ctgagagctg gtctgccatg gacatcctgg tcccactcct	100
gcagctgctg gtgctgcttc ttaccctgcc cctgcacctc atggctctgc	150
tgggctgctg gcagcccctg tgcaaaagct acttccccta cctgatggcc	200
gtgctgactc ccaagagcaa ccgcaagatg gagagcaaga aacgggagct	250
cttcagccag ataaaggggc ttacaggagc ctccgggaaa gtggccctac	300
tggagctggg ctgcggaacc ggagccaact ttcagttcta cccaccgggc	350
tgcagggtca cctgcctaga cccaaatccc cactttgaga agttcctgac	400
aaagagcatg gctgagaaca ggcacctcca atatgagcgg tttgtggtgg	450
ctcctggaga ggacatgaga cagctggctg atggctccat ggatgtggtg	500
gtctgcactc tgggtgctgt ctctgtgcag agcccaagga aggtcctgca	550
ggaggtccgg agagtactga gaccgggagg tgtgtctctt ttctgggagc	600

-continued

```

atgtggcaga accatatgga agctgggcct tcatgtggca gcaagttttc      650
gagcccacct ggaacacat tggggatggc tgctgcctca ccagagagac      700
ctggaaggat cttgagaacg cccagttctc cgaaatccaa atggaacgac      750
agccccctcc cttgaagtgg ctacctgttg ggccccacat catgggaaag      800
gctgtcaaac aatctttccc aagctccaag gcaactcatt gtccttccc      850
cagcctccaa ttagaacaag ccaccacca gcctatctat cttccactga      900
gagggaccta gcagaatgag agaagacatt catgtaccac ctactagtcc      950
ctctctcccc aacctctgcc agggcaatct ctaactcaa tccgccttc     1000
gacagtgaaa aagctctact tctacgtga cccagggagg aaacactagg     1050
accctgttgt atcctcaact gcaagtttct ggactagtct cccaacgttt     1100
gcctcccaat gttgtccctt tccttcgttc ccatggtaaa gtcctctctg     1150
ctttcctcct gaggctacac ccatgcgtct ctaggaactg gtcacaaaag     1200
tcatggtgcc tgcacccctg ccaagccccc ctgacctctc ctccccacta     1250
ccaccttctt cctgagctgg gggcaccagg gagaatcaga gatgctgggg     1300
atgccagagc aagactcaaa gaggcagagg ttttgttctc aaatatTTTT     1350
taataaatag acgaaaccac g                                     1371

```

<210> SEQ ID NO 28

<211> LENGTH: 277

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 28

```

Met Asp Ile Leu Val Pro Leu Leu Gln Leu Leu Val Leu Leu Leu
 1             5             10             15
Thr Leu Pro Leu His Leu Met Ala Leu Leu Gly Cys Trp Gln Pro
             20             25             30
Leu Cys Lys Ser Tyr Phe Pro Tyr Leu Met Ala Val Leu Thr Pro
             35             40             45
Lys Ser Asn Arg Lys Met Glu Ser Lys Lys Arg Glu Leu Phe Ser
             50             55             60
Gln Ile Lys Gly Leu Thr Gly Ala Ser Gly Lys Val Ala Leu Leu
             65             70             75
Glu Leu Gly Cys Gly Thr Gly Ala Asn Phe Gln Phe Tyr Pro Pro
             80             85             90
Gly Cys Arg Val Thr Cys Leu Asp Pro Asn Pro His Phe Glu Lys
             95            100            105
Phe Leu Thr Lys Ser Met Ala Glu Asn Arg His Leu Gln Tyr Glu
            110            115            120
Arg Phe Val Val Ala Pro Gly Glu Asp Met Arg Gln Leu Ala Asp
            125            130            135
Gly Ser Met Asp Val Val Val Cys Thr Leu Val Leu Cys Ser Val
            140            145            150
Gln Ser Pro Arg Lys Val Leu Gln Glu Val Arg Arg Val Leu Arg
            155            160            165
Pro Gly Gly Val Leu Phe Phe Trp Glu His Val Ala Glu Pro Tyr
            170            175            180

```

-continued

Gly Ser Trp Ala Phe Met Trp Gln Gln Val Phe Glu Pro Thr Trp
185 190 195

Lys His Ile Gly Asp Gly Cys Cys Leu Thr Arg Glu Thr Trp Lys
200 205 210

Asp Leu Glu Asn Ala Gln Phe Ser Glu Ile Gln Met Glu Arg Gln
215 220 225

Pro Pro Pro Leu Lys Trp Leu Pro Val Gly Pro His Ile Met Gly
230 235 240

Lys Ala Val Lys Gln Ser Phe Pro Ser Ser Lys Ala Leu Ile Cys
245 250 255

Ser Phe Pro Ser Leu Gln Leu Glu Gln Ala Thr His Gln Pro Ile
260 265 270

Tyr Leu Pro Leu Arg Gly Thr
275

<210> SEQ ID NO 29
<211> LENGTH: 494
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 29

caatgtttgc ctatccacct cccccaagcc cctttaccta tgctgctgct 50

aacgctgctg ctgctgctgc tgctgcttaa aggctcatgc ttggagtggg 100

gactggtcgg tgcccagaaa gtctcttctg ccactgacgc ccccatcagg 150

gattgggcct tctttccccc ttcttttctg tgtctcctgc ctcatcggcc 200

tgcccatgacc tgcagccaag ccagccccg tggggaaggg gagaaagtgg 250

gggatggcta agaaagctgg gagataggga acagaagagg gtagtgggtg 300

ggctaggggg gctgccttat ttaaagtggt tgtttatgat tcttatacta 350

atttatacaa agatattaag gccctgttca ttaagaaatt gttcccttcc 400

cctgtgttca atgtttgtaa agattgttct gtgtaaatat gtctttataa 450

taaacagtta aaagctgaaa aaaaaaaaaa aaaaaaaaaa aaaa 494

<210> SEQ ID NO 30
<211> LENGTH: 73
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 30

Met Leu Leu Leu Thr Leu Leu Leu Leu Leu Leu Lys Gly
1 5 10 15

Ser Cys Leu Glu Trp Gly Leu Val Gly Ala Gln Lys Val Ser Ser
20 25 30

Ala Thr Asp Ala Pro Ile Arg Asp Trp Ala Phe Phe Pro Pro Ser
35 40 45

Phe Leu Cys Leu Leu Pro His Arg Pro Ala Met Thr Cys Ser Gln
50 55 60

Ala Gln Pro Arg Gly Glu Gly Glu Lys Val Gly Asp Gly
65 70

<210> SEQ ID NO 31
<211> LENGTH: 1660

-continued

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 31

gtttgaattc cttcaactat acccacagtc caaaagcaga ctactgtgt	50
cccaggctac cagttcctcc aagcaagtca tttcccttat ttaaccgatg	100
tgtccctcaa acacctgagt gctactccct atttgcattc gttttgataa	150
atgatgttga caccctccac cgaattctaa gtggaatcat gtcgggaaga	200
gatacaatcc ttggcctgtg tatcctcgca ttacgcttgt ctttggccat	250
gatgtttacc ttcagattca tcaccaccct tctgggtcac attttcattt	300
cattgggttat tttgggattg ttgtttgtct gcggtgtttt atggtggctg	350
tattatgact ataccaacga cctcagcata gaattggaca cagaaaggga	400
aaatatgaag tgcgtgctgg ggtttgctat cgtatccaca ggcacacgg	450
cagtgcctgt cgtcttgatt tttgttctca gaaagagaat aaaattgaca	500
gttagagcttt tccaaatcac aaataaagcc atcagcagtg ctcccttcct	550
gctgttccag cactgtgga catttgccat cctcattttc tcttggttcc	600
tctgggtggc tgtgctgctg agcctgggaa ctgcaggagc tgcccagggt	650
atggaagggc gccaaagtga atataagccc ctttcgggca ttcggtacat	700
gtggctgtac catttaattg gcctcatctg gactagttaa ttcattcctg	750
cgtgccagca aatgactata gctggggcag tggttacttg ttatttcaac	800
agaagtaaaa atgatcctcc tgatcatccc atcctttcgt ctctctccat	850
tctcttcttc taccatcaag gaaccgttgt gaaagggta tttttaatct	900
ctgtgggtgag gattccgaga atcattgtca tgtacatgca aaacgcactg	950
aaagaacagc agcatggtgc attgtccagg tacctgttcc gatgctgcta	1000
ctgctgtttc tgggtgtctg acaaatacct gctccatctc aaccagaatg	1050
catatactac aactgctatt aatgggacag atttctgtac atcagcaaaa	1100
gatgcattca aaatcttgtc caagaactca agtcacttta catctattaa	1150
ctgctttgga gacttcataa tttttctagg aaagggtta gtggtgtgtt	1200
tactgtttt tggaggactc atggctttta actacaatcg ggcattccag	1250
gtgtgggcag tccctctgtt attggtagct tttttgcct acttagtagc	1300
ccatagtttt ttatctgtgt ttgaaactgt gctggatgca cttttcctgt	1350
gttttgcctg tgatctggaa acaaatgatg gatcgtcaga aaagccctac	1400
tttatggatc aagaatttct gagtttcgta aaaaggagca acaaattaaa	1450
caatgcaagg gcacagcagg acaagcactc attaaggaat gaggagggaa	1500
cagaactcca ggccattgtg agatagatac ccatttaggt atctgtacct	1550
ggaaaacatt tccttctaag agccatttac agaatagaag atgagaccac	1600
tagagaaaaa ttagtgaatt tttttttaa agacctata aacctattc	1650
ttcctcaaaa	1660

<210> SEQ ID NO 32

<211> LENGTH: 445

-continued

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 32

```

Met Ser Gly Arg Asp Thr Ile Leu Gly Leu Cys Ile Leu Ala Leu
 1           5           10           15

Ala Leu Ser Leu Ala Met Met Phe Thr Phe Arg Phe Ile Thr Thr
 20           25           30

Leu Leu Val His Ile Phe Ile Ser Leu Val Ile Leu Gly Leu Leu
 35           40           45

Phe Val Cys Gly Val Leu Trp Trp Leu Tyr Tyr Asp Tyr Thr Asn
 50           55           60

Asp Leu Ser Ile Glu Leu Asp Thr Glu Arg Glu Asn Met Lys Cys
 65           70           75

Val Leu Gly Phe Ala Ile Val Ser Thr Gly Ile Thr Ala Val Leu
 80           85           90

Leu Val Leu Ile Phe Val Leu Arg Lys Arg Ile Lys Leu Thr Val
 95           100          105

Glu Leu Phe Gln Ile Thr Asn Lys Ala Ile Ser Ser Ala Pro Phe
 110          115          120

Leu Leu Phe Gln Pro Leu Trp Thr Phe Ala Ile Leu Ile Phe Phe
 125          130          135

Trp Val Leu Trp Val Ala Val Leu Leu Ser Leu Gly Thr Ala Gly
 140          145          150

Ala Ala Gln Val Met Glu Gly Gly Gln Val Glu Tyr Lys Pro Leu
 155          160          165

Ser Gly Ile Arg Tyr Met Trp Ser Tyr His Leu Ile Gly Leu Ile
 170          175          180

Trp Thr Ser Glu Phe Ile Leu Ala Cys Gln Gln Met Thr Ile Ala
 185          190          195

Gly Ala Val Val Thr Cys Tyr Phe Asn Arg Ser Lys Asn Asp Pro
 200          205          210

Pro Asp His Pro Ile Leu Ser Ser Leu Ser Ile Leu Phe Phe Tyr
 215          220          225

His Gln Gly Thr Val Val Lys Gly Ser Phe Leu Ile Ser Val Val
 230          235          240

Arg Ile Pro Arg Ile Ile Val Met Tyr Met Gln Asn Ala Leu Lys
 245          250          255

Glu Gln Gln His Gly Ala Leu Ser Arg Tyr Leu Phe Arg Cys Cys
 260          265          270

Tyr Cys Cys Phe Trp Cys Leu Asp Lys Tyr Leu Leu His Leu Asn
 275          280          285

Gln Asn Ala Tyr Thr Thr Thr Ala Ile Asn Gly Thr Asp Phe Cys
 290          295          300

Thr Ser Ala Lys Asp Ala Phe Lys Ile Leu Ser Lys Asn Ser Ser
 305          310          315

His Phe Thr Ser Ile Asn Cys Phe Gly Asp Phe Ile Ile Phe Leu
 320          325          330

Gly Lys Val Leu Val Val Cys Phe Thr Val Phe Gly Gly Leu Met
 335          340          345

Ala Phe Asn Tyr Asn Arg Ala Phe Gln Val Trp Ala Val Pro Leu
 350          355          360

```

-continued

Leu Leu Val Ala Phe Phe Ala Tyr Leu Val Ala His Ser Phe Leu
365 370 375
Ser Val Phe Glu Thr Val Leu Asp Ala Leu Phe Leu Cys Phe Ala
380 385 390
Val Asp Leu Glu Thr Asn Asp Gly Ser Ser Glu Lys Pro Tyr Phe
395 400 405
Met Asp Gln Glu Phe Leu Ser Phe Val Lys Arg Ser Asn Lys Leu
410 415 420
Asn Asn Ala Arg Ala Gln Gln Asp Lys His Ser Leu Arg Asn Glu
425 430 435
Glu Gly Thr Glu Leu Gln Ala Ile Val Arg
440 445

<210> SEQ ID NO 33
<211> LENGTH: 2773
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 33

gttcgattag ctcctctgag aagaagagaa aaggttcttg gacctctccc	50
tgtttcttcc ttagaataat ttgtatggga tttgtgatgc aggaaagcct	100
aagggaaaaa gaatattcat tctgtgtggt gaaaaatttt tgaaaaaaa	150
attgccttct tcaaacaagg gtgtcattct gatatttatg aggactgttg	200
ttctcactat gaaggcatct gttattgaaa tgttccttgt tttgctggtg	250
actggagtac attcaaacaa agaaacggca aagaagatta aaaggcccaa	300
gttcactgtg cctcagatca actgcgatgt caaagccgga aagatcatcg	350
atcctgagtt cattgtgaaa tgtccagcag gatgccaaga ccccaaatac	400
catgtttatg gactgacgt gtatgcatcc tactccagtg tgtgtggcgc	450
tgccgtacac agtggtgtgc ttgataatcc aggagggaaa atacttgttc	500
ggaagggttc tggacagtct gggtacaaag ggagttatcc caacggtgtc	550
caatcgttat ccctaccacg atggagagaa tcctttatcg tcttagaaag	600
taaacccaaa aagggtgtaa cctaccatc agctcttaca tactcatcat	650
cgaaaagtcc agctgcccaa gcaggtgaga ccacaaaagc ctatcagagg	700
ccacctattc cagggacaac tgcacagccg gtcactctga tgcagcttct	750
ggctgtcact gtagctgtgg ccacccccac caccttgcca aggccatccc	800
cttctgtctg ttctaccacc agcatcccca gaccacaatc agtgggccac	850
aggagccagg agatggatct ctggtccact gccacctaca caagcagcca	900
aaacaggccc agagctgata caggtatcca aaggcaagat ccttcaggag	950
ctgccttcca gaaacctggt ggagcggatg tcagcctggg acttgttcca	1000
aaagaagaat tgagcacaca gtctttggag ccagtatccc tgggagatcc	1050
aaactgcaaa attgacttgt cgtttttaat tgatggggagc accagcattg	1100
gcaaacggcg attccgaatc cagaagcagc tcctggctga tgttgcccaa	1150
gctcttgaca ttggccctgc cggtcactg atgggtgttg tccagtatgg	1200
agacaacctt gctactcact ttaacctcaa gacacacacg aattctcgag	1250

-continued

```
atctgaagac agccatagag aaaattactc agagaggagg actttctaata 1300
gtaggtcggg ccatctcctt tgtgaccaag aacttctttt ccaaagccaa 1350
tggaacaga agcggggctc ccaatgtggt ggtggtgatg gtggatggct 1400
ggccacgga caaagtggag gaggcttcaa gacttgcgag agagtcagga 1450
atcaacattt tcttcacac cattgaaggt gctgctgaaa atgagaagca 1500
gtatgtggtg gagcccaact ttgcaaaaca ggccgtgtgc agaacaacg 1550
gcttctactc gctccacgtg cagagctggt ttggcctcca caagaccctg 1600
cagcctctgg tgaagcgggt ctgcgacact gaccgcctgg cctgcagcaa 1650
gacctgcttg aactcggctg acattggctt cgtcatcgac ggctccagca 1700
gtgtggggac gggcaacttc cgcaccgtcc tccagtttgt gaccaacctc 1750
accaagagt ttgagatttc cgacacggac acgcgcacgc gggccgtgca 1800
gtacacctac gaacagcggc tggagtttgg gttcgacaag tacagcagca 1850
agcctgacat cctcaacgcc atcaagaggg tgggctactg gagtgggtggc 1900
accagcacgg gggctgccat caacttcgcc ctggagcagc tcttcaagaa 1950
gtccaagccc aacaagagga agttaatgat cctcatcacc gacgggaggt 2000
cctacgacga cgtccggatc ccagccatgg ctgcccacat gaagggagtg 2050
atcacctatg cgataggcgt tgctggggct gcccaagagg agctagaagt 2100
cattgccact caccocgcca gagaccactc cttctttgtg gacgagtttg 2150
acaacctcca tcagtatgtc ccaggatca tccagaacat ttgtacagag 2200
ttcaactcac agcctcggaa ctgaattcag agcaggcaga gcaccagcaa 2250
gtgctgcttt actaactgac gtgttggaac accccaccgc ttaatggggc 2300
acgcacgggt catcaagtct tgggcagggc atggagaaac aaatgtcttg 2350
ttattattct ttgccatcat gctttttcat attccaaaac ttggagttac 2400
aaagatgatc acaaacgtat agaatgagcc aaaaggctac atcatgttga 2450
gggtgctgga gattttacat ttgacaatt gttttcaaaa taaatgttcg 2500
gaatacagtg cagcccttac gacaggctta cgtagagctt ttgtgagatt 2550
tttaagttgt tatttctgat ttgaactctg taaccctcag caagtttcat 2600
ttttgtcatg acaatgtagg aattgctgaa ttaaatgttt agaaggatga 2650
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2700
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2750
aaaaaaaaa aaaaaaaaaa aag 2773
```

<210> SEQ ID NO 34

<211> LENGTH: 678

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 34

```
Met Arg Thr Val Val Leu Thr Met Lys Ala Ser Val Ile Glu Met
  1             5             10             15
```

```
Phe Leu Val Leu Leu Val Thr Gly Val His Ser Asn Lys Glu Thr
      20             25             30
```

-continued

Ala Lys Lys Ile Lys Arg Pro Lys Phe Thr Val Pro Gln Ile Asn	35	40	45
Cys Asp Val Lys Ala Gly Lys Ile Ile Asp Pro Glu Phe Ile Val	50	55	60
Lys Cys Pro Ala Gly Cys Gln Asp Pro Lys Tyr His Val Tyr Gly	65	70	75
Thr Asp Val Tyr Ala Ser Tyr Ser Ser Val Cys Gly Ala Ala Val	80	85	90
His Ser Gly Val Leu Asp Asn Ser Gly Gly Lys Ile Leu Val Arg	95	100	105
Lys Val Ala Gly Gln Ser Gly Tyr Lys Gly Ser Tyr Ser Asn Gly	110	115	120
Val Gln Ser Leu Ser Leu Pro Arg Trp Arg Glu Ser Phe Ile Val	125	130	135
Leu Glu Ser Lys Pro Lys Lys Gly Val Thr Tyr Pro Ser Ala Leu	140	145	150
Thr Tyr Ser Ser Ser Lys Ser Pro Ala Ala Gln Ala Gly Glu Thr	155	160	165
Thr Lys Ala Tyr Gln Arg Pro Pro Ile Pro Gly Thr Thr Ala Gln	170	175	180
Pro Val Thr Leu Met Gln Leu Leu Ala Val Thr Val Ala Val Ala	185	190	195
Thr Pro Thr Thr Leu Pro Arg Pro Ser Pro Ser Ala Ala Ser Thr	200	205	210
Thr Ser Ile Pro Arg Pro Gln Ser Val Gly His Arg Ser Gln Glu	215	220	225
Met Asp Leu Trp Ser Thr Ala Thr Tyr Thr Ser Ser Gln Asn Arg	230	235	240
Pro Arg Ala Asp Pro Gly Ile Gln Arg Gln Asp Pro Ser Gly Ala	245	250	255
Ala Phe Gln Lys Pro Val Gly Ala Asp Val Ser Leu Gly Leu Val	260	265	270
Pro Lys Glu Glu Leu Ser Thr Gln Ser Leu Glu Pro Val Ser Leu	275	280	285
Gly Asp Pro Asn Cys Lys Ile Asp Leu Ser Phe Leu Ile Asp Gly	290	295	300
Ser Thr Ser Ile Gly Lys Arg Arg Phe Arg Ile Gln Lys Gln Leu	305	310	315
Leu Ala Asp Val Ala Gln Ala Leu Asp Ile Gly Pro Ala Gly Pro	320	325	330
Leu Met Gly Val Val Gln Tyr Gly Asp Asn Pro Ala Thr His Phe	335	340	345
Asn Leu Lys Thr His Thr Asn Ser Arg Asp Leu Lys Thr Ala Ile	350	355	360
Glu Lys Ile Thr Gln Arg Gly Gly Leu Ser Asn Val Gly Arg Ala	365	370	375
Ile Ser Phe Val Thr Lys Asn Phe Phe Ser Lys Ala Asn Gly Asn	380	385	390
Arg Ser Gly Ala Pro Asn Val Val Val Val Met Val Asp Gly Trp	395	400	405

-continued

Pro	Thr	Asp	Lys	Val	Glu	Glu	Ala	Ser	Arg	Leu	Ala	Arg	Glu	Ser	
				410					415					420	
Gly	Ile	Asn	Ile	Phe	Phe	Ile	Thr	Ile	Glu	Gly	Ala	Ala	Glu	Asn	
				425					430					435	
Glu	Lys	Gln	Tyr	Val	Val	Glu	Pro	Asn	Phe	Ala	Asn	Lys	Ala	Val	
				440					445					450	
Cys	Arg	Thr	Asn	Gly	Phe	Tyr	Ser	Leu	His	Val	Gln	Ser	Trp	Phe	
				455					460					465	
Gly	Leu	His	Lys	Thr	Leu	Gln	Pro	Leu	Val	Lys	Arg	Val	Cys	Asp	
				470					475					480	
Thr	Asp	Arg	Leu	Ala	Cys	Ser	Lys	Thr	Cys	Leu	Asn	Ser	Ala	Asp	
				485					490					495	
Ile	Gly	Phe	Val	Ile	Asp	Gly	Ser	Ser	Ser	Val	Gly	Thr	Gly	Asn	
				500					505					510	
Phe	Arg	Thr	Val	Leu	Gln	Phe	Val	Thr	Asn	Leu	Thr	Lys	Glu	Phe	
				515					520					525	
Glu	Ile	Ser	Asp	Thr	Asp	Thr	Arg	Ile	Gly	Ala	Val	Gln	Tyr	Thr	
				530					535					540	
Tyr	Glu	Gln	Arg	Leu	Glu	Phe	Gly	Phe	Asp	Lys	Tyr	Ser	Ser	Lys	
				545					550					555	
Pro	Asp	Ile	Leu	Asn	Ala	Ile	Lys	Arg	Val	Gly	Tyr	Trp	Ser	Gly	
				560					565					570	
Gly	Thr	Ser	Thr	Gly	Ala	Ala	Ile	Asn	Phe	Ala	Leu	Glu	Gln	Leu	
				575					580					585	
Phe	Lys	Lys	Ser	Lys	Pro	Asn	Lys	Arg	Lys	Leu	Met	Ile	Leu	Ile	
				590					595					600	
Thr	Asp	Gly	Arg	Ser	Tyr	Asp	Asp	Val	Arg	Ile	Pro	Ala	Met	Ala	
				605					610					615	
Ala	His	Leu	Lys	Gly	Val	Ile	Thr	Tyr	Ala	Ile	Gly	Val	Ala	Trp	
				620					625					630	
Ala	Ala	Gln	Glu	Glu	Leu	Glu	Val	Ile	Ala	Thr	His	Pro	Ala	Arg	
				635					640					645	
Asp	His	Ser	Phe	Phe	Val	Asp	Glu	Phe	Asp	Asn	Leu	His	Gln	Tyr	
				650					655					660	
Val	Pro	Arg	Ile	Ile	Gln	Asn	Ile	Cys	Thr	Glu	Phe	Asn	Ser	Gln	
				665					670					675	

Pro Arg Asn

<210> SEQ ID NO 35
<211> LENGTH: 2095
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 35

ccgagcacag gagattgcct gcgttagga ggtggctgcg ttgtgggaaa	50
agctatcaag gaagaaattg ccaaaccatg tctttttttc tgttttcaga	100
gtagtccaca acagatctga gtgttttaat taagcatgga atacagaaaa	150
caacaaaaaa cttaagcttt aatttcacatc ggaattccac agttttctta	200
gtcccttgga cccggttgac ctgttggtcgc ttcccgctgg ctgctctatc	250
acgtggtgct ctccgactac tcaccccgag tgtaaagaac ctcgggtcgc	300

-continued

cggtgcttctg agctgctgtg gatggcctcg gctctctgga ctgtccttcc	350
gagtaggatg tcaactgagat ccctcaaatg gagcctcctg ctgctgtcac	400
tcctgagttt ctttgtgatg tggtaacctca gccttcccca ctacaatgtg	450
atagaacgcg tgaactggat gtacttctat gagtatgagc cgatttacag	500
acaagacttt cacttcacac ttcgagagca ttcaaactgc tctcatcaaa	550
atccatttct ggtcattctg gtgacctccc acccttcaga tgtgaaagcc	600
aggcaggcca ttagagttac ttggggtgaa aaaaagtctt ggtggggata	650
tgaggttctt acatttttct tattaggcca agaggctgaa aaggaagaca	700
aatgtttggc attgtcctta gaggatgaac accttcttta tggtgacata	750
atccgacaag attttttaga cacatataat aacctgacct tgaaaaccat	800
tatggcattc aggtgggtaa ctgagttttg cccaatgcc aagtacgtaa	850
tgaagacaga cactgatgtt ttcatacaata ctggcaattt agtgaagtat	900
cttttaaacc taaaccactc agagaagttt ttcacaggtt atcctcta	950
tgataattat tcctatagag gattttacca aaaaacccat atttcttacc	1000
aggagtatcc tttcaagggt ttccctccat actgcagtgg gttgggttat	1050
ataatgtcca gagatttggg gccaaaggatc tatgaaatga tgggtcacgt	1100
aaaaccatc aagtttgaag atgtttatgt cgggatctgt ttgaatttat	1150
taaaagtga cattcatatt ccagaagaca caaatctttt ctttctatat	1200
agaatccatt tggatgtctg tcaactgaga cgtgtgattg cagccatgg	1250
cttttcttcc aaggagatca tcaacttttg gcaggtcatt ctaaggaaca	1300
ccacatgcca ttattaactt cacattctac aaaaagccta gaaggacagg	1350
ataccttgtg gaaagtgtta aataaagtag gtactgtgga aaattcatgg	1400
ggaggtcagt gtgctggcct aactgaact gaaactcatg aaaaaccag	1450
actggagact ggagggttac acttgtgatt tattagtcag gcccttcaaa	1500
gatgatattg ggaggaatta aatataaagg aattggaggt ttttgctaaa	1550
gaaattaata ggaccaaaaca atttgacat gtcatctgt agactagaat	1600
ttcttaaaag ggtgttactg agttataagc tcaactaggct gtaaaaaaa	1650
aacaatgtag agttttatct attgaacaat gtactgactt gaaggttttg	1700
tgtatatctt atgtggatta ccaatttaaa aatatatgta gttctgtgtc	1750
aaaaaacttc ttactgaag ttatactgaa caaaatttta cctgtttttg	1800
gtcatttata aagtacttca agatgttgca gtatttcaca gttattatta	1850
tttaaaatta cttcaacttt gtgtttttta atgttttgac gatttcaata	1900
caagataaaa aggatagtga atcattcttt acatgcaaac attttccagt	1950
tacttaactg atcagtttat tattgataca tcaactccatt aatgtaaagt	2000
cataggtcatt tattgcatat cagtaatctc ttggactttg ttaaatattt	2050
tactgtggta atatagagaa gaattaaagc aagaaaatct gaaaa	2095

<210> SEQ ID NO 36

<211> LENGTH: 331

<212> TYPE: PRT

-continued

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 36

Met Ala Ser Ala Leu Trp Thr Val Leu Pro Ser Arg Met Ser Leu
1 5 10 15
Arg Ser Leu Lys Trp Ser Leu Leu Leu Leu Ser Leu Leu Ser Phe
20 25 30
Phe Val Met Trp Tyr Leu Ser Leu Pro His Tyr Asn Val Ile Glu
35 40 45
Arg Val Asn Trp Met Tyr Phe Tyr Glu Tyr Glu Pro Ile Tyr Arg
50 55 60
Gln Asp Phe His Phe Thr Leu Arg Glu His Ser Asn Cys Ser His
65 70 75
Gln Asn Pro Phe Leu Val Ile Leu Val Thr Ser His Pro Ser Asp
80 85 90
Val Lys Ala Arg Gln Ala Ile Arg Val Thr Trp Gly Glu Lys Lys
95 100 105
Ser Trp Trp Gly Tyr Glu Val Leu Thr Phe Phe Leu Leu Gly Gln
110 115 120
Glu Ala Glu Lys Glu Asp Lys Met Leu Ala Leu Ser Leu Glu Asp
125 130 135
Glu His Leu Leu Tyr Gly Asp Ile Ile Arg Gln Asp Phe Leu Asp
140 145 150
Thr Tyr Asn Asn Leu Thr Leu Lys Thr Ile Met Ala Phe Arg Trp
155 160 165
Val Thr Glu Phe Cys Pro Asn Ala Lys Tyr Val Met Lys Thr Asp
170 175 180
Thr Asp Val Phe Ile Asn Thr Gly Asn Leu Val Lys Tyr Leu Leu
185 190 195
Asn Leu Asn His Ser Glu Lys Phe Phe Thr Gly Tyr Pro Leu Ile
200 205 210
Asp Asn Tyr Ser Tyr Arg Gly Phe Tyr Gln Lys Thr His Ile Ser
215 220 225
Tyr Gln Glu Tyr Pro Phe Lys Val Phe Pro Pro Tyr Cys Ser Gly
230 235 240
Leu Gly Tyr Ile Met Ser Arg Asp Leu Val Pro Arg Ile Tyr Glu
245 250 255
Met Met Gly His Val Lys Pro Ile Lys Phe Glu Asp Val Tyr Val
260 265 270
Gly Ile Cys Leu Asn Leu Leu Lys Val Asn Ile His Ile Pro Glu
275 280 285
Asp Thr Asn Leu Phe Phe Leu Tyr Arg Ile His Leu Asp Val Cys
290 295 300
Gln Leu Arg Arg Val Ile Ala Ala His Gly Phe Ser Ser Lys Glu
305 310 315
Ile Ile Thr Phe Trp Gln Val Met Leu Arg Asn Thr Thr Cys His
320 325 330

Tyr

<210> SEQ ID NO 37

<211> LENGTH: 2846

<212> TYPE: DNA

-continued

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 37

cgctcgggca ccagccgcgg caaggatgga gctggggtgc tggacgcagt	50
tggggctcac tttctctcag ctcccttctca tctcgtcctt gccaaagagag	100
tacacagtca ttaatgaagc ctgccctgga gcagagtgga atatcatgtg	150
tcgggagtg c tgtgaatatg atcagattga gtgcgtctgc cccggaaaga	200
gggaagtctg gggttatacc atcccttgct gcaggaatga ggagaatgag	250
tgtgactcct gcctgatcca ccaggttgt accatctttg aaaactgcaa	300
gagctgcgca aatggctcat ggggggttac cttggatgac ttctatgtga	350
aggggttcta ctgtgcagag tgccgagcag gctggtagcg aggagactgc	400
atgcgatgtg gccaggttct gcgagcccca aagggtcaga ttttgttgga	450
aagctatccc ctaaagctc actgtgaatg gaccattcat gctaaacctg	500
ggtttgtcat ccaactaaga tttgtcatgt tgagtctgga gtttgactac	550
atgtgccagt atgactatgt tgaggttcgt gatggagaca accgcgatgg	600
ccagatcac aagcgtgtct gtggcaacga gcggccagct cctatccaga	650
gcataaggatc ctactccac gtccctctcc actccgatgg ctccaagaat	700
tttgacggtt tccatgccat ttatgaggag atcacagcat gctcctcatc	750
cccttgtttc catgacggca cgtgcgtcct tgacaaggct ggatcttaca	800
agtggtcctg cttggcaggc tatactgggc agcgtgtga aaatctcctt	850
gaagaaagaa actgctcaga ccctgggggc ccagtcaatg ggtaccagaa	900
aataacaggg ggcctgggc ttatcaacgg acgccatgct aaaattggca	950
ccgtgggtgc tttcttttgt aacaactcct atgttcttag tggcaatgag	1000
aaaagaactt gccagcagaa tggagagtgg tcagggaaac agcccatctg	1050
cataaaagcc tgccgagaac caaagatttc agacctggtg agaaggagag	1100
ttcttccgat gcaggttcag tcaagggaga caccattaca ccagctatac	1150
tcagcggcct tcagcaagca gaaactgcag agtgccccta ccaagaagcc	1200
agcccttccc tttggagatc tgcccatggg ataccaacat ctgcataccc	1250
agctccagta tgagtgcac tcaccttct accccgcct gggcagcagc	1300
aggaggacat gtctgaggac tgggaagtgg agtgggcggg caccatcctg	1350
catccctatc tgcgggaaaa ttgagaacat cactgctcca aagacccaag	1400
ggttgcgctg gccgtggcag gcagccatct acaggaggac cagcggggtg	1450
catgacggca gcctacacaa gggagcgtgg ttcctagtct gcagcgggtg	1500
cctggtgaat gagcgactg tgggtggtgg tgccactgt gttactgacc	1550
tggggaaggt caccatgatc aagacagcag acctgaaagt tgttttgggg	1600
aaattctacc gggatgatga ccgggatgag aagaccatcc agagcctaca	1650
gatttctgct atcattctgc atcccaacta tgaccccatc ctgcttgatg	1700
ctgacatcgc catcctgaag ctccatgaca aggcccgat cagcacccga	1750
gtccagccca tctgcctcgc tgccagtcgg gatctcagca ctcccttcca	1800

-continued

```

ggagtcccac atcactgtgg ctggctggaa tgtcctggca gacgtgagga      1850
gccctggctt caagaacgac aactgcgct ctggggtggt cagtgtggtg      1900
gactcgctgc tgtgtgagga gcagcatgag gaccatggca tcccagtgag      1950
tgtcactgat aacatgttct gtgccagctg ggaaccact gcccttctg      2000
atatctgcac tgcagagaca ggaggcatcg cggctgtgtc cttcccggga      2050
cgagcatctc ctgagccacg ctggcatctg atgggactgg tcagctggag      2100
ctatgataaa acatgcagcc acaggctctc cactgccttc accaagggtg      2150
tgccttttaa agactggatt gaaagaaata tgaatgaac catgctcatg      2200
cactccttga gaagtgttct tgtatatccg tctgtacgtg tgtcattgctg      2250
tgaagcagtg tgggcctgaa gtgtgatttg gcctgtgaac ttggctgtgc      2300
cagggcttct gacttcaggg acaaaactca gtgaagggtg agtagacctc      2350
cattgctggt aggctgatgc cgcgtccact actaggacag ccaattggaa      2400
gatgccaggg cttgcaagaa gtaagtttct tcaaagaaga ccatatacaa      2450
aacctctcca ctccactgac ctgggtgtct tcccactt tcagttatac      2500
gaatgccatc agcttgacca gggaagatct gggcttcatg aggccccttt      2550
tgaggctctc aagtctaga gagctgcctg tgggacagcc cagggcagca      2600
gagctgggat gtggtgcatg ctttgtgta catggccaca gtacagtctg      2650
gtccttttcc ttcccatct cttgtacaca ttttaataaa ataagggttg      2700
gcttctgaac tacaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      2750
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      2800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa      2846

```

<210> SEQ ID NO 38

<211> LENGTH: 720

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 38

```

Met Glu Leu Gly Cys Trp Thr Gln Leu Gly Leu Thr Phe Leu Gln
 1             5             10            15
Leu Leu Leu Ile Ser Ser Leu Pro Arg Glu Tyr Thr Val Ile Asn
                20            25            30
Glu Ala Cys Pro Gly Ala Glu Trp Asn Ile Met Cys Arg Glu Cys
                35            40            45
Cys Glu Tyr Asp Gln Ile Glu Cys Val Cys Pro Gly Lys Arg Glu
                50            55            60
Val Val Gly Tyr Thr Ile Pro Cys Cys Arg Asn Glu Glu Asn Glu
                65            70            75
Cys Asp Ser Cys Leu Ile His Pro Gly Cys Thr Ile Phe Glu Asn
                80            85            90
Cys Lys Ser Cys Arg Asn Gly Ser Trp Gly Gly Thr Leu Asp Asp
                95            100           105
Phe Tyr Val Lys Gly Phe Tyr Cys Ala Glu Cys Arg Ala Gly Trp
                110           115           120
Tyr Gly Gly Asp Cys Met Arg Cys Gly Gln Val Leu Arg Ala Pro
                125           130           135

```

-continued

Lys Gly Gln Ile	Leu Leu Glu Ser Tyr	Pro Leu Asn Ala His Cys	
140		145	150
Glu Trp Thr Ile	His Ala Lys Pro Gly	Phe Val Ile Gln Leu Arg	
155		160	165
Phe Val Met Leu	Ser Leu Glu Phe Asp	Tyr Met Cys Gln Tyr Asp	
170		175	180
Tyr Val Glu Val	Arg Asp Gly Asp Asn	Arg Asp Gly Gln Ile Ile	
185		190	195
Lys Arg Val Cys	Gly Asn Glu Arg Pro	Ala Pro Ile Gln Ser Ile	
200		205	210
Gly Ser Ser Leu	His Val Leu Phe His	Ser Asp Gly Ser Lys Asn	
215		220	225
Phe Asp Gly Phe	His Ala Ile Tyr Glu	Glu Ile Thr Ala Cys Ser	
230		235	240
Ser Ser Pro Cys	Phe His Asp Gly Thr	Cys Val Leu Asp Lys Ala	
245		250	255
Gly Ser Tyr Lys	Cys Ala Cys Leu Ala	Gly Tyr Thr Gly Gln Arg	
260		265	270
Cys Glu Asn Leu	Leu Glu Glu Arg Asn	Cys Ser Asp Pro Gly Gly	
275		280	285
Pro Val Asn Gly	Tyr Gln Lys Ile Thr	Gly Gly Pro Gly Leu Ile	
290		295	300
Asn Gly Arg His	Ala Lys Ile Gly Thr	Val Val Ser Phe Phe Cys	
305		310	315
Asn Asn Ser Tyr	Val Leu Ser Gly Asn	Glu Lys Arg Thr Cys Gln	
320		325	330
Gln Asn Gly Glu	Trp Ser Gly Lys Gln	Pro Ile Cys Ile Lys Ala	
335		340	345
Cys Arg Glu Pro	Lys Ile Ser Asp Leu	Val Arg Arg Arg Val Leu	
350		355	360
Pro Met Gln Val	Gln Ser Arg Glu Thr	Pro Leu His Gln Leu Tyr	
365		370	375
Ser Ala Ala Phe	Ser Lys Gln Lys Leu	Gln Ser Ala Pro Thr Lys	
380		385	390
Lys Pro Ala Leu	Pro Phe Gly Asp Leu	Pro Met Gly Tyr Gln His	
395		400	405
Leu His Thr Gln	Leu Gln Tyr Glu Cys	Ile Ser Pro Phe Tyr Arg	
410		415	420
Arg Leu Gly Ser	Ser Arg Arg Thr Cys	Leu Arg Thr Gly Lys Trp	
425		430	435
Ser Gly Arg Ala	Pro Ser Cys Ile Pro	Ile Cys Gly Lys Ile Glu	
440		445	450
Asn Ile Thr Ala	Pro Lys Thr Gln Gly	Leu Arg Trp Pro Trp Gln	
455		460	465
Ala Ala Ile Tyr	Arg Arg Thr Ser Gly	Val His Asp Gly Ser Leu	
470		475	480
His Lys Gly Ala	Trp Phe Leu Val Cys	Ser Gly Ala Leu Val Asn	
485		490	495
Glu Arg Thr Val	Val Val Ala Ala His	Cys Val Thr Asp Leu Gly	
500		505	510

-continued

Lys	Val	Thr	Met	Ile	Lys	Thr	Ala	Asp	Leu	Lys	Val	Val	Leu	Gly	
			515						520					525	
Lys	Phe	Tyr	Arg	Asp	Asp	Asp	Arg	Asp	Glu	Lys	Thr	Ile	Gln	Ser	
			530						535					540	
Leu	Gln	Ile	Ser	Ala	Ile	Ile	Leu	His	Pro	Asn	Tyr	Asp	Pro	Ile	
			545						550					555	
Leu	Leu	Asp	Ala	Asp	Ile	Ala	Ile	Leu	Lys	Leu	Leu	Asp	Lys	Ala	
			560						565					570	
Arg	Ile	Ser	Thr	Arg	Val	Gln	Pro	Ile	Cys	Leu	Ala	Ala	Ser	Arg	
			575						580					585	
Asp	Leu	Ser	Thr	Ser	Phe	Gln	Glu	Ser	His	Ile	Thr	Val	Ala	Gly	
			590						595					600	
Trp	Asn	Val	Leu	Ala	Asp	Val	Arg	Ser	Pro	Gly	Phe	Lys	Asn	Asp	
			605						610					615	
Thr	Leu	Arg	Ser	Gly	Val	Val	Ser	Val	Val	Asp	Ser	Leu	Leu	Cys	
			620						625					630	
Glu	Glu	Gln	His	Glu	Asp	His	Gly	Ile	Pro	Val	Ser	Val	Thr	Asp	
			635						640					645	
Asn	Met	Phe	Cys	Ala	Ser	Trp	Glu	Pro	Thr	Ala	Pro	Ser	Asp	Ile	
			650						655					660	
Cys	Thr	Ala	Glu	Thr	Gly	Gly	Ile	Ala	Ala	Val	Ser	Phe	Pro	Gly	
			665						670					675	
Arg	Ala	Ser	Pro	Glu	Pro	Arg	Trp	His	Leu	Met	Gly	Leu	Val	Ser	
			680						685					690	
Trp	Ser	Tyr	Asp	Lys	Thr	Cys	Ser	His	Arg	Leu	Ser	Thr	Ala	Phe	
			695						700					705	
Thr	Lys	Val	Leu	Pro	Phe	Lys	Asp	Trp	Ile	Glu	Arg	Asn	Met	Lys	
			710						715					720	

<210> SEQ ID NO 39
<211> LENGTH: 2571
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 39

ggttctctaca tcctctcatc tgagaatcag agagcataat cttcttacgg	50
gcccgtgatt tattaacgtg gcttaatctg aaggttctca gtcaaattct	100
ttgtgatcta ctgattgtgg gggcatggca aggtttgctt aaaggagctt	150
ggctggtttg ggcccttgta gctgacagaa ggtggccagg gagaatgcag	200
cacactgctc ggagaatgaa ggcgcttctg ttgctggtct tgccttggtc	250
cagtctgct aactacattg acaatgtggg caacctgcac ttctgtatt	300
cagaactctg taaaggtgcc tcccactacg gcctgaccaa agataggaag	350
aggcgctcac aagatggctg tccagacggc tgtgcgagcc tcacagccac	400
ggctccctcc ccagaggttt ctgcagctgc caccatctcc ttaatgacag	450
acgagcctgg cctagacaac cctgcctacg tgtcctcggc agaggacggg	500
cagccagcaa tcagcccagt ggactctggc cggagcaacc gaactagggc	550
acggcccttt gagagatcca ctattagaag cagatcattt aaaaaataa	600
atcgagcttt gagtgttctt cgaaggacaa agagcgggag tgcagttgcc	650

-continued

aaccatgccg accagggcag ggaaaattct gaaaacacca ctgccctga	700
agtctttcca aggttgtacc acctgattcc agatggtgaa attaccagca	750
tcaagatcaa tcgagtagat cccagtgaag gcctctctat taggctggtg	800
ggaggtagcg aaacccact ggtccatctc attatccaac acatttatcg	850
tgatggggtg atcgccagag acggccggct actgccagga gacatcattc	900
taaaggtaa cgggatggac atcagcaatg tcctcacia ctacgctgtg	950
cgtctcctcg ggcagccctg ccaggtgctg tggtgactg tgatgcgtga	1000
acagaagttc cgcagcagga acaatggaca ggccccgat gcctacagac	1050
cccagatga cagctttcat gtgattctca acaaaagtag ccccgaggag	1100
cagcttgga taaaactggt gcgcaagggt gatgagcctg gggttttcat	1150
cttcaatgtg ctggatggcg gtgtggcata tcgacatggt cagcttgagg	1200
agaatgacgg tgtgttagcc atcaatggac atgatcttcg atatggcagc	1250
ccagaaagt cggctcatct gattcaggcc agtgaaagac gtgttcacct	1300
cgtcgtgtcc cgccagggtc ggcagcggag ccctgacatc ttccaggaag	1350
ccggctggaa cagcaatggc agctggctcc cagggccagg ggagaggagc	1400
aacactccca agccctcca tcctacaatt acttgatcat agaagtggtg	1450
aaatatcaa aaagaccccg gtgaatctct cggcatgacc gtcgcagggg	1500
gagcatcaca tagagaatgg gatttgctca tctatgtcat cagtgttgag	1550
cccgaggagg tcataagcag agatggaaga ataaaaacag gtgacatttt	1600
gttgaatgtg gatggggctg aactgacaga ggtcagccgg agtgaggcag	1650
tggcattatt gaaaagaaca tcctcctcga tagtactcaa agctttggaa	1700
gtcaagagt atgagcccca ggaagactgc agcagcccag cagccctgga	1750
ctccaaccac aacatggccc caccagtgta ctggtcccca tcctgggtca	1800
tggtgctgga attaccacgg tgcttgata actgtaaaga tattgtatta	1850
cgaagaaca cagctggaag tctgggcttc tgcattgtag gaggttatga	1900
agaatacaat ggaacaaac cttttttcat caaatccatt gttgaaggaa	1950
caccagcata caatgatgga agaattagat gtggtgatat tcttcttgct	2000
gtcaatggta gaagtacatc aggaatgata catgcttgct tggcaagact	2050
gctgaaagaa cttaaggaa gaattactct aactattggt tcttgccctg	2100
gcactttttt atagaatcaa tgatgggtca gaggaaca gaaaaatcac	2150
aaataggcta agaagttgaa aactatatt tatcttgta gtttttata	2200
ttaaagaaag aatacattgt aaaaatgtca ggaaaagtat gatcatctaa	2250
tgaaagccag ttacacctca gaaaatatga ttccaaaaaa attaaaacta	2300
ctagtttttt ttcagtgtgg aggatttctc attactctac aacattgttt	2350
atattttttc tattcaataa aaagccctaa aacaactaaa atgattgatt	2400
tgtatacccc actgaattca agctgattta aatttaaaat ttggtatatg	2450
ctgaagtctg ccaagggtac attatggcca tttttaattt acagctaaaa	2500
tattttttaa aatgcattgc tgagaaacgt tgctttcatc aaacaagaat	2550

-continued

aaatatatttt cagaagtta a

2571

<210> SEQ ID NO 40

<211> LENGTH: 632

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 40

Met Lys Ala Leu Leu Leu Val Leu Pro Trp Leu Ser Pro Ala
1 5 10 15

Asn Tyr Ile Asp Asn Val Gly Asn Leu His Phe Leu Tyr Ser Glu
20 25 30

Leu Cys Lys Gly Ala Ser His Tyr Gly Leu Thr Lys Asp Arg Lys
35 40 45

Arg Arg Ser Gln Asp Gly Cys Pro Asp Gly Cys Ala Ser Leu Thr
50 55 60

Ala Thr Ala Pro Ser Pro Glu Val Ser Ala Ala Ala Thr Ile Ser
65 70 75

Leu Met Thr Asp Glu Pro Gly Leu Asp Asn Pro Ala Tyr Val Ser
80 85 90

Ser Ala Glu Asp Gly Gln Pro Ala Ile Ser Pro Val Asp Ser Gly
95 100 105

Arg Ser Asn Arg Thr Arg Ala Arg Pro Phe Glu Arg Ser Thr Ile
110 115 120

Arg Ser Arg Ser Phe Lys Lys Ile Asn Arg Ala Leu Ser Val Leu
125 130 135

Arg Arg Thr Lys Ser Gly Ser Ala Val Ala Asn His Ala Asp Gln
140 145 150

Gly Arg Glu Asn Ser Glu Asn Thr Thr Ala Pro Glu Val Phe Pro
155 160 165

Arg Leu Tyr His Leu Ile Pro Asp Gly Glu Ile Thr Ser Ile Lys
170 175 180

Ile Asn Arg Val Asp Pro Ser Glu Ser Leu Ser Ile Arg Leu Val
185 190 195

Gly Gly Ser Glu Thr Pro Leu Val His Ile Ile Ile Gln His Ile
200 205 210

Tyr Arg Asp Gly Val Ile Ala Arg Asp Gly Arg Leu Leu Pro Gly
215 220 225

Asp Ile Ile Leu Lys Val Asn Gly Met Asp Ile Ser Asn Val Pro
230 235 240

His Asn Tyr Ala Val Arg Leu Leu Arg Gln Pro Cys Gln Val Leu
245 250 255

Trp Leu Thr Val Met Arg Glu Gln Lys Phe Arg Ser Arg Asn Asn
260 265 270

Gly Gln Ala Pro Asp Ala Tyr Arg Pro Arg Asp Asp Ser Phe His
275 280 285

Val Ile Leu Asn Lys Ser Ser Pro Glu Glu Gln Leu Gly Ile Lys
290 295 300

Leu Val Arg Lys Val Asp Glu Pro Gly Val Phe Ile Phe Asn Val
305 310 315

Leu Asp Gly Gly Val Ala Tyr Arg His Gly Gln Leu Glu Glu Asn
320 325 330

-continued

Asp	Arg	Val	Leu	Ala	Ile	Asn	Gly	His	Asp	Leu	Arg	Tyr	Gly	Ser
			335						340					345
Pro	Glu	Ser	Ala	Ala	His	Leu	Ile	Gln	Ala	Ser	Glu	Arg	Arg	Val
			350						355					360
His	Leu	Val	Val	Ser	Arg	Gln	Val	Arg	Gln	Arg	Ser	Pro	Asp	Ile
			365						370					375
Phe	Gln	Glu	Ala	Gly	Trp	Asn	Ser	Asn	Gly	Ser	Trp	Ser	Pro	Gly
			380						385					390
Pro	Gly	Glu	Arg	Ser	Asn	Thr	Pro	Lys	Pro	Leu	His	Pro	Thr	Ile
			395						400					405
Thr	Cys	His	Glu	Lys	Val	Val	Asn	Ile	Gln	Lys	Asp	Pro	Gly	Glu
			410						415					420
Ser	Leu	Gly	Met	Thr	Val	Ala	Gly	Gly	Ala	Ser	His	Arg	Glu	Trp
			425						430					435
Asp	Leu	Pro	Ile	Tyr	Val	Ile	Ser	Val	Glu	Pro	Gly	Gly	Val	Ile
			440						445					450
Ser	Arg	Asp	Gly	Arg	Ile	Lys	Thr	Gly	Asp	Ile	Leu	Leu	Asn	Val
			455						460					465
Asp	Gly	Val	Glu	Leu	Thr	Glu	Val	Ser	Arg	Ser	Glu	Ala	Val	Ala
			470						475					480
Leu	Leu	Lys	Arg	Thr	Ser	Ser	Ser	Ile	Val	Leu	Lys	Ala	Leu	Glu
			485						490					495
Val	Lys	Glu	Tyr	Glu	Pro	Gln	Glu	Asp	Cys	Ser	Ser	Pro	Ala	Ala
			500						505					510
Leu	Asp	Ser	Asn	His	Asn	Met	Ala	Pro	Pro	Ser	Asp	Trp	Ser	Pro
			515						520					525
Ser	Trp	Val	Met	Trp	Leu	Glu	Leu	Pro	Arg	Cys	Leu	Tyr	Asn	Cys
			530						535					540
Lys	Asp	Ile	Val	Leu	Arg	Arg	Asn	Thr	Ala	Gly	Ser	Leu	Gly	Phe
			545						550					555
Cys	Ile	Val	Gly	Gly	Tyr	Glu	Glu	Tyr	Asn	Gly	Asn	Lys	Pro	Phe
			560						565					570
Phe	Ile	Lys	Ser	Ile	Val	Glu	Gly	Thr	Pro	Ala	Tyr	Asn	Asp	Gly
			575						580					585
Arg	Ile	Arg	Cys	Gly	Asp	Ile	Leu	Leu	Ala	Val	Asn	Gly	Arg	Ser
			590						595					600
Thr	Ser	Gly	Met	Ile	His	Ala	Cys	Leu	Ala	Arg	Leu	Leu	Lys	Glu
			605						610					615
Leu	Lys	Gly	Arg	Ile	Thr	Leu	Thr	Ile	Val	Ser	Trp	Pro	Gly	Thr
			620						625					630
Phe	Leu													

<210> SEQ ID NO 41
<211> LENGTH: 1964
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 41

accaggcatt gtatcttcag ttgtcatcaa gttcgcaatc agattggaaa	50
agctcaactt gaagctttct tgctgcagt gaagcagaga gatagatatt	100
attcacgtaa taaaaaacat gggcttcaac ctgactttcc acctttocta	150

-continued

caaattccga ttactgttgc tgttgacttt gtgcctgaca gtggttgggt	200
gggccaccag taactacttc gtgggtgccca ttcaagagat tcctaaagca	250
aaggagtcca tggctaattt ccataagacc ctcatTTtTgTg ggaaggga	300
aactctgactaatgaagcat ccacgaagaa ggtagaactt gacaactgtc	350
cttctgtgtc tccttacctc agaggccaga gcaagctcat tttcaacca	400
gatctcactt tggaagaggt acaggcagaa aatcccaaag tgtccagagg	450
ccggtatcgc cctcaggaat gtaaagcttt acagagggtc gccatcctcg	500
ttccccaccg gaacagagag aaacacctga tgtacctgct ggaacatctg	550
catcccttcc tgcagaggca gcagctggat tatggcatct acgtcatcca	600
ccaggctgaa ggtaaaaagt ttaatcgagc caaactcttg aatgtgggct	650
atctagaagc cctcaaggaa gaaaattggg actgctttat attocacgat	700
gtggacctgg tacccgagaa tgactttaac ctttacaagt gtgaggagca	750
tccaagcat ctggtggttg gcaggaacag cactgggtac aggttacgtt	800
acagtggata ttttgggggt gttactgccc taagcagaga gcagtttttc	850
aagggtgaatg gattctctaa caactactgg ggatggggag gcgaagacga	900
tgacctcaga ctcagggttg agctccaaag aatgaaaatt tcccggtccc	950
tgccctgaagt gggtaaatat acaatgggtc tccacactag agacaaaggc	1000
aatgagggtga acgcagaacg gatgaagctc ttacaccaag tgtcacgagt	1050
ctggagaaca gatgggttga gtagtgttcc ttataaatta gtatctgtgg	1100
aacacaatcc tttatatatc aacatcacag tggatttctg gtttggtgca	1150
tgaccctgga tcttttgggt atgtttggaa gaactgattc tttgtttgca	1200
ataatttttg cctagagact tcaaatagta gcacacatta agaacctgtt	1250
acagctcatt gttgagctga atttttcctt tttgtatttt cttagcagag	1300
ctcctgggtga tgtagagtat aaaacagttg taacaagaca gctttcttag	1350
tcattttgat catgagggtt aaatattgta atatggatac ttgaaggact	1400
ttatataaaa ggatgactca aaggataaaa tgaacgctat ttgaggactc	1450
tggttgaagg agatttatTT aaatttgaag taatatatta tgggataaaa	1500
ggccacagga aataagactg ctgaatgtct gagagaacca gagttgttct	1550
cgtccaaggt agaaaggtac gaagatacaa tactgttatt catTTtTcct	1600
gtacaatcat ctgtgaagtg gtggtgtcag gtgagaaggc gtccacaaaa	1650
gaggggagaa aaggcgacga atcaggacac agtgaacttg ggaatgaaga	1700
ggtagcagga ggggtggagt tcggctgcaa aggcagcagt agctgagctg	1750
gttgacagtg ctgatagcct tcaggggagg acctgccag gtatgccttc	1800
cagtgatgcc caccagagaa tacattctct attagtTTTT aaagagTTTT	1850
tgtaaaatga ttttgtacaa gtaggatatg aattagcagt ttacaagttt	1900
acatattaac taataataaa tatgtctatc aaatacctct gtagtaaaat	1950
gtgaaaaagc aaaa	1964

-continued

<211> LENGTH: 344

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 42

```

Met Gly Phe Asn Leu Thr Phe His Leu Ser Tyr Lys Phe Arg Leu
 1           5           10          15
Leu Leu Leu Leu Thr Leu Cys Leu Thr Val Val Gly Trp Ala Thr
 20          25          30
Ser Asn Tyr Phe Val Gly Ala Ile Gln Glu Ile Pro Lys Ala Lys
 35          40          45
Glu Phe Met Ala Asn Phe His Lys Thr Leu Ile Leu Gly Lys Gly
 50          55          60
Lys Thr Leu Thr Asn Glu Ala Ser Thr Lys Lys Val Glu Leu Asp
 65          70          75
Asn Cys Pro Ser Val Ser Pro Tyr Leu Arg Gly Gln Ser Lys Leu
 80          85          90
Ile Phe Lys Pro Asp Leu Thr Leu Glu Glu Val Gln Ala Glu Asn
 95          100         105
Pro Lys Val Ser Arg Gly Arg Tyr Arg Pro Gln Glu Cys Lys Ala
 110         115         120
Leu Gln Arg Val Ala Ile Leu Val Pro His Arg Asn Arg Glu Lys
 125         130         135
His Leu Met Tyr Leu Leu Glu His Leu His Pro Phe Leu Gln Arg
 140         145         150
Gln Gln Leu Asp Tyr Gly Ile Tyr Val Ile His Gln Ala Glu Gly
 155         160         165
Lys Lys Phe Asn Arg Ala Lys Leu Leu Asn Val Gly Tyr Leu Glu
 170         175         180
Ala Leu Lys Glu Glu Asn Trp Asp Cys Phe Ile Phe His Asp Val
 185         190         195
Asp Leu Val Pro Glu Asn Asp Phe Asn Leu Tyr Lys Cys Glu Glu
 200         205         210
His Pro Lys His Leu Val Val Gly Arg Asn Ser Thr Gly Tyr Arg
 215         220         225
Leu Arg Tyr Ser Gly Tyr Phe Gly Gly Val Thr Ala Leu Ser Arg
 230         235         240
Glu Gln Phe Phe Lys Val Asn Gly Phe Ser Asn Asn Tyr Trp Gly
 245         250         255
Trp Gly Gly Glu Asp Asp Asp Leu Arg Leu Arg Val Glu Leu Gln
 260         265         270
Arg Met Lys Ile Ser Arg Pro Leu Pro Glu Val Gly Lys Tyr Thr
 275         280         285
Met Val Phe His Thr Arg Asp Lys Gly Asn Glu Val Asn Ala Glu
 290         295         300
Arg Met Lys Leu Leu His Gln Val Ser Arg Val Trp Arg Thr Asp
 305         310         315
Gly Leu Ser Ser Cys Ser Tyr Lys Leu Val Ser Val Glu His Asn
 320         325         330
Pro Leu Tyr Ile Asn Ile Thr Val Asp Phe Trp Phe Gly Ala
 335         340

```

-continued

<210> SEQ ID NO 43
<211> LENGTH: 485
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 43

gctcaagacc cagcagtggg acagccagac agacggcacg atggcactga 50
gctcccagat ctgggccgct tgccctcctgc tcctcctcct cctcgccagc 100
ctgaccagtg gctctgtttt cccacaacag acgggacaac ttgcagagct 150
gcaaccccag gacagagctg gagccagggc cagctggatg cccatgttcc 200
agaggcgaag gaggcgagac acccacttcc ccattctgcat tttctgctgc 250
ggctgctgtc atcgatcaaa gtgtgggatg tgctgcaaga cgtagaacct 300
acctgccctg ccccgctccc ctcccttctt tatttattcc tgctgcccga 350
gaacataggt cttgaataa aatggctggt tcttttgttt tccaaaaaaa 400
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 450
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 485

<210> SEQ ID NO 44
<211> LENGTH: 84
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 44

Met Ala Leu Ser Ser Gln Ile Trp Ala Ala Cys Leu Leu Leu Leu
1 5 10 15
Leu Leu Leu Ala Ser Leu Thr Ser Gly Ser Val Phe Pro Gln Gln
20 25 30
Thr Gly Gln Leu Ala Glu Leu Gln Pro Gln Asp Arg Ala Gly Ala
35 40 45
Arg Ala Ser Trp Met Pro Met Phe Gln Arg Arg Arg Arg Arg Asp
50 55 60
Thr His Phe Pro Ile Cys Ile Phe Cys Cys Gly Cys Cys His Arg
65 70 75
Ser Lys Cys Gly Met Cys Cys Lys Thr
80

<210> SEQ ID NO 45
<211> LENGTH: 1076
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 45

gtggcttcat ttcagtggct gacttccaga gagcaatatg gctggttccc 50
caacatgcct caccctcatc tatatccttt ggcagctcac agggtcagca 100
gcctctggag ccgtgaaaga gctggtcggt tccgttggtg gggccgtgac 150
tttccccctg aagtccaaag taaagcaagt tgactctatt gtctggacct 200
tcaacacaac ccctcttgtc accatacagc cagaaggggg cactatcata 250
gtgacccaaa atcgtaatag ggagagagta gacttccag atggaggcta 300
ctccctgaag ctgagcaaac tgaagaagaa tgactcaggg atctactatg 350
tggggatata cagctcatca ctccagcagc cctccacca ggagtagctg 400

-continued

```
ctgcatgtct acgagcacct gtcaaagcct aaagtcacca tgggtctgca      450
gagcaataag aatggcacct gtgtgaccaa tctgacatgc tgcattggaac      500
atggggaaga ggatgtgatt tatacctgga aggccctggg gcaagcagcc      550
aatgagtgcc ataatgggtc catcctcccc atctcctgga gatggggaga      600
aagtgatatg accttcattc gcgttgccag gaaccctgtc agcagaaaact      650
tctcaagccc catccttgcc aggaagctct gtgaaggtgc tgctgatgac      700
ccagattcct ccatggctct cctgtgtctc ctggttggtc ccctcctgct      750
cagtctcttt gtactggggc tatttctttg gtttctgaag agagagagac      800
aagaagagta cattgaagag aagaagagag tggacatttg tcgggaaact      850
cctaacatat gcccccattc tggagagaac acagagtacg acacaatccc      900
tcacactaat agaacaatcc taaaggaaga tccagcaaat acggtttact      950
ccactgtgga aataccgaaa aagatggaaa atccccactc actgctcacg     1000
atgccagaca caccaaggct atttgcttat gagaatgtta tctagacagc     1050
agtgcactcc cctaagtctc tgctca                                1076
```

<210> SEQ ID NO 46

<211> LENGTH: 335

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 46

```
Met Ala Gly Ser Pro Thr Cys Leu Thr Leu Ile Tyr Ile Leu Trp
 1             5             10            15
Gln Leu Thr Gly Ser Ala Ala Ser Gly Pro Val Lys Glu Leu Val
          20          25          30
Gly Ser Val Gly Gly Ala Val Thr Phe Pro Leu Lys Ser Lys Val
          35          40          45
Lys Gln Val Asp Ser Ile Val Trp Thr Phe Asn Thr Thr Pro Leu
          50          55          60
Val Thr Ile Gln Pro Glu Gly Gly Thr Ile Ile Val Thr Gln Asn
          65          70          75
Arg Asn Arg Glu Arg Val Asp Phe Pro Asp Gly Gly Tyr Ser Leu
          80          85          90
Lys Leu Ser Lys Leu Lys Lys Asn Asp Ser Gly Ile Tyr Tyr Val
          95          100         105
Gly Ile Tyr Ser Ser Ser Leu Gln Gln Pro Ser Thr Gln Glu Tyr
          110         115         120
Val Leu His Val Tyr Glu His Leu Ser Lys Pro Lys Val Thr Met
          125         130         135
Gly Leu Gln Ser Asn Lys Asn Gly Thr Cys Val Thr Asn Leu Thr
          140         145         150
Cys Cys Met Glu His Gly Glu Glu Asp Val Ile Tyr Thr Trp Lys
          155         160         165
Ala Leu Gly Gln Ala Ala Asn Glu Ser His Asn Gly Ser Ile Leu
          170         175         180
Pro Ile Ser Trp Arg Trp Gly Glu Ser Asp Met Thr Phe Ile Cys
          185         190         195
```

-continued

Val	Ala	Arg	Asn	Pro	Val	Ser	Arg	Asn	Phe	Ser	Ser	Pro	Ile	Leu
			200						205					210
Ala	Arg	Lys	Leu	Cys	Glu	Gly	Ala	Ala	Asp	Asp	Pro	Asp	Ser	Ser
			215						220					225
Met	Val	Leu	Leu	Cys	Leu	Leu	Leu	Val	Pro	Leu	Leu	Leu	Ser	Leu
			230						235					240
Phe	Val	Leu	Gly	Leu	Phe	Leu	Trp	Phe	Leu	Lys	Arg	Glu	Arg	Gln
			245						250					255
Glu	Glu	Tyr	Ile	Glu	Glu	Lys	Lys	Arg	Val	Asp	Ile	Cys	Arg	Glu
			260						265					270
Thr	Pro	Asn	Ile	Cys	Pro	His	Ser	Gly	Glu	Asn	Thr	Glu	Tyr	Asp
			275						280					285
Thr	Ile	Pro	His	Thr	Asn	Arg	Thr	Ile	Leu	Lys	Glu	Asp	Pro	Ala
			290						295					300
Asn	Thr	Val	Tyr	Ser	Thr	Val	Glu	Ile	Pro	Lys	Lys	Met	Glu	Asn
			305						310					315
Pro	His	Ser	Leu	Leu	Thr	Met	Pro	Asp	Thr	Pro	Arg	Leu	Phe	Ala
			320						325					330
Tyr	Glu	Asn	Val	Ile										
			335											

<210> SEQ ID NO 47
<211> LENGTH: 766
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 47

ggctcgagcg tttctgagcc aggggtgacc atgacctgct gcgaaggatg	50
gacatcctgc aatggattca gcctgctggt tctactgctg ttaggagtag	100
ttctcaatgc gatacctcta attgtcagct tagttgagga agaccaattt	150
tctcaaaacc ccatctcttg ctttgagtgg tggttccag gaattatag	200
agcaggctctg atggccattc cagcaacaac aatgtccttg acagcaagaa	250
aaagagcgtg ctgcaacaac agaactggaa tgtttctttc atcatttttc	300
agtgtgatca cagtcattgg tgctctgtat tgcagtctga tatccatcca	350
ggctctctta aaaggtcctc tcatgtgtaa ttctccaagc aacagtaatg	400
ccaattgtga attttcattg aaaaacatca gtgacattca tccagaatcc	450
ttcaacttgc agtgggtttt caatgactct tgtgcacctc ctactggttt	500
caataaacc accagtaacg acaccatggc gagtggctgg agagcatcta	550
gtttccactt cgattctgaa gaaaacaaac ataggcttat ccactttctca	600
gtatttttag gtctattgct tgttggaatt ctggaggctc tgtttgggct	650
cagtcagata gtcacgggtt tccttggtctg tctgtgtgga gtctctaagc	700
gaagaagtca aattgtgtag tttaatggga ataaaatgta agtatcagta	750
gtttgaaaaa aaaaaa	766

<210> SEQ ID NO 48
<211> LENGTH: 229
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 48

Met	Thr	Cys	Cys	Glu	Gly	Trp	Thr	Ser	Cys	Asn	Gly	Phe	Ser	Leu
1				5					10					15
Leu	Val	Leu	Leu	Leu	Leu	Gly	Val	Val	Leu	Asn	Ala	Ile	Pro	Leu
				20					25					30
Ile	Val	Ser	Leu	Val	Glu	Glu	Asp	Gln	Phe	Ser	Gln	Asn	Pro	Ile
				35					40					45
Ser	Cys	Phe	Glu	Trp	Trp	Phe	Pro	Gly	Ile	Ile	Gly	Ala	Gly	Leu
				50					55					60
Met	Ala	Ile	Pro	Ala	Thr	Thr	Met	Ser	Leu	Thr	Ala	Arg	Lys	Arg
				65					70					75
Ala	Cys	Cys	Asn	Asn	Arg	Thr	Gly	Met	Phe	Leu	Ser	Ser	Phe	Phe
				80					85					90
Ser	Val	Ile	Thr	Val	Ile	Gly	Ala	Leu	Tyr	Cys	Met	Leu	Ile	Ser
				95					100					105
Ile	Gln	Ala	Leu	Leu	Lys	Gly	Pro	Leu	Met	Cys	Asn	Ser	Pro	Ser
				110					115					120
Asn	Ser	Asn	Ala	Asn	Cys	Glu	Phe	Ser	Leu	Lys	Asn	Ile	Ser	Asp
				125					130					135
Ile	His	Pro	Glu	Ser	Phe	Asn	Leu	Gln	Trp	Phe	Phe	Asn	Asp	Ser
				140					145					150
Cys	Ala	Pro	Pro	Thr	Gly	Phe	Asn	Lys	Pro	Thr	Ser	Asn	Asp	Thr
				155					160					165
Met	Ala	Ser	Gly	Trp	Arg	Ala	Ser	Ser	Phe	His	Phe	Asp	Ser	Glu
				170					175					180
Glu	Asn	Lys	His	Arg	Leu	Ile	His	Phe	Ser	Val	Phe	Leu	Gly	Leu
				185					190					195
Leu	Leu	Val	Gly	Ile	Leu	Glu	Val	Leu	Phe	Gly	Leu	Ser	Gln	Ile
				200					205					210
Val	Ile	Gly	Phe	Leu	Gly	Cys	Leu	Cys	Gly	Val	Ser	Lys	Arg	Arg
				215					220					225

Ser Gln Ile Val

<210> SEQ ID NO 49
<211> LENGTH: 636
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 49

atccgttctc	tgcgctgcc	gtcaggtga	gccctcgcca	aggtgacctc	50
gcagacact	ggtgaaggag	cagtgaggaa	cctgcagagt	cacacagttg	100
ctgaccaatt	gagctgtgag	cctggagcag	atccgtgggc	tgcagacccc	150
cgccccagt	cctctcccc	tgcagccctg	cccctcgaac	tgtgacatgg	200
agagagtgc	cctggccctt	ctcctactgg	caggcctgac	tgccttgga	250
gccaatgacc	catttgccaa	taaagacgat	cccttctact	atgactggaa	300
aaacctgcag	ctgagcggac	tgatctgcgg	agggctcctg	gccattgctg	350
ggatcgcggc	agttctgagt	ggcaaatgca	aatacaagag	cagccagaag	400
cagcacagtc	ctgtacctga	gaaggccatc	ccactcatca	ctccaggtc	450
tgccactact	tgctgagcac	aggactggcc	tcagggatg	gcctgaagcc	500

-continued

taacactggc cccagcacc tcctcccctg ggaggcctta tcctcaagga	550
aggactttctc tccaagggca ggctgttagg cccctttctg atcaggaggc	600
ttcttttatga attaaactcg ccccaccacc ccctca	636

<210> SEQ ID NO 50
<211> LENGTH: 89
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 50

Met	Glu	Arg	Val	Thr	Leu	Ala	Leu	Leu	Leu	Leu	Ala	Gly	Leu	Thr
1				5					10				15	
Ala	Leu	Glu	Ala	Asn	Asp	Pro	Phe	Ala	Asn	Lys	Asp	Asp	Pro	Phe
			20						25				30	
Tyr	Tyr	Asp	Trp	Lys	Asn	Leu	Gln	Leu	Ser	Gly	Leu	Ile	Cys	Gly
			35						40				45	
Gly	Leu	Leu	Ala	Ile	Ala	Gly	Ile	Ala	Ala	Val	Leu	Ser	Gly	Lys
			50						55				60	
Cys	Lys	Tyr	Lys	Ser	Ser	Gln	Lys	Gln	His	Ser	Pro	Val	Pro	Glu
			65						70				75	
Lys	Ala	Ile	Pro	Leu	Ile	Thr	Pro	Gly	Ser	Ala	Thr	Thr	Cys	
			80						85					

<210> SEQ ID NO 51
<211> LENGTH: 1734
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 51

gtggactctg agaagcccag gcagttgagg acaggagaga gaaggctgca	50
gaccagaggg gagggaggac agggagtcgg aaggaggagg acagaggagg	100
gcacagagac gcagagcaag ggcggcaagg aggagaccct ggtgggagga	150
agacactctg gagagagagg gggctgggca gagatgaagt tccaggggcc	200
cctggcctgc ctctgctgg ccctctgcct gggcagtggg gaggctggcc	250
ccctgcagag cggagaggaa agcactggga caaatattgg ggaggccctt	300
ggacatggcc tgggagacgc cctgagcgaa ggggtgggaa aggccattgg	350
caaagaggcc ggaggggcag ctggctctaa agtcagtgag gcccttggcc	400
aagggaccag agaagcagtt ggcactggag tcaggcaggt tccaggcttt	450
ggcgacagag atgctttggg caacagggtc ggggaagcag cccatgctct	500
gggaaacact gggcacgaga ttggcagaca ggcagaagat gtcattcgac	550
acggagcaga tgctgtccgc ggctcctggc aggggggtgcc tggccacagt	600
ggtgcttggg aaacttctgg aggccatggc atctttggct ctcaaggtgg	650
ccttgagggc cagggccagg gcaatcctgg aggtctgggg actccgtggg	700
tccacggata ccccgaaac tcagcaggca gctttggaat gaatcctcag	750
ggagctccct ggggtcaagg aggcaatgga gggccaccaa actttgggac	800
caacactcag ggagctgtgg ccagcctgg ctatggttca gtgagagcca	850
gcaaccagaa tgaagggtgc acgaatcccc caccatctgg ctcaggtgga	900

-continued

ggctccagca actctggggg aggcagcggc tcacagtcgg gcagcagtgg	950
cagtggcagc aatggtgaca acaacaatgg cagcagcagt ggtggcagca	1000
cag cagcagtggc agcagcagtg gcggcagcag tggcggcagc	1050
gca gcagtggcaa cagtgggtgc agcagaggtg acagcggcag	1100
tcc tgggggatcca gcaccggctc ctctcccggc aaccacggtg	1150
cgg aggaaatgga cataaacccg ggtgtgaaaa gccagggaat	1200
gcg ggagcgggga atctgggatt cagggttca gaggacaggg	1250
agc aacatgaggg aaataagcaa agagggcaat cgcctccttg	1300
tgg agacaattat cgggggcaag ggtcgagctg gggcagtgga	1350
acg ctgttggtgg agtcaatact gtgaactctg agacgtctcc	1400
ttt aactttgaca ctttctggaa gaattttaaa tccaagctgg	1450
caa ctgggatgcc ataaacaagg accagagaag ctctcgcatc	1500
ctc cagacaagga gccaccagat tggatgggag ccccccact	1550
taa aacaccaccc tctcatcact aatctcagcc cttgcccttg	1600
cct tagctgcccc acaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1650
aaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1700
aaa aaaaaaaaaa aaaaaaaaaa aaaa	1734

EQ ID NO 52
ENGTH: 440
YPE: PRT
RGANISM: Homo Sapien

EQUENCE: 52

Met Lys Phe Gln Gly Pro Leu Ala Cys Leu Leu Leu Ala Leu Cys	1 5 10 15
Leu Gly Ser Gly Glu Ala Gly Pro Leu Gln Ser Gly Glu Glu Ser	20 25 30
Thr Gly Thr Asn Ile Gly Glu Ala Leu Gly His Gly Leu Gly Asp	35 40 45
Ala Leu Ser Glu Gly Val Gly Lys Ala Ile Gly Lys Glu Ala Gly	50 55 60
Gly Ala Ala Gly Ser Lys Val Ser Glu Ala Leu Gly Gln Gly Thr	65 70 75
Arg Glu Ala Val Gly Thr Gly Val Arg Gln Val Pro Gly Phe Gly	80 85 90
Ala Ala Asp Ala Leu Gly Asn Arg Val Gly Glu Ala Ala His Ala	95 100 105
Leu Gly Asn Thr Gly His Glu Ile Gly Arg Gln Ala Glu Asp Val	110 115 120
Ile Arg His Gly Ala Asp Ala Val Arg Gly Ser Trp Gln Gly Val	125 130 135
Pro Gly His Ser Gly Ala Trp Glu Thr Ser Gly Gly His Gly Ile	140 145 150
Phe Gly Ser Gln Gly Gly Leu Gly Gly Gln Gly Gln Gly Asn Pro	155 160 165

-continued

Gly Gly Leu Gly Thr Pro Trp Val His Gly Tyr Pro Gly Asn Ser
170 175 180

Ala Gly Ser Phe Gly Met Asn Pro Gln Gly Ala Pro Trp Gly Gln
185 190 195

Gly Gly Asn Gly Gly Pro Pro Asn Phe Gly Thr Asn Thr Gln Gly
200 205 210

Ala Val Ala Gln Pro Gly Tyr Gly Ser Val Arg Ala Ser Asn Gln
215 220 225

Asn Glu Gly Cys Thr Asn Pro Pro Pro Ser Gly Ser Gly Gly Gly
230 235 240

Ser Ser Asn Ser Gly Gly Gly Ser Gly Ser Gln Ser Gly Ser Ser
245 250 255

Gly Ser Gly Ser Asn Gly Asp Asn Asn Asn Gly Ser Ser Ser Gly
260 265 270

Gly Ser Ser Ser Gly Ser Ser Ser Gly Ser Ser Ser Gly Gly Ser
275 280 285

Ser Gly Gly Ser Ser Gly Gly Ser Ser Gly Asn Ser Gly Gly Ser
290 295 300

Arg Gly Asp Ser Gly Ser Glu Ser Ser Trp Gly Ser Ser Thr Gly
305 310 315

Ser Ser Ser Gly Asn His Gly Gly Ser Gly Gly Gly Asn Gly His
320 325 330

Lys Pro Gly Cys Glu Lys Pro Gly Asn Glu Ala Arg Gly Ser Gly
335 340 345

Glu Ser Gly Ile Gln Gly Phe Arg Gly Gln Gly Val Ser Ser Asn
350 355 360

Met Arg Glu Ile Ser Lys Glu Gly Asn Arg Leu Leu Gly Gly Ser
365 370 375

Gly Asp Asn Tyr Arg Gly Gln Gly Ser Ser Trp Gly Ser Gly Gly
380 385 390

Gly Asp Ala Val Gly Gly Val Asn Thr Val Asn Ser Glu Thr Ser
395 400 405

Pro Gly Met Phe Asn Phe Asp Thr Phe Trp Lys Asn Phe Lys Ser
410 415 420

Lys Leu Gly Phe Ile Asn Trp Asp Ala Ile Asn Lys Asp Gln Arg
425 430 435

Ser Ser Arg Ile Pro
440

<210> SEQ ID NO 53

<211> LENGTH: 1676

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 53

ggagaagagg ttgtgtggga caagctgctc ccgacagaag gatgtcgctg	50
ctgagcctgc cctggctggg cctcagaccg gtggcaatgt ccccatggct	100
actcctgctg ctggttgtgg gctcctggct actcgcccg atcctggctt	150
ggacctatgc cttctataac aactgccgcc ggctccagtg tttccacag	200
ccccaaaac ggaactggtt ttggggtcac ctgggctga tcactcctac	250
agaggagggc ttgaaggact cgaccagat gtcggccacc tattcccagg	300

-continued

```
gctttacggt atggctgggt cccatcatcc cttcatcgt tttatgccac      350
cctgacacca tccggtctat caccaatgcc tcagctgcca ttgcacccaa      400
ggataatctc ttcacaggt tcctgaagcc ctggctggga gaagggatac      450
tgctgagtgg cggtgacaag tggagccgcc accgtcggat gctgacgccc      500
gccttcatt tcaacatcct gaagtccat ataacgatct tcaacaagag      550
tgcaaacatc atgcttgaca agtggcagca cctggcctca gagggcagca      600
gtcgtctgga catgtttgag cacatcagcc tcatgacctt ggacagtcta      650
cagaaatgca tcttcagctt tgacagccat tgtcaggaga ggcccagtga      700
atatattgcc accatcttgg agctcagtc ccttgtagag aaaagaagcc      750
agcatatcct ccagcacatg gactttctgt attacctctc ccatgacggg      800
cggcgcttcc acagggcctg ccgctgtgtg catgacttca cagacgctgt      850
catccgggag cggcgctcga ccctccccac tcagggtatt gatgattttt      900
tcaaagacaa agccaagtcc aagactttgg atttcattga tgtgcttctg      950
ctgagcaagg atgaagatgg gaaggcattg tcagatgagg atataagagc     1000
agaggctgac accttcatgt ttggaggcca tgacaccacg gccagtgggc     1050
tctcctgggt cctgtacaac ctgcgagggc acccagaata ccaggagcgc     1100
tgccgacagg aggtgcaaga gcttctgaag gaccgcgatc ctaaagagat     1150
tgaatgggac gacctggccc agctgccctt cctgacctg tgcgtgaagg     1200
agagcctgag gttacatccc ccagctccct tcatctcccg atgctgcacc     1250
caggacattg ttctcccaga tggccgagtc atcccaaag gcattacctg     1300
cctcatcgat attatagggg tccatcacia cccaactgtg tggccggatc     1350
ctgagggtcta cgacccttc cgctttgacc cagagaacag caaggggagg     1400
tcacctctgg cttttattcc tttctccgca gggcccagga actgcatcgg     1450
gcaggcgttc gccatggcgg agatgaaagt ggtcctggcg ttgatgctgc     1500
tgcaacttccg gttcctgcca gaccacactg agccccgcag gaagctggaa     1550
ttgatcatgc gcgccgaggg cgggctttgg ctgcgggtgg agccctgaa     1600
tgtaggcttg cagtgacttt ctgacccatc cacctgtttt ttgcagatt     1650
gtcatgaata aaacggtgct gtcaaaa                                1676
```

<210> SEQ ID NO 54
<211> LENGTH: 524
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 54

```
Met Ser Leu Leu Ser Leu Pro Trp Leu Gly Leu Arg Pro Val Ala
 1             5             10             15
Met Ser Pro Trp Leu Leu Leu Leu Val Val Gly Ser Trp Leu
             20             25             30
Leu Ala Arg Ile Leu Ala Trp Thr Tyr Ala Phe Tyr Asn Asn Cys
             35             40             45
Arg Arg Leu Gln Cys Phe Pro Gln Pro Pro Lys Arg Asn Trp Phe
             50             55             60
```

-continued

Trp Gly His Leu Gly Leu Ile Thr Pro Thr Glu Glu Gly Leu Lys	65	70	75
Asp Ser Thr Gln Met Ser Ala Thr Tyr Ser Gln Gly Phe Thr Val	80	85	90
Trp Leu Gly Pro Ile Ile Pro Phe Ile Val Leu Cys His Pro Asp	95	100	105
Thr Ile Arg Ser Ile Thr Asn Ala Ser Ala Ala Ile Ala Pro Lys	110	115	120
Asp Asn Leu Phe Ile Arg Phe Leu Lys Pro Trp Leu Gly Glu Gly	125	130	135
Ile Leu Leu Ser Gly Gly Asp Lys Trp Ser Arg His Arg Arg Met	140	145	150
Leu Thr Pro Ala Phe His Phe Asn Ile Leu Lys Ser Tyr Ile Thr	155	160	165
Ile Phe Asn Lys Ser Ala Asn Ile Met Leu Asp Lys Trp Gln His	170	175	180
Leu Ala Ser Glu Gly Ser Ser Arg Leu Asp Met Phe Glu His Ile	185	190	195
Ser Leu Met Thr Leu Asp Ser Leu Gln Lys Cys Ile Phe Ser Phe	200	205	210
Asp Ser His Cys Gln Glu Arg Pro Ser Glu Tyr Ile Ala Thr Ile	215	220	225
Leu Glu Leu Ser Ala Leu Val Glu Lys Arg Ser Gln His Ile Leu	230	235	240
Gln His Met Asp Phe Leu Tyr Tyr Leu Ser His Asp Gly Arg Arg	245	250	255
Phe His Arg Ala Cys Arg Leu Val His Asp Phe Thr Asp Ala Val	260	265	270
Ile Arg Glu Arg Arg Arg Thr Leu Pro Thr Gln Gly Ile Asp Asp	275	280	285
Phe Phe Lys Asp Lys Ala Lys Ser Lys Thr Leu Asp Phe Ile Asp	290	295	300
Val Leu Leu Leu Ser Lys Asp Glu Asp Gly Lys Ala Leu Ser Asp	305	310	315
Glu Asp Ile Arg Ala Glu Ala Asp Thr Phe Met Phe Gly Gly His	320	325	330
Asp Thr Thr Ala Ser Gly Leu Ser Trp Val Leu Tyr Asn Leu Ala	335	340	345
Arg His Pro Glu Tyr Gln Glu Arg Cys Arg Gln Glu Val Gln Glu	350	355	360
Leu Leu Lys Asp Arg Asp Pro Lys Glu Ile Glu Trp Asp Asp Leu	365	370	375
Ala Gln Leu Pro Phe Leu Thr Met Cys Val Lys Glu Ser Leu Arg	380	385	390
Leu His Pro Pro Ala Pro Phe Ile Ser Arg Cys Cys Thr Gln Asp	395	400	405
Ile Val Leu Pro Asp Gly Arg Val Ile Pro Lys Gly Ile Thr Cys	410	415	420
Leu Ile Asp Ile Ile Gly Val His His Asn Pro Thr Val Trp Pro	425	430	435

-continued

<210> SEQ ID NO 57

<211> LENGTH: 3334

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 57

cggtctgagc tcgagccgaa tcggctcgag gggcagtgga gcaccagca	50
ggccccaac atgctctgtc tgtgcctgta cgtgccggtc atcggggaag	100
cccagaccga gttccagtac tttagctga aggggctccc tgccgagctg	150
aagtccattt tcaagctcag tgtcttcac cctcccagg aattctccac	200
ctaccgccag tggaagcaga aaattgtaca agctggagat aaggaccttg	250
atgggcagct agactttgaa gaatttgtcc attatctcca agatcatgag	300
aagaagctga ggctggtgtt taagattttg gacaaaaaga atgatggacg	350
cattgacgcg caggagatca tgcagtccct gcgggacttg ggagtcaaga	400
tatctgaaca gcaggcagaa aaaattctca agagcatgga taaaaacggc	450
acgatgacca tcgactggaa cgagtggaga gactaccacc tcctccaccc	500
cgtagaaaa atccccgaga tcatctctca ctggaagcat tccacgatct	550
ttgatgtggg tgagaatcta acggtcccgg atgagttcac agtgaggag	600
aggcagacgg ggatgtgggt gagacacctg gtggcaggag gtggggcagg	650
ggccgtatcc agaacctgca cggccccctt ggacaggctc aagggtgctca	700
tgacaggcca tgctccccc agcaacaaca tgggcatcgt tggtagcttc	750
actcagatga ttcgagaagg aggggccagg tcaactctggc ggggcaatgg	800
catcaacgct ctcaaaattg cccccgaatc agccatcaa ttcatggcct	850
atgagcagat caagcgccct gttggtagtg accaggagac tctgaggatt	900
cacgagaggc ttgtggcagg gtccttgga ggggccatcg ccagagcag	950
catctaccca atggaggtcc tgaagaccgg gatggcgctg cggaagacag	1000
gccagtactc aggaatgctg gactgcgcca ggaggatcct ggccagagag	1050
gggtggggcg ccttctacaa aggtatgtc cccaacatgc tgggcatcat	1100
cccctatgcc ggcatcgacc ttgcagtcta cgagacgctc aagaatgcct	1150
ggctgcagca ctatgcagtg aacagcgcg accccggcgt gtttgtgtc	1200
ctggcctgtg gcaccatgtc cagtacctgt ggccagctgg ccagctacc	1250
cctggcccta gtcaggaccc ggatgcaggc gcaagcctct attgaggcg	1300
ctccggaggt gaccatgagc agcctcttca aacatatcct gcggaccgag	1350
ggggccttcg ggctgtacag ggggtggcc cccaacttca tgaaggtcat	1400
cccagctgtg agcatcagct acgtgtgtta cgagaacctg aagatcacc	1450
tgggcgtgca gtcgcggtga cggggggagg gcccccggc agtggactcg	1500
ctgatcctgg gccgcagcct ggggtgtgca gccatctcat tctgtgaatg	1550
tgccaacact aagctgtctc gagccaagct gtgaaaacc tagacgcacc	1600
cgcaggagg gtggggagag ctggcaggcc cagggttgt cctgctgacc	1650
ccagcagacc ctctgtttg ttccagcgaa gaccacaggc attccttagg	1700

-continued

gtccagggtc agcaggctcc gggctcacat gtgtaaggac aggacatddd	1750
ctgcagtgcc tgccaatagt gagcttggag cctggaggcc ggcttagttc	1800
ttccatttca cccttgacgc cagctgtttg ccaaggcccc tgccctctgg	1850
tctgccgtgc atctccctgt gccctcttgc tgcctgcctg tctgctgag	1900
taagggtgga ggagggctac agcccacatc ccacccctc gtccaatccc	1950
ataatccatg atgaaagggt aggtcacgtg gcctcccagg cctgacttcc	2000
caacctacag cattgacgcc aacttggctg tgaaggaaga ggaaggatc	2050
tggccttgtg gtcactggca tctgagccct gctgatggct ggggctctcg	2100
ggcatgcttg ggagtgcagg gggctcgggc tgcctggcct ggctgcacag	2150
aaggcaagtg ctggggctca tgggtgcttg agctggcctg gaccctgtca	2200
ggatggggcc caccacagaa ccaactcac tgtccccact gtggcatgag	2250
ggcagtggag caccatgttt gagggcgaag ggcagagcgt ttgtgtgttc	2300
tggggaggga aggaaaagggt gttggaggcc ttaattatgg actgttggga	2350
aaagggtttt gtccagaagg acaagccgga caaatgagcg acttctgtgc	2400
ttccagagga agacgaggga gcaggagctt ggctgactgc tcagagtctg	2450
ttctgacgcc ctgggggttc ctgtccaacc ccagcagggg cgcagcggga	2500
ccagcccac attccacttg tgtcactgct tggaacctat ttatdddgt	2550
ttatdddtaa cagagttatg tcctaactat tttatagat ttgtttaatt	2600
aatagcttgt cattttcaag ttcatttttt attcataatt atgttcatgg	2650
ttgattgtac cttcccaagc ccgccagtg ggatgggagg aggaggagaa	2700
ggggggcctt gggccgctgc agtcacatct gtccagagaa attccttttg	2750
ggactggagg cagaaaagcg gccagaaggc agcagccctg gctcctttcc	2800
tttggcagggt tggggaagggt cttgccccca gccttaggat ttcagggttt	2850
gactgggggc gtggagagag agggaggaac ctcaataacc ttgaagggtg	2900
aatccagtta tttcctgcgc tgcgagggtt tctttatttc actctttct	2950
gaatgtcaag gcagtgagggt gcctctcact gtgaatttgt ggtgggcggg	3000
ggctggagga gaggtgggg ggctggctcc gtccctccca gccttctgct	3050
gcccttgctt aacaatgccg gccaaactggc gacctcacgg ttgcacttcc	3100
attccaccag aatgacctga tgaggaaatc ttcaatagga tgcaaagatc	3150
aatgcaaaaa ttgttatata tgaacatata actggagtcg tcaaaaagca	3200
aattaagaaa gaattggacg ttagaagttg tcatttaaag cagccttcta	3250
ataaagttgt ttcaaagctg aaaaaaaaa aaaaaaaaa aaaaaaaaa	3300
aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaa	3334

<210> SEQ ID NO 58

<211> LENGTH: 469

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 58

Met Leu Cys Leu Cys Leu Tyr Val Pro Val Ile Gly Glu Ala Gln

-continued

1	5	10	15
Thr Glu Phe Gln Tyr	Phe Glu Ser Lys	Gly Leu Pro Ala Glu	Leu
	20	25	30
Lys Ser Ile Phe Lys	Leu Ser Val Phe	Ile Pro Ser Gln Glu	Phe
	35	40	45
Ser Thr Tyr Arg Gln	Trp Lys Gln Lys	Ile Val Gln Ala Gly	Asp
	50	55	60
Lys Asp Leu Asp Gly	Gln Leu Asp Phe	Glu Glu Phe Val His	Tyr
	65	70	75
Leu Gln Asp His Glu	Lys Lys Leu Arg	Leu Val Phe Lys Ile	Leu
	80	85	90
Asp Lys Lys Asn Asp	Gly Arg Ile Asp	Ala Gln Glu Ile Met	Gln
	95	100	105
Ser Leu Arg Asp Leu	Gly Val Lys Ile	Ser Glu Gln Gln Ala	Glu
	110	115	120
Lys Ile Leu Lys Ser	Met Asp Lys Asn	Gly Thr Met Thr Ile	Asp
	125	130	135
Trp Asn Glu Trp Arg	Asp Tyr His Leu	Leu His Pro Val Glu	Asn
	140	145	150
Ile Pro Glu Ile Ile	Leu Tyr Trp Lys	His Ser Thr Ile Phe	Asp
	155	160	165
Val Gly Glu Asn Leu	Thr Val Pro Asp	Glu Phe Thr Val Glu	Glu
	170	175	180
Arg Gln Thr Gly Met	Trp Trp Arg His	Leu Val Ala Gly Gly	Gly
	185	190	195
Ala Gly Ala Val Ser	Arg Thr Cys Thr	Ala Pro Leu Asp Arg	Leu
	200	205	210
Lys Val Leu Met Gln	Val His Ala Ser	Arg Ser Asn Asn Met	Gly
	215	220	225
Ile Val Gly Gly Phe	Thr Gln Met Ile	Arg Glu Gly Gly Ala	Arg
	230	235	240
Ser Leu Trp Arg Gly	Asn Gly Ile Asn	Val Leu Lys Ile Ala	Pro
	245	250	255
Glu Ser Ala Ile Lys	Phe Met Ala Tyr	Glu Gln Ile Lys Arg	Leu
	260	265	270
Val Gly Ser Asp Gln	Glu Thr Leu Arg	Ile His Glu Arg Leu	Val
	275	280	285
Ala Gly Ser Leu Ala	Gly Ala Ile Ala	Gln Ser Ser Ile Tyr	Pro
	290	295	300
Met Glu Val Leu Lys	Thr Arg Met Ala	Leu Arg Lys Thr Gly	Gln
	305	310	315
Tyr Ser Gly Met Leu	Asp Cys Ala Arg	Arg Ile Leu Ala Arg	Glu
	320	325	330
Gly Val Ala Ala Phe	Tyr Lys Gly Tyr	Val Pro Asn Met Leu	Gly
	335	340	345
Ile Ile Pro Tyr Ala	Gly Ile Asp Leu	Ala Val Tyr Glu Thr	Leu
	350	355	360
Lys Asn Ala Trp Leu	Gln His Tyr Ala	Val Asn Ser Ala Asp	Pro
	365	370	375
Gly Val Phe Val Leu	Leu Ala Cys Gly	Thr Met Ser Ser Thr	Cys
	380	385	390

-continued

Gly	Gln	Leu	Ala	Ser	Tyr	Pro	Leu	Ala	Leu	Val	Arg	Thr	Arg	Met
				395					400					405
Gln	Ala	Gln	Ala	Ser	Ile	Glu	Gly	Ala	Pro	Glu	Val	Thr	Met	Ser
				410					415					420
Ser	Leu	Phe	Lys	His	Ile	Leu	Arg	Thr	Glu	Gly	Ala	Phe	Gly	Leu
				425					430					435
Tyr	Arg	Gly	Leu	Ala	Pro	Asn	Phe	Met	Lys	Val	Ile	Pro	Ala	Val
				440					445					450
Ser	Ile	Ser	Tyr	Val	Val	Tyr	Glu	Asn	Leu	Lys	Ile	Thr	Leu	Gly
				455					460					465
Val	Gln	Ser	Arg											

<210> SEQ ID NO 59
<211> LENGTH: 1658
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 59

ggaaggcagc	ggcagctcca	ctcagccagt	accagatac	gctgggaacc	50
ttccccagcc	atggcttccc	tggggcagat	cctcttctg	agcataatta	100
gcatcatcat	tattctggct	ggagcaattg	cactcatcat	tggttttgt	150
atttcaggga	gacactccat	cacagtcact	actgtgcct	cagctgggaa	200
cattggggag	gatggaatcc	tgagctgcac	ttttgaacct	gacatcaaac	250
tttctgatat	cgtgatacaa	tggtgaagg	aaggtgtttt	aggottggtc	300
catgagttca	aagaaggcaa	agatgagctg	tcggagcagg	atgaaatgtt	350
cagaggccgg	acagcagtgt	ttgctgatca	agtgatagtt	ggcaatgcct	400
ctttgcgggt	gaaaaacgtg	caactcacag	atgtgggcac	ctacaaatgt	450
tatatcatca	cttctaaagg	caaggggaat	gctaaccctg	agtataaaac	500
tgagaccttc	agcatgccgg	aagtgaatgt	ggactataat	gccagctcag	550
agaccttgcg	gtgtgaggct	ccccgatggt	tccccagcc	cacagtggtc	600
tgggcatccc	aagttgacca	gggagccaac	ttctcggaag	tctccaatac	650
cagctttgag	ctgaactctg	agaatgtgac	catgaagggt	gtgtctgtgc	700
tctacaatgt	tacgatcaac	aacacatact	cctgtatgat	tgaaaatgac	750
attgccaaag	caacagggga	tatcaaagt	acagaatcgg	agatcaaaag	800
gcggagtcac	ctacagctgc	taaactcaaa	ggcttctctg	tgtgtctctt	850
ctttctttgc	catcagctgg	gcacttctgc	ctctcagccc	ttacctgatg	900
ctaaaaatat	gtgccttggc	cacaaaaaag	catgcaaagt	cattgttaca	950
acagggatct	acagaactat	ttcaccacca	gatatgacct	agttttatat	1000
ttctggggag	aaatgaattc	atatctagaa	gtctggagtg	agcaaacaag	1050
agcaagaaac	aaaaagaagc	caaaagcaga	aggctccaat	atgaacaaga	1100
taaatctatc	ttcaaagaca	tattagaagt	tgggaaaata	attcatgtga	1150
actagacaag	tgtgttaaga	gtgataagta	aaatgcacgt	ggagacaagt	1200
gcatccccag	atctcaggga	cctccccctg	cctgtcacct	ggggagtgag	1250

-continued

aggacaggat agtgcattgtt ctttgtctct gaatttttag ttatatgtgc	1300
tgtaatgttg ctctgaggaa gcccttgaa agtctatccc aacatatcca	1350
catcttatat tccacaaatt aagctgtagt atgtacccta agacgctgct	1400
aattgactgc cacttcgcaa ctcaggggcg gctgcatttt agtaatgggt	1450
caaatgattc actttttatg atgcttccaa aggtgccttg gcttctcttc	1500
ccaactgaca aatgccaaag ttgagaaaaa tgatcataat tttagcataa	1550
acagagcagt cggggacacc gattttataa ataaactgag caccttcttt	1600
ttaaacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1650
aaaaaaaa	1658

<210> SEQ ID NO 60

<211> LENGTH: 282

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 60

Met Ala Ser Leu Gly Gln Ile Leu Phe Trp Ser Ile Ile Ser Ile	1	5	10	15
Ile Ile Ile Leu Ala Gly Ala Ile Ala Leu Ile Ile Gly Phe Gly	20	25	30	
Ile Ser Gly Arg His Ser Ile Thr Val Thr Thr Val Ala Ser Ala	35	40	45	
Gly Asn Ile Gly Glu Asp Gly Ile Leu Ser Cys Thr Phe Glu Pro	50	55	60	
Asp Ile Lys Leu Ser Asp Ile Val Ile Gln Trp Leu Lys Glu Gly	65	70	75	
Val Leu Gly Leu Val His Glu Phe Lys Glu Gly Lys Asp Glu Leu	80	85	90	
Ser Glu Gln Asp Glu Met Phe Arg Gly Arg Thr Ala Val Phe Ala	95	100	105	
Asp Gln Val Ile Val Gly Asn Ala Ser Leu Arg Leu Lys Asn Val	110	115	120	
Gln Leu Thr Asp Ala Gly Thr Tyr Lys Cys Tyr Ile Ile Thr Ser	125	130	135	
Lys Gly Lys Gly Asn Ala Asn Leu Glu Tyr Lys Thr Gly Ala Phe	140	145	150	
Ser Met Pro Glu Val Asn Val Asp Tyr Asn Ala Ser Ser Glu Thr	155	160	165	
Leu Arg Cys Glu Ala Pro Arg Trp Phe Pro Gln Pro Thr Val Val	170	175	180	
Trp Ala Ser Gln Val Asp Gln Gly Ala Asn Phe Ser Glu Val Ser	185	190	195	
Asn Thr Ser Phe Glu Leu Asn Ser Glu Asn Val Thr Met Lys Val	200	205	210	
Val Ser Val Leu Tyr Asn Val Thr Ile Asn Asn Thr Tyr Ser Cys	215	220	225	
Met Ile Glu Asn Asp Ile Ala Lys Ala Thr Gly Asp Ile Lys Val	230	235	240	
Thr Glu Ser Glu Ile Lys Arg Arg Ser His Leu Gln Leu Leu Asn	245	250	255	

-continued

Ser Lys Ala Ser Leu Cys Val Ser Ser Phe Phe Ala Ile Ser Trp
260 265 270

Ala Leu Leu Pro Leu Ser Pro Tyr Leu Met Leu Lys
275 280

<210> SEQ ID NO 61
<211> LENGTH: 1617
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 61

tgacgtcaga atcaccatgg ccagctatcc ttaccggcag ggctgcccag	50
gagctgcagg acaagcacca ggagccccto cgggtagcta ctaccctgga	100
cccccaata gtgagggca gtatggtagt gggctacccc ctggtggtgg	150
ttatgggggt cctgcccctg gagggcctta tggaccacca gctggtggag	200
ggccctatgg acacccaat cctgggatgt tcccctctgg aactccagga	250
ggaccatatg gcggtgcagc tcccgggggc ccctatggtc agccacotcc	300
aagttcctac ggtgcccagc agcctgggct ttatggacag ggtggcgccc	350
ctcccaatgt ggatcctgag gcctactcct gggtccagtc ggtggactca	400
gatcacagtg gctatatctc catgaaggag ctaaagcagg ccctggtcaa	450
ctgcaattgg tcttcattca atgatgagac ctgcctcatg atgataaaca	500
tgtttgacaa gaccaagtca ggccgcacgc atgtctacgg cttctcagcc	550
ctgtggaaa tcatccagca gtggaagaac ctcttcagc agtatgaccg	600
ggaccgctcg ggctccatta gctacacaga gotgcagcaa gctctgtccc	650
aaatgggcta caacctgagc cccagttca ccagcttct ggtctccgc	700
tactgcccac gctctgccaa tcctgccatg cagcttgacc gcttcatcca	750
ggtgtgcacc cagctgcagg tgctgacaga ggccttcgg gagaaggaca	800
cagctgtaca aggcaacatc cggctcagct tcgaggactt cgtcaccatg	850
acagcttctc ggatgctatg acccaaccat ctgtggagag tggagtgcac	900
cagggacctt tcctggcttc ttagagttag agaagtatgt ggacatctct	950
tcttttctg tccctctaga agaacattct cccttgcttg atgcaacact	1000
gttccaaaag aggggtggaga gtcctgcac atagccacca aatagtgagg	1050
accggggctg agggcacaca gataggggccc tgatggagga gaggatagaa	1100
gttgaatgtc ctgatggcca tgagcagttg agtggcacag cctggcacca	1150
ggagcaggtc cttgtaatgg agttagtgtc cagtcagctg agctccaccc	1200
tgatgccagt ggtgagtgtt catcggcctg ttaccgtag tacctgtgtt	1250
ccctcaccag gccatcctgt caaacgagcc cattttctcc aaagtggaat	1300
ctgaccaagc atgagagaga tctgtctatg ggaccagtgg cttggattct	1350
gccacaccca taaatccttg tgtgttaact tctagctgcc tggggctggc	1400
cctgctcaga caaatctgct ccctgggcat ctttgccag gcttctgccc	1450
cctgcagctg ggacccctca cttgcctgcc atgctctgct cggttcagct	1500
ctccaggaga cagtggtoac ctctccctgc caatactttt ttttaatttg	1550

-continued

atttttttttc atttggggcc aaaagtccag tgaaattgta agcttcaata 1600

aaaggatgaa actctga 1617

<210> SEQ ID NO 62

<211> LENGTH: 284

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 62

Met Ala Ser Tyr Pro Tyr Arg Gln Gly Cys Pro Gly Ala Ala Gly
1 5 10 15Gln Ala Pro Gly Ala Pro Pro Gly Ser Tyr Tyr Pro Gly Pro Pro
20 25 30Asn Ser Gly Gly Gln Tyr Gly Ser Gly Leu Pro Pro Gly Gly Gly
35 40 45Tyr Gly Gly Pro Ala Pro Gly Gly Pro Tyr Gly Pro Pro Ala Gly
50 55 60Gly Gly Pro Tyr Gly His Pro Asn Pro Gly Met Phe Pro Ser Gly
65 70 75Thr Pro Gly Gly Pro Tyr Gly Gly Ala Ala Pro Gly Gly Pro Tyr
80 85 90Gly Gln Pro Pro Pro Ser Ser Tyr Gly Ala Gln Gln Pro Gly Leu
95 100 105Tyr Gly Gln Gly Gly Ala Pro Pro Asn Val Asp Pro Glu Ala Tyr
110 115 120Ser Trp Phe Gln Ser Val Asp Ser Asp His Ser Gly Tyr Ile Ser
125 130 135Met Lys Glu Leu Lys Gln Ala Leu Val Asn Cys Asn Trp Ser Ser
140 145 150Phe Asn Asp Glu Thr Cys Leu Met Met Ile Asn Met Phe Asp Lys
155 160 165Thr Lys Ser Gly Arg Ile Asp Val Tyr Gly Phe Ser Ala Leu Trp
170 175 180Lys Phe Ile Gln Gln Trp Lys Asn Leu Phe Gln Gln Tyr Asp Arg
185 190 195Asp Arg Ser Gly Ser Ile Ser Tyr Thr Glu Leu Gln Gln Ala Leu
200 205 210Ser Gln Met Gly Tyr Asn Leu Ser Pro Gln Phe Thr Gln Leu Leu
215 220 225Val Ser Arg Tyr Cys Pro Arg Ser Ala Asn Pro Ala Met Gln Leu
230 235 240Asp Arg Phe Ile Gln Val Cys Thr Gln Leu Gln Val Leu Thr Glu
245 250 255Ala Phe Arg Glu Lys Asp Thr Ala Val Gln Gly Asn Ile Arg Leu
260 265 270Ser Phe Glu Asp Phe Val Thr Met Thr Ala Ser Arg Met Leu
275 280

<210> SEQ ID NO 63

<211> LENGTH: 1234

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 63

caggatgcag	ggccgcgtgg	cagggagctg	cgctcctctg	ggcctgctcc	50
tggtctgtct	tcattctccca	ggcctctttg	cccggagcat	cggtgttggtg	100
gaggagaaa	tttcccaaaa	cttcgggacc	aacttgccctc	agctcggaca	150
accttcctcc	actggcccct	ctaactctga	acatccgcag	cccgtctgtg	200
accctaggtc	taatgacttg	gcaagggttc	ctctgaagct	cagcgtgcct	250
ccatcagatg	gcttcccacc	tgcaggaggt	tctgcagtgc	agagggtggcc	300
tccatcgtgg	gggtgcctg	ccatggattc	ctggccccct	gaggatcctt	350
ggcagatgat	ggctgctgcg	gctgaggacc	gcctggggga	agcgtgcct	400
gaagaactct	cttacctctc	cagtgcctgcg	gccctcgctc	cgggcagtgg	450
ccctttgcct	ggggagtctt	ctcccgatgc	cacaggccctc	tcacctgagg	500
cttcactcct	ccaccaggac	tcggagtcca	gacgactgcc	ccgttcta	550
tcactgggag	ccgggggaaa	aatcctttcc	caacgccctc	cctggtctct	600
catccacagg	gttctgcctg	atcacccctg	gggtaccctg	aatccagtg	650
tgctctgggg	aggtggaggc	cctgggactg	gttggggaac	gaggcccatg	700
ccacaccctg	agggaatctg	gggtatcaat	aatcaacccc	caggtaccag	750
ctgggggaaat	attaatcggg	atccaggagg	cagctgggga	aatattaatc	800
ggatcccagg	aggcagctgg	gggaatatta	atcggtatcc	aggaggcagc	850
tgggggaata	ttcatctata	cccaggatgc	aataacccat	ttcctcctgg	900
agttctccgc	cctcctggct	cttcttgga	catccagct	ggcttcctta	950
atcctccaag	ccctaggttg	cagtggggct	agagcacgat	agagggaaac	1000
ccaacattgg	gagttagagt	cctgctcccg	ccccttgctg	tgtgggctca	1050
atccaggccc	tgtaacatg	tttccagcac	tatccccact	tttcagtgcc	1100
tcccctgctc	atctccaata	aaataaaagc	acttatgaaa	aaaaaaaaaa	1150
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1200
aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaa		1234

<210> SEQ ID NO 64

<211> LENGTH: 325

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 64

Met	Gln	Gly	Arg	Val	Ala	Gly	Ser	Cys	Ala	Pro	Leu	Gly	Leu	Leu
1				5					10					15
Leu	Val	Cys	Leu	His	Leu	Pro	Gly	Leu	Phe	Ala	Arg	Ser	Ile	Gly
			20						25					30
Val	Val	Glu	Glu	Lys	Val	Ser	Gln	Asn	Phe	Gly	Thr	Asn	Leu	Pro
			35						40					45
Gln	Leu	Gly	Gln	Pro	Ser	Ser	Thr	Gly	Pro	Ser	Asn	Ser	Glu	His
			50						55					60
Pro	Gln	Pro	Ala	Leu	Asp	Pro	Arg	Ser	Asn	Asp	Leu	Ala	Arg	Val
			65						70					75
Pro	Leu	Lys	Leu	Ser	Val	Pro	Pro	Ser	Asp	Gly	Phe	Pro	Pro	Ala

-continued

80					85					90				
Gly	Gly	Ser	Ala	Val	Gln	Arg	Trp	Pro	Pro	Ser	Trp	Gly	Leu	Pro
95					100					105				
Ala	Met	Asp	Ser	Trp	Pro	Pro	Glu	Asp	Pro	Trp	Gln	Met	Met	Ala
110					115					120				
Ala	Ala	Ala	Glu	Asp	Arg	Leu	Gly	Glu	Ala	Leu	Pro	Glu	Glu	Leu
125					130					135				
Ser	Tyr	Leu	Ser	Ser	Ala	Ala	Ala	Leu	Ala	Pro	Gly	Ser	Gly	Pro
140					145					150				
Leu	Pro	Gly	Glu	Ser	Ser	Pro	Asp	Ala	Thr	Gly	Leu	Ser	Pro	Glu
155					160					165				
Ala	Ser	Leu	Leu	His	Gln	Asp	Ser	Glu	Ser	Arg	Arg	Leu	Pro	Arg
170					175					180				
Ser	Asn	Ser	Leu	Gly	Ala	Gly	Gly	Lys	Ile	Leu	Ser	Gln	Arg	Pro
185					190					195				
Pro	Trp	Ser	Leu	Ile	His	Arg	Val	Leu	Pro	Asp	His	Pro	Trp	Gly
200					205					210				
Thr	Leu	Asn	Pro	Ser	Val	Ser	Trp	Gly	Gly	Gly	Gly	Pro	Gly	Thr
215					220					225				
Gly	Trp	Gly	Thr	Arg	Pro	Met	Pro	His	Pro	Glu	Gly	Ile	Trp	Gly
230					235					240				
Ile	Asn	Asn	Gln	Pro	Pro	Gly	Thr	Ser	Trp	Gly	Asn	Ile	Asn	Arg
245					250					255				
Tyr	Pro	Gly	Gly	Ser	Trp	Gly	Asn	Ile	Asn	Arg	Tyr	Pro	Gly	Gly
260					265					270				
Ser	Trp	Gly	Asn	Ile	Asn	Arg	Tyr	Pro	Gly	Gly	Ser	Trp	Gly	Asn
275					280					285				
Ile	His	Leu	Tyr	Pro	Gly	Ile	Asn	Asn	Pro	Phe	Pro	Pro	Gly	Val
290					295					300				
Leu	Arg	Pro	Pro	Gly	Ser	Ser	Trp	Asn	Ile	Pro	Ala	Gly	Phe	Pro
305					310					315				
Asn	Pro	Pro	Ser	Pro	Arg	Leu	Gln	Trp	Gly					
320					325									

<210> SEQ ID NO 65
<211> LENGTH: 422
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 65

aaggagagggc caccgggact tcagtgtctc ctccatccca ggagcgcagt	50
ggccactatg gggctctgggc tgccccttgt cctcctcttg accctccttg	100
gcagctcaca tggaacaggg ccgggtatga ctttgcaact gaagctgaag	150
gagtcttttc tgacaaattc ctctatgag tccagcttcc tggaattgct	200
tgaaaagctc tgcctcctcc tccatctccc ttcagggacc agcgtcaccc	250
tccaccatgc aagatctcaa caccatgttg tctgcaacac atgacagcca	300
ttgaagcctg tgtccttctt ggcccgggct tttgggccgg ggaatgcagga	350
ggcaggcccc gaccctgtct ttcagcaggc cccaccctc ctgagtggca	400
ataaataaaa ttcggtatgc tg	422

-continued

<210> SEQ ID NO 66
<211> LENGTH: 78
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 66

Met	Gly	Ser	Gly	Leu	Pro	Leu	Val	Leu	Leu	Leu	Thr	Leu	Leu	Gly
1				5					10					15
Ser	Ser	His	Gly	Thr	Gly	Pro	Gly	Met	Thr	Leu	Gln	Leu	Lys	Leu
			20						25					30
Lys	Glu	Ser	Phe	Leu	Thr	Asn	Ser	Ser	Tyr	Glu	Ser	Ser	Phe	Leu
			35						40					45
Glu	Leu	Leu	Glu	Lys	Leu	Cys	Leu	Leu	Leu	His	Leu	Pro	Ser	Gly
			50						55					60
Thr	Ser	Val	Thr	Leu	His	His	Ala	Arg	Ser	Gln	His	His	Val	Val
			65						70					75
Cys	Asn	Thr												

<210> SEQ ID NO 67
<211> LENGTH: 744
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 67

acggaccgag	ggttcgaggg	agggacacgg	accaggaacc	tgagctaggt	50
caaagacgcc	cgggccaggt	gccccgtcgc	aggtgcccct	ggccggagat	100
gcggtagtag	gggcgagcgc	gagaagcccc	ttcctcggcg	ctgccaaacc	150
gccaccacgc	ccatggcgaa	ccccgggctg	gggctgcttc	tgcgctggg	200
cctgccgttc	ctgctggccc	gctggggccg	agcctggggg	caaatacaga	250
ccacttctgc	aaatgagaat	agcactgttt	tgccttcatc	caccagctcc	300
agctccgatg	gcaacctgcg	tccggaagcc	atcactgcta	tcatcgtggt	350
cttctccctc	ttggctgcct	tgctcctggc	tgtggggctg	gcactgttgg	400
tgcggaagct	tcgggagaag	cggcagacgg	agggcaccta	ccggcccagt	450
agcgaggagc	agttctccca	tgcagccgag	gcccggggccc	ctcaggactc	500
caaggagacg	gtgcagggct	gcctgcccac	ctaggtcccc	tctcctgcac	550
ctgtctccct	tcattgctgt	gtgaccttgg	ggaaaggcag	tgccctctct	600
gggcagtcag	atccacccag	tgcttaatat	cagggaagaa	ggtacttcaa	650
agactctgcc	cctgaggtca	agagaggatg	gggctattca	cttttatata	700
tttatataaa	attagtagtg	agatgtaaaa	aaaaaaaaaa	aaaa	744

<210> SEQ ID NO 68
<211> LENGTH: 123
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 68

Met	Ala	Asn	Pro	Gly	Leu	Gly	Leu	Leu	Leu	Ala	Leu	Gly	Leu	Pro
1				5					10					15
Phe	Leu	Leu	Ala	Arg	Trp	Gly	Arg	Ala	Trp	Gly	Gln	Ile	Gln	Thr
			20						25					30

-continued

Thr	Ser	Ala	Asn	Glu	Asn	Ser	Thr	Val	Leu	Pro	Ser	Ser	Thr	Ser	
			35						40					45	
Ser	Ser	Ser	Asp	Gly	Asn	Leu	Arg	Pro	Glu	Ala	Ile	Thr	Ala	Ile	
			50						55					60	
Ile	Val	Val	Phe	Ser	Leu	Leu	Ala	Ala	Leu	Leu	Leu	Ala	Val	Gly	
			65						70					75	
Leu	Ala	Leu	Leu	Val	Arg	Lys	Leu	Arg	Glu	Lys	Arg	Gln	Thr	Glu	
			80						85					90	
Gly	Thr	Tyr	Arg	Pro	Ser	Ser	Glu	Glu	Gln	Phe	Ser	His	Ala	Ala	
			95						100					105	
Glu	Ala	Arg	Ala	Pro	Gln	Asp	Ser	Lys	Glu	Thr	Val	Gln	Gly	Cys	
			110						115					120	
Leu	Pro	Ile													

<210> SEQ ID NO 69
<211> LENGTH: 3265
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 69

gccaggaata actagagagg aacaatgggg ttattcagag gttttgtttt	50
cctcttagtt ctgtgcctgc tgcaccagtc aaatacttcc ttcattaagc	100
tgaataataa tggctttgaa gatattgtca ttgttataga tcctagtgtg	150
ccagaagatg aaaaaataat tgaacaaata gaggatatgg tgactacagc	200
ttctacgtac ctgtttgaag ccacagaaaa aagatTTTTT ttcaaaaatg	250
tatctatatt aattcctgag aattggaagg aaaatcctca gtacaaaagg	300
ccaaaacatg aaaccataa acatgctgat gttatagttg caccacctac	350
actcccaggt agagatgaac catacaccaa gcagttcaca gaatgtggag	400
agaaaggcga atacattcac ttcacccctg accttctact tggaaaaaaa	450
caaaatgaat atggaccacc aggcaaactg tttgtccatg agtgggctca	500
cctccggtgg ggagtgtttg atgagtacaa tgaagatcag cctttctacc	550
gtgctaagtc aaaaaaatc gaagcaacaa ggtgttccgc aggtatctct	600
ggtagaaata gagtttataa gtgtcaagga ggcagctgtc ttagtagagc	650
atgcagaatt gattctacaa caaaactgta tggaaaagat tgtcaattct	700
ttcctgataa agtacaaca gaaaagcat ccataatggt tatgcaaagt	750
attgattctg ttgttgaatt ttgtaacgaa aaaaccata atcaagaagc	800
tccaagccta caaaacataa agtgcaattt tagaagtaca tgggaggtga	850
ttagcaattc tgaggatttt aaaaacacca tacccatggt gacaccacct	900
cctccacctg tcttctcatt gctgaagatc agtcaaagaa ttgtgtgctt	950
agttcttgat aagtctggaa gcatgggggg taaggaccgc ctaaatecgaa	1000
tgaatcaagc agcaaaacat ttcctgctgc agactgttga aaatggatcc	1050
tgggtgggga tggttcactt tgatagtact gccactattg taaataagct	1100
aatccaaata aaaagcagtg atgaaagaaa cacactcatg gcaggattac	1150
ctacatatoc tctgggagga acttccatct gctctggaat taaatatgca	1200

-continued

tttcaggatga ttggagagct acattcccaa ctcgatgat ccgaagtact	1250
gctgctgact gatggggagg ataactctgc aagttcttgt attgatgaag	1300
tgaacaaaag tggggccatt gttcatttta ttgctttggg aagagctgct	1350
gatgaagcag taatagagat gagcaagata acaggaggaa gtcattttta	1400
tgtttcagat gaagctcaga acaatggcct cattgatgct tttggggctc	1450
ttacatcagg aaatactgat ctctcccaga agtcccttca gctcgaaagt	1500
aagggattaa cactgaatag taatgcctgg atgaacgaca ctgtcataat	1550
tgatagtaca gtgggaaagg acacgttctt tctcatcaca tggaacagtc	1600
tgccctccag tatttctctc tgggatccca gtggaacaat aatggaaaat	1650
ttcacagtgg atgcaacttc caaatggcc tatctcagta ttccaggaaac	1700
tgcaagggtg ggcacttggg catacaatct tcaagccaaa gcgaaccag	1750
aaacattaac tattacagta acttctcgag cagcaaatc ttctgtgcct	1800
ccaatcacag tgaatgctaa aatgaataag gacgtaaaca gtttcccag	1850
cccaatgatt gtttacgcag aaattctaca aggatatgta cctgttcttg	1900
gagccaatgt gactgctttc attgaatcac agaatggaca tacagaagtt	1950
ttggaacttt tggataatgg tgcaggcgct gattctttca agaattgatg	2000
agtctactcc aggtatttta cagcatatac agaaaatggc agatatagct	2050
taaaagttcg ggctcatgga ggagcaaaac ctgccaggct aaaattacg	2100
cctccactga atagagccgc gtacatacca ggctgggtag tgaacgggga	2150
aattgaagca aaccgccaa gacctgaaat tgatgaggat actcagacca	2200
ccttgaggga tttcagccga acagcatccg gaggtgcatt tgtggatca	2250
caagtcccaa gccttccctt gcctgaccaa taccaccaa gtcaaatcac	2300
agaccttgat gccacagtgc atgaggataa gattattctt acatggacag	2350
caccaggaga taattttgat gttggaagg ttcaacgtta tatcataaga	2400
ataagtgcaa gtattcttga tctaagagac agttttgatg atgctcttca	2450
agtaaaact actgatctgt caccaaagga ggccaactcc aaggaaagct	2500
ttgcatttaa accagaaaat atctcagaag aaaatgcaac ccacatattt	2550
attgccatta aaagtataga taaaagcaat ttgacatcaa aagtatccaa	2600
cattgcacaa gtaactttgt ttatccctca agcaaatcct gatgacattg	2650
atcctacacc tactcctact cctactccta ctctgataa aagtcataat	2700
tctggagtta atatttctac gctggatttg tctgtgattg ggtctgttgt	2750
aattgttaac tttattttaa gtaccaccat ttgaacctta acgaagaaaa	2800
aaatcttcaa gtagacctag aagagagttt taaaaaaca aacaatgtaa	2850
gtaaaggata tttctgaatc ttaaaattca tcccatgtgt gatcataaac	2900
tcataaaaaa aattttaaga tgtcggaaaa ggatactttg attaaataaa	2950
aactctatg gatagttaa aactgtcaag attaaattt aatagtttca	3000
tttattttgt attttatttg taagaaatag tgatgaacaa agatcctttt	3050
tcatactgat acctggttgt atattatttg atgcaacagt tttctgaaat	3100

-continued

gatatttcaa attgcatcaa gaaattaaaa tcattctatct gagtagtcaa 3150
aatacaagta aaggagagca aataaacaac atttgaaaa aaaaaaaaaa 3200
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 3250
aaaaaaaaa aaaaaa 3265

<210> SEQ ID NO 70
<211> LENGTH: 919
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 70

Met Gly Leu Phe Arg Gly Phe Val Phe Leu Leu Val Leu Cys Leu
1 5 10 15
Leu His Gln Ser Asn Thr Ser Phe Ile Lys Leu Asn Asn Asn Gly
20 25 30
Phe Glu Asp Ile Val Ile Val Ile Asp Pro Ser Val Pro Glu Asp
35 40 45
Glu Lys Ile Ile Glu Gln Ile Glu Asp Met Val Thr Thr Ala Ser
50 55 60
Thr Tyr Leu Phe Glu Ala Thr Glu Lys Arg Phe Phe Phe Lys Asn
65 70 75
Val Ser Ile Leu Ile Pro Glu Asn Trp Lys Glu Asn Pro Gln Tyr
80 85 90
Lys Arg Pro Lys His Glu Asn His Lys His Ala Asp Val Ile Val
95 100 105
Ala Pro Pro Thr Leu Pro Gly Arg Asp Glu Pro Tyr Thr Lys Gln
110 115 120
Phe Thr Glu Cys Gly Glu Lys Gly Glu Tyr Ile His Phe Thr Pro
125 130 135
Asp Leu Leu Leu Gly Lys Lys Gln Asn Glu Tyr Gly Pro Pro Gly
140 145 150
Lys Leu Phe Val His Glu Trp Ala His Leu Arg Trp Gly Val Phe
155 160 165
Asp Glu Tyr Asn Glu Asp Gln Pro Phe Tyr Arg Ala Lys Ser Lys
170 175 180
Lys Ile Glu Ala Thr Arg Cys Ser Ala Gly Ile Ser Gly Arg Asn
185 190 195
Arg Val Tyr Lys Cys Gln Gly Gly Ser Cys Leu Ser Arg Ala Cys
200 205 210
Arg Ile Asp Ser Thr Thr Lys Leu Tyr Gly Lys Asp Cys Gln Phe
215 220 225
Phe Pro Asp Lys Val Gln Thr Glu Lys Ala Ser Ile Met Phe Met
230 235 240
Gln Ser Ile Asp Ser Val Val Glu Phe Cys Asn Glu Lys Thr His
245 250 255
Asn Gln Glu Ala Pro Ser Leu Gln Asn Ile Lys Cys Asn Phe Arg
260 265 270
Ser Thr Trp Glu Val Ile Ser Asn Ser Glu Asp Phe Lys Asn Thr
275 280 285
Ile Pro Met Val Thr Pro Pro Pro Pro Pro Val Phe Ser Leu Leu
290 295 300

-continued

Lys	Ile	Ser	Gln	Arg	Ile	Val	Cys	Leu	Val	Leu	Asp	Lys	Ser	Gly	305	310	315
Ser	Met	Gly	Gly	Lys	Asp	Arg	Leu	Asn	Arg	Met	Asn	Gln	Ala	Ala	320	325	330
Lys	His	Phe	Leu	Leu	Gln	Thr	Val	Glu	Asn	Gly	Ser	Trp	Val	Gly	335	340	345
Met	Val	His	Phe	Asp	Ser	Thr	Ala	Thr	Ile	Val	Asn	Lys	Leu	Ile	350	355	360
Gln	Ile	Lys	Ser	Ser	Asp	Glu	Arg	Asn	Thr	Leu	Met	Ala	Gly	Leu	365	370	375
Pro	Thr	Tyr	Pro	Leu	Gly	Gly	Thr	Ser	Ile	Cys	Ser	Gly	Ile	Lys	380	385	390
Tyr	Ala	Phe	Gln	Val	Ile	Gly	Glu	Leu	His	Ser	Gln	Leu	Asp	Gly	395	400	405
Ser	Glu	Val	Leu	Leu	Leu	Thr	Asp	Gly	Glu	Asp	Asn	Thr	Ala	Ser	410	415	420
Ser	Cys	Ile	Asp	Glu	Val	Lys	Gln	Ser	Gly	Ala	Ile	Val	His	Phe	425	430	435
Ile	Ala	Leu	Gly	Arg	Ala	Ala	Asp	Glu	Ala	Val	Ile	Glu	Met	Ser	440	445	450
Lys	Ile	Thr	Gly	Gly	Ser	His	Phe	Tyr	Val	Ser	Asp	Glu	Ala	Gln	455	460	465
Asn	Asn	Gly	Leu	Ile	Asp	Ala	Phe	Gly	Ala	Leu	Thr	Ser	Gly	Asn	470	475	480
Thr	Asp	Leu	Ser	Gln	Lys	Ser	Leu	Gln	Leu	Glu	Ser	Lys	Gly	Leu	485	490	495
Thr	Leu	Asn	Ser	Asn	Ala	Trp	Met	Asn	Asp	Thr	Val	Ile	Ile	Asp	500	505	510
Ser	Thr	Val	Gly	Lys	Asp	Thr	Phe	Phe	Leu	Ile	Thr	Trp	Asn	Ser	515	520	525
Leu	Pro	Pro	Ser	Ile	Ser	Leu	Trp	Asp	Pro	Ser	Gly	Thr	Ile	Met	530	535	540
Glu	Asn	Phe	Thr	Val	Asp	Ala	Thr	Ser	Lys	Met	Ala	Tyr	Leu	Ser	545	550	555
Ile	Pro	Gly	Thr	Ala	Lys	Val	Gly	Thr	Trp	Ala	Tyr	Asn	Leu	Gln	560	565	570
Ala	Lys	Ala	Asn	Pro	Glu	Thr	Leu	Thr	Ile	Thr	Val	Thr	Ser	Arg	575	580	585
Ala	Ala	Asn	Ser	Ser	Val	Pro	Pro	Ile	Thr	Val	Asn	Ala	Lys	Met	590	595	600
Asn	Lys	Asp	Val	Asn	Ser	Phe	Pro	Ser	Pro	Met	Ile	Val	Tyr	Ala	605	610	615
Glu	Ile	Leu	Gln	Gly	Tyr	Val	Pro	Val	Leu	Gly	Ala	Asn	Val	Thr	620	625	630
Ala	Phe	Ile	Glu	Ser	Gln	Asn	Gly	His	Thr	Glu	Val	Leu	Glu	Leu	635	640	645
Leu	Asp	Asn	Gly	Ala	Gly	Ala	Asp	Ser	Phe	Lys	Asn	Asp	Gly	Val	650	655	660
Tyr	Ser	Arg	Tyr	Phe	Thr	Ala	Tyr	Thr	Glu	Asn	Gly	Arg	Tyr	Ser	665	670	675

-continued

Leu	Lys	Val	Arg	Ala	His	Gly	Gly	Ala	Asn	Thr	Ala	Arg	Leu	Lys
				680					685					690
Leu	Arg	Pro	Pro	Leu	Asn	Arg	Ala	Ala	Tyr	Ile	Pro	Gly	Trp	Val
				695					700					705
Val	Asn	Gly	Glu	Ile	Glu	Ala	Asn	Pro	Pro	Arg	Pro	Glu	Ile	Asp
				710					715					720
Glu	Asp	Thr	Gln	Thr	Thr	Leu	Glu	Asp	Phe	Ser	Arg	Thr	Ala	Ser
				725					730					735
Gly	Gly	Ala	Phe	Val	Val	Ser	Gln	Val	Pro	Ser	Leu	Pro	Leu	Pro
				740					745					750
Asp	Gln	Tyr	Pro	Pro	Ser	Gln	Ile	Thr	Asp	Leu	Asp	Ala	Thr	Val
				755					760					765
His	Glu	Asp	Lys	Ile	Ile	Leu	Thr	Trp	Thr	Ala	Pro	Gly	Asp	Asn
				770					775					780
Phe	Asp	Val	Gly	Lys	Val	Gln	Arg	Tyr	Ile	Ile	Arg	Ile	Ser	Ala
				785					790					795
Ser	Ile	Leu	Asp	Leu	Arg	Asp	Ser	Phe	Asp	Asp	Ala	Leu	Gln	Val
				800					805					810
Asn	Thr	Thr	Asp	Leu	Ser	Pro	Lys	Glu	Ala	Asn	Ser	Lys	Glu	Ser
				815					820					825
Phe	Ala	Phe	Lys	Pro	Glu	Asn	Ile	Ser	Glu	Glu	Asn	Ala	Thr	His
				830					835					840
Ile	Phe	Ile	Ala	Ile	Lys	Ser	Ile	Asp	Lys	Ser	Asn	Leu	Thr	Ser
				845					850					855
Lys	Val	Ser	Asn	Ile	Ala	Gln	Val	Thr	Leu	Phe	Ile	Pro	Gln	Ala
				860					865					870
Asn	Pro	Asp	Asp	Ile	Asp	Pro	Thr	Pro	Thr	Pro	Thr	Pro	Thr	Pro
				875					880					885
Thr	Pro	Asp	Lys	Ser	His	Asn	Ser	Gly	Val	Asn	Ile	Ser	Thr	Leu
				890					895					900
Val	Leu	Ser	Val	Ile	Gly	Ser	Val	Val	Ile	Val	Asn	Phe	Ile	Leu
				905					910					915

Ser Thr Thr Ile

<210> SEQ ID NO 71

<211> LENGTH: 3877

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 71

ctccttaggt ggaaaccctg ggagtagagt actgacagca aagaccggga	50
aagaccatac gtccccgggc aggggtgaca acaggtgtca tctttttgat	100
ctcgtgtgtg gctgccttcc tatttcaagg aaagacgcca aggtaatttt	150
gaccagagg agcaatgatg tagccacctc ctaaccttcc cttcttgaac	200
ccccagtat gccaggattt actagagagt gtcaactcaa ccagcaagcg	250
gtccttcggt cttaacttgt ggttgaggga gagaacctt gtggggctgc	300
gttctcttag cagtgtcag aagtgacttg cctgaggggtg gaccagaaga	350
aaggaaaggt cccctcttgc tgttgctgc acatcaggaa ggctgtgatg	400
ggaatgaagg tgaaaacttg gagatttcac ttcagtcatt gcttctgcct	450

-continued

gcaagatcat cctttaaag tagagaagct gctctgtgtg gtggttaact	500
ccaagaggca gaactcgttc tagaaggaaa tggatgcaag cagctccggg	550
ggccccaac gcattgcttc tgtggtctag ccaggggaag cccttccgtg	600
ggggcccggt ctttagagga tgccaccggt tctggacgca tggctgattc	650
ctgaatgatg atggttcgcc gggggctgct tgcgtgatt tccgggtgg	700
tggttttgct ggtgctcctc tgctgtgcta tctctgtcct gtacatgttg	750
gcctgcacc caaaaggtag cgaggagcag ctggcactgc ccagggccaa	800
cagccccacg ggaaggagg ggtaccaggc cgtccttcag gaggggagg	850
agcagcaccg caactacgtg agcagcctga agcggcagat cgcacagctc	900
aaggaggagc tgcaggagag gaggtagcag ctccaggaatg ggcagtacca	950
agccagcgat gctgctggcc tgggtctgga caggagcccc ccagagaaaa	1000
cccaggccga cctcctggcc ttcctgcact cgcagggtga caaggcagag	1050
gtgaatgtg gcgtcaagct ggccacagag tatgcagcag tgcctttcga	1100
tagctttact ctacagaagg tgtaccagct ggagactggc cttaccgcc	1150
accccgagga gaagcctgtg aggaaggaca agcgggatga gttggtgaa	1200
gccattgaat cagccttga gaccctgaac aatcctgcag agaacagccc	1250
caatcacctg ccttacacgg cctctgattt catagaagg atctaccgaa	1300
cagaaaggga caaagggaca ttgtatgagc tcaccttcaa aggggaccac	1350
aaacacgaat tcaaacggct catcttattt cgaccattca gccccatcat	1400
gaaagtgaat aatgaaaagc tcaacatggc caacacgctt atcaatgtta	1450
tcgtgcctct agcaaaaagg gtggacaagt tccggcagtt catgcagaat	1500
ttcagggaga tgtgcattga gcaggatggg agagtccatc tcactgttgt	1550
ttactttggg aaagaagaaa taaatgaagt caaaggaata cttgaaaaca	1600
cttccaaagc tgccaacttc aggaacttta ccttcattcca gctgaatgga	1650
gaattttctc ggggaaaggg acttgatgtt ggagcccgtc tctggaaggg	1700
aagcaacgtc cttctctttt tctgtgatgt ggacatctac ttcacatctg	1750
aattcctcaa tacgtgtagg ctgaatacac agccaggga gaaggtattt	1800
tatccagttc ttttcagtca gtacaatcct ggcataatat acggccacca	1850
tgatgcagtc cctcccttgg aacagcagct ggtcataaag aaggaaactg	1900
gattttggag agactttgga tttgggatga cgtgtcagta tcggtcagac	1950
ttcatcaata taggtgggtt tgatctggac atcaaaggct ggggcggaga	2000
ggatgtgcac ctttatcgca agtatctcca cagcaacctc atagtggtag	2050
ggacgcctgt gcgaggactc ttccacctct ggcattgaga gcgctgcattg	2100
gacgagctga ccccgagca gtacaagatg tgcattgcatt ccaaggccat	2150
gaacgaggca tcccacggcc agctgggcat gctggtgttc aggcacgaga	2200
tagaggctca ccttcgcaaa cagaaacaga agacaagtag caaaaaaaca	2250
tgaactccca gagaaggatt gtgggagaca ctttttcttt ctttttgcaa	2300
ttactgaaag tggctgcaac agagaaaaga cttccataaa ggacgacaaa	2350

-continued

agaattggac tgatgggtca gagatgagaa agcctccgat ttctctctgt	2400
tgggctttttt acaacagaaa tcaaaatctc cgctttgcct gcaaaagtaa	2450
cccagttgca ccctgtgaag tgtctgacaa aggcagaatg cttgtgagat	2500
tataagccta atgggtgtga ggttttgatg gtgtttacaa tacactgaga	2550
cctgttgtttt tgtgtgctca ttgaaatatt catgatttaa gagcagtttt	2600
gtaaaaaatt cattagcatg aaaggcaagc atatttctcc tcatatgaat	2650
gagcctatca gcagggctct agtttctagg aatgctaaaa tatcagaagg	2700
caggagagga gataggctta ttatgatact agtgagtaca ttaagtaaaa	2750
taaaatggac cagaaaagaa aagaacatc aaatatcgtg tcatattttc	2800
cccaagatta accaaaaata atctgcttat ctttttggtt gtccttttaa	2850
ctgtctccgt ttttttcttt tatttaaaaa tgcacttttt ttcccttggt	2900
agttatatgc tgcttattta attaccactt tgcaagcctt acaagagagc	2950
acaagttggc ctacattttt atatttttta agaagatact ttgagatgca	3000
ttatgagaac tttcagttca aagcatcaaa ttgatgccat atccaaggac	3050
atgccaaatg ctgattctgt caggcactga atgtcaggca ttgagacata	3100
gggaaggaat ggtttgtact aatacagacg tacagatact ttctctgaag	3150
agtattttcg aagaggagca actgaacact ggaggaaaag aaaatgacac	3200
tttctgcttt acagaaaagg aaactcatc agactggtga tatcgtgatg	3250
tacctaaaag tcagaaacca cattttctcc tcagaagtag ggaccgcttt	3300
cttacctgtt taaataaacc aaagtatacc gtgtgaacca aacaatctct	3350
tttcaaaaaca ggggtgctcct cctggcttct ggcttcata agaagaaatg	3400
gagaaaaata tatatatata tatatatatt gtgaaagatc aatccatctg	3450
ccagaatcta gtgggatgga agtttttgct acatgttatc caccocaggc	3500
cagggtggaag taactgaatt attttttaaa ttaagcagtt ctactcaatc	3550
accaagatgc ttctgaaaat tgcattttat taccatttca aactattttt	3600
taaaaataaa tacagttaac atagagtggg ttcttcattc atgtgaaaat	3650
tattagccag caccagatgc atgagctaata tatctctttg agtccttgct	3700
tctgtttgct cacagtaaac tcattgttta aaagcttcaa gaacattcaa	3750
gctgttggtg tgtaaaaaa tgcattgtat tgatttgtac tggtagttta	3800
tgaaatttaa ttaaacaca ggccatgaat ggaaggtggt attgcacagc	3850
taataaaaata tgatttggtg atatgaa	3877

<210> SEQ ID NO 72

<211> LENGTH: 532

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 72

Met	Met	Met	Val	Arg	Arg	Gly	Leu	Leu	Ala	Trp	Ile	Ser	Arg	Val
1				5					10					15

Val	Val	Leu	Leu	Val	Leu	Leu	Cys	Cys	Ala	Ile	Ser	Val	Leu	Tyr
				20					25					30

-continued

Met	Leu	Ala	Cys	Thr	Pro	Lys	Gly	Asp	Glu	Glu	Gln	Leu	Ala	Leu	35	40	45
Pro	Arg	Ala	Asn	Ser	Pro	Thr	Gly	Lys	Glu	Gly	Tyr	Gln	Ala	Val	50	55	60
Leu	Gln	Glu	Trp	Glu	Glu	Gln	His	Arg	Asn	Tyr	Val	Ser	Ser	Leu	65	70	75
Lys	Arg	Gln	Ile	Ala	Gln	Leu	Lys	Glu	Glu	Leu	Gln	Glu	Arg	Ser	80	85	90
Glu	Gln	Leu	Arg	Asn	Gly	Gln	Tyr	Gln	Ala	Ser	Asp	Ala	Ala	Gly	95	100	105
Leu	Gly	Leu	Asp	Arg	Ser	Pro	Pro	Glu	Lys	Thr	Gln	Ala	Asp	Leu	110	115	120
Leu	Ala	Phe	Leu	His	Ser	Gln	Val	Asp	Lys	Ala	Glu	Val	Asn	Ala	125	130	135
Gly	Val	Lys	Leu	Ala	Thr	Glu	Tyr	Ala	Ala	Val	Pro	Phe	Asp	Ser	140	145	150
Phe	Thr	Leu	Gln	Lys	Val	Tyr	Gln	Leu	Glu	Thr	Gly	Leu	Thr	Arg	155	160	165
His	Pro	Glu	Glu	Lys	Pro	Val	Arg	Lys	Asp	Lys	Arg	Asp	Glu	Leu	170	175	180
Val	Glu	Ala	Ile	Glu	Ser	Ala	Leu	Glu	Thr	Leu	Asn	Asn	Pro	Ala	185	190	195
Glu	Asn	Ser	Pro	Asn	His	Arg	Pro	Tyr	Thr	Ala	Ser	Asp	Phe	Ile	200	205	210
Glu	Gly	Ile	Tyr	Arg	Thr	Glu	Arg	Asp	Lys	Gly	Thr	Leu	Tyr	Glu	215	220	225
Leu	Thr	Phe	Lys	Gly	Asp	His	Lys	His	Glu	Phe	Lys	Arg	Leu	Ile	230	235	240
Leu	Phe	Arg	Pro	Phe	Ser	Pro	Ile	Met	Lys	Val	Lys	Asn	Glu	Lys	245	250	255
Leu	Asn	Met	Ala	Asn	Thr	Leu	Ile	Asn	Val	Ile	Val	Pro	Leu	Ala	260	265	270
Lys	Arg	Val	Asp	Lys	Phe	Arg	Gln	Phe	Met	Gln	Asn	Phe	Arg	Glu	275	280	285
Met	Cys	Ile	Glu	Gln	Asp	Gly	Arg	Val	His	Leu	Thr	Val	Val	Tyr	290	295	300
Phe	Gly	Lys	Glu	Glu	Ile	Asn	Glu	Val	Lys	Gly	Ile	Leu	Glu	Asn	305	310	315
Thr	Ser	Lys	Ala	Ala	Asn	Phe	Arg	Asn	Phe	Thr	Phe	Ile	Gln	Leu	320	325	330
Asn	Gly	Glu	Phe	Ser	Arg	Gly	Lys	Gly	Leu	Asp	Val	Gly	Ala	Arg	335	340	345
Phe	Trp	Lys	Gly	Ser	Asn	Val	Leu	Leu	Phe	Phe	Cys	Asp	Val	Asp	350	355	360
Ile	Tyr	Phe	Thr	Ser	Glu	Phe	Leu	Asn	Thr	Cys	Arg	Leu	Asn	Thr	365	370	375
Gln	Pro	Gly	Lys	Lys	Val	Phe	Tyr	Pro	Val	Leu	Phe	Ser	Gln	Tyr	380	385	390
Asn	Pro	Gly	Ile	Ile	Tyr	Gly	His	His	Asp	Ala	Val	Pro	Pro	Leu	395	400	405
Glu	Gln	Gln	Leu	Val	Ile	Lys	Lys	Glu	Thr	Gly	Phe	Trp	Arg	Asp			

-continued

410	415	420
Phe Gly Phe Gly Met Thr Cys Gln Tyr	Arg Ser Asp Phe Ile Asn	
425	430	435
Ile Gly Gly Phe Asp Leu Asp Ile Lys	Gly Trp Gly Gly Glu Asp	
440	445	450
Val His Leu Tyr Arg Lys Tyr Leu His	Ser Asn Leu Ile Val Val	
455	460	465
Arg Thr Pro Val Arg Gly Leu Phe His	Leu Trp His Glu Lys Arg	
470	475	480
Cys Met Asp Glu Leu Thr Pro Glu Gln	Tyr Lys Met Cys Met Gln	
485	490	495
Ser Lys Ala Met Asn Glu Ala Ser His	Gly Gln Leu Gly Met Leu	
500	505	510
Val Phe Arg His Glu Ile Glu Ala His	Leu Arg Lys Gln Lys Gln	
515	520	525
Lys Thr Ser Ser Lys Lys Thr		
530		

<210> SEQ ID NO 73
<211> LENGTH: 1701
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien
<220> FEATURE:
<221> NAME/KEY: unsure
<222> LOCATION: 1528
<223> OTHER INFORMATION: unknown base

<400> SEQUENCE: 73

gagactgcag agggagataa agagagaggg caaagaggca gcaagagatt	50
tgctctgggg atccagaaac ccatgatacc ctactgaaca ccgaatcccc	100
tggaagccca cagagacaga gacagcaaga gaagcagaga taaatacact	150
cacgccagga gctcgctcgc tctctctctc tctctctcac tctccctcc	200
ctctctctct gctgtccta gtccctctagt cctcaaattc ccagtcccct	250
gcaccccttc ctgggacact atgttggtct cgcacctcct gctggagggtg	300
atttgatcc tggctgcaga tgggggtcaa cactggacgt atgagggccc	350
acatggtcag gaccattggc cagcctctta ccctgagtgt ggaacaatg	400
cccagtcgcc catcgatatt cagacagaca gtgtgacatt tgaccctgat	450
ttgctgtctc tgcagcccca cggatatgac cagcctggca ccgagccttt	500
ggacctgcac aacaatggcc acacagtgca actctctctg ccctctaccc	550
tgtatctggg tggacttccc cgaaaatatg tagctgcca gtcacacctg	600
cactggggtc agaaaggatc cccagggggg tcagaacacc agatcaacag	650
tgaagccaca ttgcagagc tccacattgt acattatgac tctgattcct	700
atgacagctt gagtgaggct gctgagaggc ctcagggcct ggctgtcctg	750
ggcatcctaa ttgagtgagg tgagactaag aatatagctt atgaacacat	800
tctgagtcac ttgcatgaag tcaggcataa agatcagaag acctcagtgc	850
ctcccttcaa cctaagagag ctgctcccca aacagctggg gcagtacttc	900
cgctacaatg gctcgctcac aactccccct tgctaccaga gtgtgctctg	950

-continued

gacagttttt tatagaaggt cccagatttc aatggaacag ctggaaaagc	1000
ttcaggggac attgttctcc acagaagagg agccctctaa gcttctggta	1050
cagaactacc gagcccttca gcctctcaat cagcgcatgg tctttgcttc	1100
tttcatccaa gcaggatcct cgtataccac aggtgaaatg ctgagtctag	1150
gtgtaggaat cttggttggc tgtctctgcc ttctcctggc tgtttatttc	1200
attgctagaa agattcggaa gaagaggctg gaaaaccgaa agagtgtggt	1250
cttcacctca gcacaagcca cgactgaggc ataaattcct tctcagatac	1300
catggatgtg gatgacttcc cttcatgcct atcaggaagc ctctaaaatg	1350
gggtgtagga tctggccaga aacactgtag gagtagtaag cagatgtcct	1400
ccttcccctg gacatctctt agagaggaat ggaccaggc tgcattcca	1450
ggaagaactg cagagccttc agcctctcca aacatgtagg aggaaatgag	1500
gaaatcgctg tgttgtaaat gcagaganca aactctgttt agttgcaggg	1550
gaagtttggg atatacccca aagtcctcta cccctcact tttatggccc	1600
tttccctaga tatactgcgg gatctctcct taggataaag agttgctgtt	1650
gaagttgtat atttttgatc aatatatttg gaaattaaag tttctgactt	1700
t	1701

<210> SEQ ID NO 74

<211> LENGTH: 337

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 74

Met	Leu	Phe	Ser	Ala	Leu	Leu	Leu	Glu	Val	Ile	Trp	Ile	Leu	Ala
1				5					10					15
Ala	Asp	Gly	Gly	Gln	His	Trp	Thr	Tyr	Glu	Gly	Pro	His	Gly	Gln
				20					25					30
Asp	His	Trp	Pro	Ala	Ser	Tyr	Pro	Glu	Cys	Gly	Asn	Asn	Ala	Gln
				35					40					45
Ser	Pro	Ile	Asp	Ile	Gln	Thr	Asp	Ser	Val	Thr	Phe	Asp	Pro	Asp
				50					55					60
Leu	Pro	Ala	Leu	Gln	Pro	His	Gly	Tyr	Asp	Gln	Pro	Gly	Thr	Glu
				65					70					75
Pro	Leu	Asp	Leu	His	Asn	Asn	Gly	His	Thr	Val	Gln	Leu	Ser	Leu
				80					85					90
Pro	Ser	Thr	Leu	Tyr	Leu	Gly	Gly	Leu	Pro	Arg	Lys	Tyr	Val	Ala
				95					100					105
Ala	Gln	Leu	His	Leu	His	Trp	Gly	Gln	Lys	Gly	Ser	Pro	Gly	Gly
				110					115					120
Ser	Glu	His	Gln	Ile	Asn	Ser	Glu	Ala	Thr	Phe	Ala	Glu	Leu	His
				125					130					135
Ile	Val	His	Tyr	Asp	Ser	Asp	Ser	Tyr	Asp	Ser	Leu	Ser	Glu	Ala
				140					145					150
Ala	Glu	Arg	Pro	Gln	Gly	Leu	Ala	Val	Leu	Gly	Ile	Leu	Ile	Glu
				155					160					165
Val	Gly	Glu	Thr	Lys	Asn	Ile	Ala	Tyr	Glu	His	Ile	Leu	Ser	His
				170					175					180

-continued

Leu	His	Glu	Val	Arg	His	Lys	Asp	Gln	Lys	Thr	Ser	Val	Pro	Pro	
				185					190					195	
Phe	Asn	Leu	Arg	Glu	Leu	Leu	Pro	Lys	Gln	Leu	Gly	Gln	Tyr	Phe	
				200					205					210	
Arg	Tyr	Asn	Gly	Ser	Leu	Thr	Thr	Pro	Pro	Cys	Tyr	Gln	Ser	Val	
				215					220					225	
Leu	Trp	Thr	Val	Phe	Tyr	Arg	Arg	Ser	Gln	Ile	Ser	Met	Glu	Gln	
				230					235					240	
Leu	Glu	Lys	Leu	Gln	Gly	Thr	Leu	Phe	Ser	Thr	Glu	Glu	Glu	Pro	
				245					250					255	
Ser	Lys	Leu	Leu	Val	Gln	Asn	Tyr	Arg	Ala	Leu	Gln	Pro	Leu	Asn	
				260					265					270	
Gln	Arg	Met	Val	Phe	Ala	Ser	Phe	Ile	Gln	Ala	Gly	Ser	Ser	Tyr	
				275					280					285	
Thr	Thr	Gly	Glu	Met	Leu	Ser	Leu	Gly	Val	Gly	Ile	Leu	Val	Gly	
				290					295					300	
Cys	Leu	Cys	Leu	Leu	Leu	Ala	Val	Tyr	Phe	Ile	Ala	Arg	Lys	Ile	
				305					310					315	
Arg	Lys	Lys	Arg	Leu	Glu	Asn	Arg	Lys	Ser	Val	Val	Phe	Thr	Ser	
				320					325					330	
Ala	Gln	Ala	Thr	Thr	Glu	Ala									
				335											

<210> SEQ ID NO 75
<211> LENGTH: 1743
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 75

tgccgctgcc gccgctgctg ctgttgctcc tggcggcgcc ttggggacgg	50
gcagttccct gtgtctctgg tggtttgctt aaacctgcaa acatcacctt	100
cttatccatc aacatgaaga atgtcctaca atggactcca ccagagggtc	150
ttcaaggagt taaagttact tacactgtgc agtatttcat cacaaattgg	200
cccaccagag gtggcactga ctacagatga gaagtccatt tctgttgctc	250
tgacagctcc agagaagtgg aagagaaatc cagaagacct tcctgtttcc	300
atgcaacaaa tatactccaa tctgaagtat aacgtgtctg tgttgaatac	350
taaatcaaac agaacgtggt ccagtggtgt gaccaaccac acgctggtgc	400
tcacctggct ggagccgaac actctttact gcgtacacgt ggagtccttc	450
gtcccagggc ccctcgcgcg tgctcagcct tctgagaagc agtgtgccag	500
gactttgaaa gatcaatcat cagagttcaa ggctaaaatc atcttctggt	550
atgttttgcc catatctatt accgtgttct ttttttctgt gatgggctat	600
tccatctacc gatatatcca cgttgcaaaa gagaaacacc cagcaaattt	650
gattttgatt tatggaaatg aatttgacaa aagattcttt gtgcctgctg	700
aaaaaatcgt gattaacttt atcacctca atatctcgga tgattctaaa	750
atttctcatc aggatatgag ttactggga aaaagcagtg atgtatccag	800
ccttaatgat cctcagccca gcgggaacct gagggcccct caggagggaag	850
aggaggtgaa acatttaggg tatgcttcgc atttgatgga aattttttgt	900

-continued

gactctgaag aaaacacgga aggtacttct ctcacccagc aagagtccct 950
cagcagaaca atacccccgg ataaaacagt cattgaatat gaatatgatg 1000
tcagaaccac tgacatttgt gcggggcctg aagagcagga gctcagtttg 1050
caggaggagg tgtccacaca aggaacatta ttggagtcgc aggcagcggt 1100
ggcagtcctg ggcccgcaaa cgttacagta ctcatacacc cctcagctcc 1150
aagacttaga cccctggcg caggagcaca cagactcgga ggaggggccc 1200
gaggaagagc catcgacgac cctggtcgac tgggatcccc aaactggcag 1250
gctgtgtatt ccttcgctgt ccagcttcga ccaggattca gagggctgcg 1300
agccttctga gggggatggg ctcgagagg agggctctct atctagactc 1350
tatgaggagc cggtccaga caggccacca ggagaaaatg aaacctatct 1400
catgcaattc atggaggaat gggggttata tgtgcagatg gaaaactgat 1450
gccaacactt ccttttgcc tttgtttcct gtgcaacaa gtgagtcacc 1500
cctttgatcc cagccataaa gtacctggga tgaaagaagt tttttccagt 1550
ttgtcagtg ctgtgagaat tacttatttc ttttctctat tctcatagca 1600
cgtgtgtgat tggttcatgc atgtaggct ctaacaatg atgggtggcc 1650
tctggagtcc aggggctggc cggttgttct atgcagagaa agcagtcaat 1700
aatgtttgc cagactgggt gcagaattta ttcaggtggg tgt 1743

<210> SEQ ID NO 76

<211> LENGTH: 442

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 76

Met Ser Tyr Asn Gly Leu His Gln Arg Val Phe Lys Glu Leu Lys
1 5 10 15
Leu Leu Thr Leu Cys Ser Ile Ser Ser Gln Ile Gly Pro Pro Glu
20 25 30
Val Ala Leu Thr Thr Asp Glu Lys Ser Ile Ser Val Val Leu Thr
35 40 45
Ala Pro Glu Lys Trp Lys Arg Asn Pro Glu Asp Leu Pro Val Ser
50 55 60
Met Gln Gln Ile Tyr Ser Asn Leu Lys Tyr Asn Val Ser Val Leu
65 70 75
Asn Thr Lys Ser Asn Arg Thr Trp Ser Gln Cys Val Thr Asn His
80 85 90
Thr Leu Val Leu Thr Trp Leu Glu Pro Asn Thr Leu Tyr Cys Val
95 100 105
His Val Glu Ser Phe Val Pro Gly Pro Pro Arg Arg Ala Gln Pro
110 115 120
Ser Glu Lys Gln Cys Ala Arg Thr Leu Lys Asp Gln Ser Ser Glu
125 130 135
Phe Lys Ala Lys Ile Ile Phe Trp Tyr Val Leu Pro Ile Ser Ile
140 145 150
Thr Val Phe Leu Phe Ser Val Met Gly Tyr Ser Ile Tyr Arg Tyr
155 160 165

-continued

Ile	His	Val	Gly	Lys	Glu	Lys	His	Pro	Ala	Asn	Leu	Ile	Leu	Ile	
			170						175					180	
Tyr	Gly	Asn	Glu	Phe	Asp	Lys	Arg	Phe	Phe	Val	Pro	Ala	Glu	Lys	
			185						190					195	
Ile	Val	Ile	Asn	Phe	Ile	Thr	Leu	Asn	Ile	Ser	Asp	Asp	Ser	Lys	
			200						205					210	
Ile	Ser	His	Gln	Asp	Met	Ser	Leu	Leu	Gly	Lys	Ser	Ser	Asp	Val	
			215						220					225	
Ser	Ser	Leu	Asn	Asp	Pro	Gln	Pro	Ser	Gly	Asn	Leu	Arg	Pro	Pro	
			230						235					240	
Gln	Glu	Glu	Glu	Glu	Val	Lys	His	Leu	Gly	Tyr	Ala	Ser	His	Leu	
			245						250					255	
Met	Glu	Ile	Phe	Cys	Asp	Ser	Glu	Glu	Asn	Thr	Glu	Gly	Thr	Ser	
			260						265					270	
Leu	Thr	Gln	Gln	Glu	Ser	Leu	Ser	Arg	Thr	Ile	Pro	Pro	Asp	Lys	
			275						280					285	
Thr	Val	Ile	Glu	Tyr	Glu	Tyr	Asp	Val	Arg	Thr	Thr	Asp	Ile	Cys	
			290						295					300	
Ala	Gly	Pro	Glu	Glu	Gln	Glu	Leu	Ser	Leu	Gln	Glu	Glu	Val	Ser	
			305						310					315	
Thr	Gln	Gly	Thr	Leu	Leu	Glu	Ser	Gln	Ala	Ala	Leu	Ala	Val	Leu	
			320						325					330	
Gly	Pro	Gln	Thr	Leu	Gln	Tyr	Ser	Tyr	Thr	Pro	Gln	Leu	Gln	Asp	
			335						340					345	
Leu	Asp	Pro	Leu	Ala	Gln	Glu	His	Thr	Asp	Ser	Glu	Glu	Gly	Pro	
			350						355					360	
Glu	Glu	Glu	Pro	Ser	Thr	Thr	Leu	Val	Asp	Trp	Asp	Pro	Gln	Thr	
			365						370					375	
Gly	Arg	Leu	Cys	Ile	Pro	Ser	Leu	Ser	Ser	Phe	Asp	Gln	Asp	Ser	
			380						385					390	
Glu	Gly	Cys	Glu	Pro	Ser	Glu	Gly	Asp	Gly	Leu	Gly	Glu	Glu	Gly	
			395						400					405	
Leu	Leu	Ser	Arg	Leu	Tyr	Glu	Glu	Pro	Ala	Pro	Asp	Arg	Pro	Pro	
			410						415					420	
Gly	Glu	Asn	Glu	Thr	Tyr	Leu	Met	Gln	Phe	Met	Glu	Glu	Trp	Gly	
			425						430					435	
Leu	Tyr	Val	Gln	Met	Glu	Asn									
			440												

<210> SEQ ID NO 77
<211> LENGTH: 1636
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 77

gaggagcggg ccgaggactc cagcgtgccc aggtctggca tcctgcactt	50
gctgccctct gacacctggg aagatggccg gcccgtagac cttcaccctt	100
ctctgtggtt tgctggcagc caccttgatc caagccaccc tcagtccac	150
tgcatgtctc atcctcggcc caaaagtcat caaagaaaag ctgacacagg	200
agctgaagga ccacaacgcc accagcatcc tgcagcagct gccgctgctc	250
agtgccatgc gggaaaagcc agccggaggc atccctgtgc tgggcagcct	300

-continued

ggtgaacacc gtcctgaagc acatcatctg gctgaaggtc atcacagcta 350
acatcctcca gctgcaggtg aagccctcgg ccaatgacca ggagctgcta 400
gtcaagatcc ccctggacat ggtggctgga ttcaacacgc ccctgggtcaa 450
gaccatcgtg gagtccaca tgacgactga ggcccaagcc accatccgca 500
tggacaccag tgcaagtggc cccacccgcc tggctcctcag tgactgtgcc 550
accagccatg ggagcctgcg catccaactg ctgtataagc tctccttcct 600
ggtgaagccc ttagctaagc aggtcatgaa cctcctagtg ccatccctgc 650
ccaatctagt gaaaaaccag ctgtgtcccg tgatcgaggc ttccttcaat 700
ggcatgtatg cagacctcct gcagctgggtg aaggtgccca ttccctcag 750
cattgacctg ctgagtttg accttctgta tcctgccatc aagggtgaca 800
ccattcagct ctacctgggg gccaaagttgt tggactcaca gggaaaggtg 850
accaagtggg tcaataactc tgacgcttcc ctgacaatgc ccaccctgga 900
caacatcccg ttcagcctca tcgtgagtca ggacgtgggtg aaagctgcag 950
tggctgctgt gctctctcca gaagaattca tggctcctgtt ggactctgtg 1000
cttcctgaga gtgcccatcg gctgaagtca agcatcgggc tgatcaatga 1050
aaaggctgca gataagctgg gatctacca gatcgtgaag atcctaactc 1100
aggacactcc cgagtttttt atagaccaag gccatgccaa ggtggcccaa 1150
ctgatcgtgc tggaagtgtt tccctccagt gaagccctcc gccctttgtt 1200
caccctgggc atcgaagcca gctcggaagc tcagttttac accaaaggtg 1250
accaacttat actcaacttg aataacatca gctctgatcg gatccagctg 1300
atgaactctg ggattggctg gttccaacct gatgttctga aaaacatcat 1350
cactgagatc atccactcca tcctgctgcc gaaccagaat ggcaaattaa 1400
gatctggggc ccagtgctca ttggtgaagg ccttgggatt cgaggcagct 1450
gagtcctcac tgaccaagga tgcccttggtg cttactccag cctccttggtg 1500
gaaaccacgc tctcctgtct ccagtgaaag acttggatgg cagccatcag 1550
ggaaggctgg gtcccagctg ggagtatggg tgtgagctct atagaccatc 1600
cctctctgca atcaataaac acttgctctgt gaaaaa 1636

<210> SEQ ID NO 78
<211> LENGTH: 484
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 78

Met Ala Gly Pro Trp Thr Phe Thr Leu Leu Cys Gly Leu Leu Ala
1 5 10 15
Ala Thr Leu Ile Gln Ala Thr Leu Ser Pro Thr Ala Val Leu Ile
20 25 30
Leu Gly Pro Lys Val Ile Lys Glu Lys Leu Thr Gln Glu Leu Lys
35 40 45
Asp His Asn Ala Thr Ser Ile Leu Gln Gln Leu Pro Leu Leu Ser
50 55 60
Ala Met Arg Glu Lys Pro Ala Gly Gly Ile Pro Val Leu Gly Ser

-continued

65					70					75				
Leu	Val	Asn	Thr	Val	Leu	Lys	His	Ile	Ile	Trp	Leu	Lys	Val	Ile
				80					85					90
Thr	Ala	Asn	Ile	Leu	Gln	Leu	Gln	Val	Lys	Pro	Ser	Ala	Asn	Asp
				95					100					105
Gln	Glu	Leu	Leu	Val	Lys	Ile	Pro	Leu	Asp	Met	Val	Ala	Gly	Phe
				110					115					120
Asn	Thr	Pro	Leu	Val	Lys	Thr	Ile	Val	Glu	Phe	His	Met	Thr	Thr
				125					130					135
Glu	Ala	Gln	Ala	Thr	Ile	Arg	Met	Asp	Thr	Ser	Ala	Ser	Gly	Pro
				140					145					150
Thr	Arg	Leu	Val	Leu	Ser	Asp	Cys	Ala	Thr	Ser	His	Gly	Ser	Leu
				155					160					165
Arg	Ile	Gln	Leu	Leu	Tyr	Lys	Leu	Ser	Phe	Leu	Val	Asn	Ala	Leu
				170					175					180
Ala	Lys	Gln	Val	Met	Asn	Leu	Leu	Val	Pro	Ser	Leu	Pro	Asn	Leu
				185					190					195
Val	Lys	Asn	Gln	Leu	Cys	Pro	Val	Ile	Glu	Ala	Ser	Phe	Asn	Gly
				200					205					210
Met	Tyr	Ala	Asp	Leu	Leu	Gln	Leu	Val	Lys	Val	Pro	Ile	Ser	Leu
				215					220					225
Ser	Ile	Asp	Arg	Leu	Glu	Phe	Asp	Leu	Leu	Tyr	Pro	Ala	Ile	Lys
				230					235					240
Gly	Asp	Thr	Ile	Gln	Leu	Tyr	Leu	Gly	Ala	Lys	Leu	Leu	Asp	Ser
				245					250					255
Gln	Gly	Lys	Val	Thr	Lys	Trp	Phe	Asn	Asn	Ser	Ala	Ala	Ser	Leu
				260					265					270
Thr	Met	Pro	Thr	Leu	Asp	Asn	Ile	Pro	Phe	Ser	Leu	Ile	Val	Ser
				275					280					285
Gln	Asp	Val	Val	Lys	Ala	Ala	Val	Ala	Ala	Val	Leu	Ser	Pro	Glu
				290					295					300
Glu	Phe	Met	Val	Leu	Leu	Asp	Ser	Val	Leu	Pro	Glu	Ser	Ala	His
				305					310					315
Arg	Leu	Lys	Ser	Ser	Ile	Gly	Leu	Ile	Asn	Glu	Lys	Ala	Ala	Asp
				320					325					330
Lys	Leu	Gly	Ser	Thr	Gln	Ile	Val	Lys	Ile	Leu	Thr	Gln	Asp	Thr
				335					340					345
Pro	Glu	Phe	Phe	Ile	Asp	Gln	Gly	His	Ala	Lys	Val	Ala	Gln	Leu
				350					355					360
Ile	Val	Leu	Glu	Val	Phe	Pro	Ser	Ser	Glu	Ala	Leu	Arg	Pro	Leu
				365					370					375
Phe	Thr	Leu	Gly	Ile	Glu	Ala	Ser	Ser	Glu	Ala	Gln	Phe	Tyr	Thr
				380					385					390
Lys	Gly	Asp	Gln	Leu	Ile	Leu	Asn	Leu	Asn	Asn	Ile	Ser	Ser	Asp
				395					400					405
Arg	Ile	Gln	Leu	Met	Asn	Ser	Gly	Ile	Gly	Trp	Phe	Gln	Pro	Asp
				410					415					420
Val	Leu	Lys	Asn	Ile	Ile	Thr	Glu	Ile	Ile	His	Ser	Ile	Leu	Leu
				425					430					435
Pro	Asn	Gln	Asn	Gly	Lys	Leu	Arg	Ser	Gly	Val	Pro	Val	Ser	Leu
				440					445					450

-continued

Val Lys Ala Leu Gly Phe Glu Ala Ala Glu Ser Ser Leu Thr Lys
455 460 465
Asp Ala Leu Val Leu Thr Pro Ala Ser Leu Trp Lys Pro Ser Ser
470 475 480
Pro Val Ser Gln

<210> SEQ ID NO 79
<211> LENGTH: 1475
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 79

gagagaagtc agcctggcag agagactctg aaatgaggga ttagaggtgt	50
tcaaggagca agagcttcag cctgaagaca agggagcagt ccctgaagac	100
gcttctactg agaggtctgc catggcctct cttggcctcc aacttgtggg	150
ctacatccta ggccttctgg ggcttttggg cactactggt gccatgctgc	200
tccccagctg gaaaacaagt tcttatgtcg gtgccagcat tgtgacagca	250
gttggtcttct ccaagggcct ctggatggaa tgtgccacac acagcacagg	300
catcacccag tgtgacatct atagcaccct tctgggcctg cccgctgaca	350
tccaggctgc ccaggccatg atggtgacat ccagtgaat ctccctccctg	400
gcctgcatta tctctgtggt gggcatgaga tgcacagtct tctgccagga	450
atcccagacc aaagacagag tggcggtagc aggtggagtc tttttcatcc	500
ttggaggcct cctgggattc attcctgttg cctggaatct tcatgggatc	550
ctacgggact tctactcacc actggtgcct gacagcatga aatttgagat	600
tggagaggct ctttacttgg gcattatttc ttccctgttc tcctgatag	650
ctggaatcat cctctgcttt tcctgctcat cccagagaaa tcgctccaac	700
tactacgatg cctaccaagc ccaacctctt gccacaagga gctotccaag	750
gcctgggtcaa cctcccaaag tcaagagtga gttcaattcc tacagcctga	800
cagggtatgt gtgaagaacc aggggccaga gctgggggggt ggctgggtct	850
gtgaaaaaca gtggacagca ccccagggc cagaggtgag ggacactacc	900
actggtatct gtcagaaggt gctgctgagg atagactgac tttggccatt	950
ggattgagca aaggcagaaa tgggggctag tgtaacagca tgcaggttga	1000
attgccaaag atgctcgcca tgccagcctt tctgttttcc tcaccttgct	1050
gtccccctgc cctaagtccc caacctcaa cttgaaaccc cattccctta	1100
agccaggact cagaggatcc ctttgccctc tggtttacct gggactccat	1150
ccccaaaccc actaatcaca tcccactgac tgaccctctg tgatcaaaga	1200
ccctctctct ggctgaggtt ggctcttagc tcattgctgg ggatgggaag	1250
gagaagcagt ggcttttgtg ggcattgctc taacctactt ctcaagcttc	1300
cctccaaaga aactgatttg ccctggaacc tocatccac tcttgttatg	1350
actccacagt gtccagacta atttgtgcat gaactgaaat aaaaccatcc	1400
tacggtatcc agggaacaga aagcaggatg caggatggga ggacaggaag	1450
gcagcctggg acatttaaaa aaata	1475

-continued

<210> SEQ ID NO 80
<211> LENGTH: 230
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 80

Met Ala Ser Leu Gly Leu Gln Leu Val Gly Tyr Ile Leu Gly Leu
1 5 10 15
Leu Gly Leu Leu Gly Thr Leu Val Ala Met Leu Leu Pro Ser Trp
20 25 30
Lys Thr Ser Ser Tyr Val Gly Ala Ser Ile Val Thr Ala Val Gly
35 40 45
Phe Ser Lys Gly Leu Trp Met Glu Cys Ala Thr His Ser Thr Gly
50 55 60
Ile Thr Gln Cys Asp Ile Tyr Ser Thr Leu Leu Gly Leu Pro Ala
65 70 75
Asp Ile Gln Ala Ala Gln Ala Met Met Val Thr Ser Ser Ala Ile
80 85 90
Ser Ser Leu Ala Cys Ile Ile Ser Val Val Gly Met Arg Cys Thr
95 100 105
Val Phe Cys Gln Glu Ser Arg Ala Lys Asp Arg Val Ala Val Ala
110 115 120
Gly Gly Val Phe Phe Ile Leu Gly Gly Leu Leu Gly Phe Ile Pro
125 130 135
Val Ala Trp Asn Leu His Gly Ile Leu Arg Asp Phe Tyr Ser Pro
140 145 150
Leu Val Pro Asp Ser Met Lys Phe Glu Ile Gly Glu Ala Leu Tyr
155 160 165
Leu Gly Ile Ile Ser Ser Leu Phe Ser Leu Ile Ala Gly Ile Ile
170 175 180
Leu Cys Phe Ser Cys Ser Ser Gln Arg Asn Arg Ser Asn Tyr Tyr
185 190 195
Asp Ala Tyr Gln Ala Gln Pro Leu Ala Thr Arg Ser Ser Pro Arg
200 205 210
Pro Gly Gln Pro Pro Lys Val Lys Ser Glu Phe Asn Ser Tyr Ser
215 220 225
Leu Thr Gly Tyr Val
230

<210> SEQ ID NO 81
<211> LENGTH: 1732
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 81

cccacgcgtc cgcgcctctc ccttctgctg gaccttcctt cgtctctcca 50
tctctccctc ctttccccgc gttctctttc cacctttctc ttcttccac 100
cttagacctc ccttctgtgc ctcttttctt gccaccgct gcttctgtgc 150
ccttctccga ccccgctcta gcagcagacc tcctggggtc tgtgggttga 200
tctgtggccc ctgtgcctcc gtgtcctttt cgtctccctt cctcccgact 250
ccgctcccg g accagcggcc tgacctggg gaaaggatgg ttcccagagt 300

-continued

```
gagggtcctc tcctccttgc tgggactcgc gctgctctgg ttccccctgg      350
actccccacgc tcgagcccgcc ccagacatgt tctgcctttt ccatgggaag      400
agatactccc ccggcgagag ctggcaccgc tacttgagag cacaaggcct      450
gatgtactgc ctgcgctgta cctgctcaga gggcgcccat gtgagtgttt      500
accgcctcca ctgtccgcct gtccactgcc ccagcctgtg gacggagcca      550
cagcaatgct gtcccaagtg tgtggaacct cactctcctt ctggactccg      600
ggccccacca aagtcctgcc agcacaaagg gaccatgtac caacacggag      650
agatcttcag tgcccatgag ctgttcccct ccgcctgcc caaccagtgt      700
gtcctctgca gctgcacaga ggccagatc tactgcggcc tcacaacctg      750
ccccgaacca ggctgccagc caccctccc actgccagac tcctgctgcc      800
aagcctgcaa agatgaggca agtgagcaat cggatgaaga ggacagtgtg      850
cagtcgctcc atggggtgag acatcctcag gatccatgtt ccagtgatgc      900
tgggagaaaa agaggcccg gacccccagc cccactggc ctcagcgccc      950
ctctgagctt catccctcgc cacttcagac ccaagggagc aggcagcaca     1000
actgtcaaga tcgtcctgaa ggagaacat aagaaagcct gtgtgcatgg     1050
cgggaagacg tactcccacg gggaggtgtg gcacccggcc ttccgtgcct     1100
tcggcccctt gccctgcac ctatgcacct gtgaggatgg ccgccaggac     1150
tgccagcgtg tgacctgtcc caccgagtac ccctgccgtc accccgagaa     1200
agtggtctgg aagtgtctga agatttgccc agaggacaaa gcagaccctg     1250
gccacagtga gatcagttct accaggtgtc ccaaggcacc gggccgggtc     1300
ctcgtccaca catcggtatc cccaagccca gacaacctgc gtcgctttgc     1350
cctggaacac gaggcctcgg acttggtgga gatctacctc tggaagctgg     1400
taaaagatga ggaactgag gctcagagag gtgaagtacc tggcccaagg     1450
ccacacagcc agaatcttcc acttgactca gatcaagaaa gtcaggaagc     1500
aagacttcca gaaagaggca cagcacttcc gactgctcgc tggcccccac     1550
gaaggtcact ggaacgtctt cctagcccag accctggagc tgaaggtcac     1600
ggccagtcca gacaaagtga ccaagacata acaaagacct aacagttgca     1650
gatatgagct gtataattgt tgttattata tattaataaa taagaagttg     1700
cattaccctc aaaaaaaaaa aaaaaaaaaa aa                        1732
```

<210> SEQ ID NO 82

<211> LENGTH: 451

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 82

```
Met Val Pro Glu Val Arg Val Leu Ser Ser Leu Leu Gly Leu Ala
  1             5             10            15
```

```
Leu Leu Trp Phe Pro Leu Asp Ser His Ala Arg Ala Arg Pro Asp
          20            25            30
```

```
Met Phe Cys Leu Phe His Gly Lys Arg Tyr Ser Pro Gly Glu Ser
          35            40            45
```

-continued

Trp	His	Pro	Tyr	Leu	Glu	Pro	Gln	Gly	Leu	Met	Tyr	Cys	Leu	Arg	50	55	60
Cys	Thr	Cys	Ser	Glu	Gly	Ala	His	Val	Ser	Cys	Tyr	Arg	Leu	His	65	70	75
Cys	Pro	Pro	Val	His	Cys	Pro	Gln	Pro	Val	Thr	Glu	Pro	Gln	Gln	80	85	90
Cys	Cys	Pro	Lys	Cys	Val	Glu	Pro	His	Thr	Pro	Ser	Gly	Leu	Arg	95	100	105
Ala	Pro	Pro	Lys	Ser	Cys	Gln	His	Asn	Gly	Thr	Met	Tyr	Gln	His	110	115	120
Gly	Glu	Ile	Phe	Ser	Ala	His	Glu	Leu	Phe	Pro	Ser	Arg	Leu	Pro	125	130	135
Asn	Gln	Cys	Val	Leu	Cys	Ser	Cys	Thr	Glu	Gly	Gln	Ile	Tyr	Cys	140	145	150
Gly	Leu	Thr	Thr	Cys	Pro	Glu	Pro	Gly	Cys	Pro	Ala	Pro	Leu	Pro	155	160	165
Leu	Pro	Asp	Ser	Cys	Cys	Gln	Ala	Cys	Lys	Asp	Glu	Ala	Ser	Glu	170	175	180
Gln	Ser	Asp	Glu	Glu	Asp	Ser	Val	Gln	Ser	Leu	His	Gly	Val	Arg	185	190	195
His	Pro	Gln	Asp	Pro	Cys	Ser	Ser	Asp	Ala	Gly	Arg	Lys	Arg	Gly	200	205	210
Pro	Gly	Thr	Pro	Ala	Pro	Thr	Gly	Leu	Ser	Ala	Pro	Leu	Ser	Phe	215	220	225
Ile	Pro	Arg	His	Phe	Arg	Pro	Lys	Gly	Ala	Gly	Ser	Thr	Thr	Val	230	235	240
Lys	Ile	Val	Leu	Lys	Glu	Lys	His	Lys	Lys	Ala	Cys	Val	His	Gly	245	250	255
Gly	Lys	Thr	Tyr	Ser	His	Gly	Glu	Val	Trp	His	Pro	Ala	Phe	Arg	260	265	270
Ala	Phe	Gly	Pro	Leu	Pro	Cys	Ile	Leu	Cys	Thr	Cys	Glu	Asp	Gly	275	280	285
Arg	Gln	Asp	Cys	Gln	Arg	Val	Thr	Cys	Pro	Thr	Glu	Tyr	Pro	Cys	290	295	300
Arg	His	Pro	Glu	Lys	Val	Ala	Gly	Lys	Cys	Cys	Lys	Ile	Cys	Pro	305	310	315
Glu	Asp	Lys	Ala	Asp	Pro	Gly	His	Ser	Glu	Ile	Ser	Ser	Thr	Arg	320	325	330
Cys	Pro	Lys	Ala	Pro	Gly	Arg	Val	Leu	Val	His	Thr	Ser	Val	Ser	335	340	345
Pro	Ser	Pro	Asp	Asn	Leu	Arg	Arg	Phe	Ala	Leu	Glu	His	Glu	Ala	350	355	360
Ser	Asp	Leu	Val	Glu	Ile	Tyr	Leu	Trp	Lys	Leu	Val	Lys	Asp	Glu	365	370	375
Glu	Thr	Glu	Ala	Gln	Arg	Gly	Glu	Val	Pro	Gly	Pro	Arg	Pro	His	380	385	390
Ser	Gln	Asn	Leu	Pro	Leu	Asp	Ser	Asp	Gln	Glu	Ser	Gln	Glu	Ala	395	400	405
Arg	Leu	Pro	Glu	Arg	Gly	Thr	Ala	Leu	Pro	Thr	Ala	Arg	Trp	Pro	410	415	420
Pro	Arg	Arg	Ser	Leu	Glu	Arg	Leu	Pro	Ser	Pro	Asp	Pro	Gly	Ala			

-continued

atagtcacatc gccagtcac ccaggaatca gagaaagagg cctcttgga	1600
aagggcctct gcaatcccag agacaagcaa cagtgaagtc tcctcacagg	1650
caaccacgcc cttctcccc aggggtgaaa tgtaggatga atcacatccc	1700
acattcttct ttagggatat taaggctctc ctcccagatc caaagtcacc	1750
cagcagccgg ccaaggtggc ttccagatga agggggactg gcctgtccac	1800
atgggagtcg ggtgtcatgg ctgccctgag ctgggaggga agaaggctga	1850
cattacattt agtttgctct cactccatct ggctaagtga tcttgaaata	1900
ccacctctca ggtgaagaac cgtcaggaat tcccatctca caggctgtgg	1950
tgtagattaa gtagacaagg aatgtgaata atgcttagat cttattgatg	2000
acagagtgtg tcctaattgt ttgttcatta tattacactt tcagtaaaaa	2050
aa	2052

<210> SEQ ID NO 84

<211> LENGTH: 500

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 84

Met Ala Leu Met Leu Ser Leu Val Leu Ser Leu Leu Lys Leu Gly	1	5	10	15
Ser Gly Gln Trp Gln Val Phe Gly Pro Asp Lys Pro Val Gln Ala	20	25	30	
Leu Val Gly Glu Asp Ala Ala Phe Ser Cys Phe Leu Ser Pro Lys	35	40	45	
Thr Asn Ala Glu Ala Met Glu Val Arg Phe Phe Arg Gly Gln Phe	50	55	60	
Ser Ser Val Val His Leu Tyr Arg Asp Gly Lys Asp Gln Pro Phe	65	70	75	
Met Gln Met Pro Gln Tyr Gln Gly Arg Thr Lys Leu Val Lys Asp	80	85	90	
Ser Ile Ala Glu Gly Arg Ile Ser Leu Arg Leu Glu Asn Ile Thr	95	100	105	
Val Leu Asp Ala Gly Leu Tyr Gly Cys Arg Ile Ser Ser Gln Ser	110	115	120	
Tyr Tyr Gln Lys Ala Ile Trp Glu Leu Gln Val Ser Ala Leu Gly	125	130	135	
Ser Val Pro Leu Ile Ser Ile Thr Gly Tyr Val Asp Arg Asp Ile	140	145	150	
Gln Leu Leu Cys Gln Ser Ser Gly Trp Phe Pro Arg Pro Thr Ala	155	160	165	
Lys Trp Lys Gly Pro Gln Gly Gln Asp Leu Ser Thr Asp Ser Arg	170	175	180	
Thr Asn Arg Asp Met His Gly Leu Phe Asp Val Glu Ile Ser Leu	185	190	195	
Thr Val Gln Glu Asn Ala Gly Ser Ile Ser Cys Ser Met Arg His	200	205	210	
Ala His Leu Ser Arg Glu Val Glu Ser Arg Val Gln Ile Gly Asp	215	220	225	

-continued

Thr	Phe	Phe	Glu	Pro	Ile	Ser	Trp	His	Leu	Ala	Thr	Lys	Val	Leu	
			230						235					240	
Gly	Ile	Leu	Cys	Cys	Gly	Leu	Phe	Phe	Gly	Ile	Val	Gly	Leu	Lys	
			245						250					255	
Ile	Phe	Phe	Ser	Lys	Phe	Gln	Trp	Lys	Ile	Gln	Ala	Glu	Leu	Asp	
			260						265					270	
Trp	Arg	Arg	Lys	His	Gly	Gln	Ala	Glu	Leu	Arg	Asp	Ala	Arg	Lys	
			275						280					285	
His	Ala	Val	Glu	Val	Thr	Leu	Asp	Pro	Glu	Thr	Ala	His	Pro	Lys	
			290						295					300	
Leu	Cys	Val	Ser	Asp	Leu	Lys	Thr	Val	Thr	His	Arg	Lys	Ala	Pro	
			305						310					315	
Gln	Glu	Val	Pro	His	Ser	Glu	Lys	Arg	Phe	Thr	Arg	Lys	Ser	Val	
			320						325					330	
Val	Ala	Ser	Gln	Ser	Phe	Gln	Ala	Gly	Lys	His	Tyr	Trp	Glu	Val	
			335						340					345	
Asp	Gly	Gly	His	Asn	Lys	Arg	Trp	Arg	Val	Gly	Val	Cys	Arg	Asp	
			350						355					360	
Asp	Val	Asp	Arg	Arg	Lys	Glu	Tyr	Val	Thr	Leu	Ser	Pro	Asp	His	
			365						370					375	
Gly	Tyr	Trp	Val	Leu	Arg	Leu	Asn	Gly	Glu	His	Leu	Tyr	Phe	Thr	
			380						385					390	
Leu	Asn	Pro	Arg	Phe	Ile	Ser	Val	Phe	Pro	Arg	Thr	Pro	Pro	Thr	
			395						400					405	
Lys	Ile	Gly	Val	Phe	Leu	Asp	Tyr	Glu	Cys	Gly	Thr	Ile	Ser	Phe	
			410						415					420	
Phe	Asn	Ile	Asn	Asp	Gln	Ser	Leu	Ile	Tyr	Thr	Leu	Thr	Cys	Arg	
			425						430					435	
Phe	Glu	Gly	Leu	Leu	Arg	Pro	Tyr	Ile	Glu	Tyr	Pro	Ser	Tyr	Asn	
			440						445					450	
Glu	Gln	Asn	Gly	Thr	Pro	Ile	Val	Ile	Cys	Pro	Val	Thr	Gln	Glu	
			455						460					465	
Ser	Glu	Lys	Glu	Ala	Ser	Trp	Gln	Arg	Ala	Ser	Ala	Ile	Pro	Glu	
			470						475					480	
Thr	Ser	Asn	Ser	Glu	Ser	Ser	Ser	Gln	Ala	Thr	Thr	Pro	Phe	Leu	
			485						490					495	
Pro	Arg	Gly	Glu	Met											
			500												

<210> SEQ ID NO 85
<211> LENGTH: 1665
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 85

aacagacgtt ccctcgcggc cctggcacct ctaaccccag acatgctgct	50
gctgctgctg cccctgctct gggggaggga gagggcggaa ggacagacaa	100
gtaaactgct gacgatgcag agttccgtga cggtcagga aggcctgtgt	150
gtccatgtgc cctgctcctt ctctacccc tcgcatggct ggatttacc	200
tggcccagta gttcatggct actggttccg ggaaggggcc aatacagacc	250
aggatgctcc agtggccaca aacaaccag ctcgggcagt gtgggaggag	300

-continued

```

actcgggacc gattccacct ccttggggac ccacatacca agaattgcac      350
cctgagcatc agagatgcc aagaagtga tgcggggaga tacttctttc      400
gtatggagaa aggaagtata aaatggaatt ataacatca ccggctctct      450
gtgaatgtga cagccttgac ccacaggccc aacatcctca tcccaggcac      500
cctggagtcg ggctgcccc agaatctgac ctgctctgtg ccctgggcct      550
gtgagcaggg gacacccct atgatctcct ggataggac ctcctgtgtc      600
ccccggacc cctccaccac ccgctcctcg gtgctcacc tcatcccaca      650
gccccaggac catggcacca gcctcacctg tcaggtgacc ttccctgggg      700
ccagcgtgac cacgaacaag accgtccatc tcaacgtgtc ctaccgcct      750
cagaacttga ccatgactgt cttccaagga gacggcacag tatccacagt      800
cttgggaaat ggctcatctc tgtcactccc agagggccag tctctgcgcc      850
tggtctgtgc agttgatgca gttgacagca atccccctgc caggctgagc      900
ctgagctgga gaggcctgac cctgtgcccc tcacagccct caaacccggg      950
ggtgctggag ctgccttggg tgcacctgag ggatgcagct gaattcacct     1000
gcagagctca gaaccctctc ggctctcagc aggtctacct gaacgtctcc     1050
ctgcagagca aagccacatc aggagtgact cagggggtgg tcgggggagc     1100
tggagccaca gccctggtct tcctgtcctt ctgcgtcatc ttcgtttag      1150
tgaggtcctg caggaagaaa tcggcaaggc cagcagcggg cgtgggagat     1200
acgggcatag aggatgcaaa cgctgtcagg ggttcagcct ctcaggggcc     1250
cctgactgaa ccttgggcag aagacagtcc ccagaccag cctccccag      1300
cttctgcccg ctctcagtg ggggaaggag agtccagta tgcattccctc     1350
agcttccaga tggatgaagc ttgggactcg cggggacagg aggccactga     1400
caccgagtac tcggagatca agatccacag atgagaaact gcagagactc     1450
accctgattg agggatcaca gcccctccag gcaagggaga agtcagaggc     1500
tgattcttgt agaattaaca gccctcaacg tgatgagcta tgataaact      1550
atgaattatg tgcagagtga aaagcacaca ggctttagag tcaaagtatc     1600
tcaaacctga atccacactg tgccctccct tttatTTTTT taactaaaag     1650
acagacaaat tccta                                             1665

```

<210> SEQ ID NO 86

<211> LENGTH: 463

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 86

```

Met Leu Leu Leu Leu Leu Pro Leu Leu Trp Gly Arg Glu Arg Ala
 1             5             10            15
Glu Gly Gln Thr Ser Lys Leu Leu Thr Met Gln Ser Ser Val Thr
                20            25            30
Val Gln Glu Gly Leu Cys Val His Val Pro Cys Ser Phe Ser Tyr
                35            40            45
Pro Ser His Gly Trp Ile Tyr Pro Gly Pro Val Val His Gly Tyr
                50            55            60

```

-continued

Trp Phe Arg Glu Gly Ala Asn Thr Asp Gln Asp Ala Pro Val Ala	65	70	75
Thr Asn Asn Pro Ala Arg Ala Val Trp Glu Glu Thr Arg Asp Arg	80	85	90
Phe His Leu Leu Gly Asp Pro His Thr Lys Asn Cys Thr Leu Ser	95	100	105
Ile Arg Asp Ala Arg Arg Ser Asp Ala Gly Arg Tyr Phe Phe Arg	110	115	120
Met Glu Lys Gly Ser Ile Lys Trp Asn Tyr Lys His His Arg Leu	125	130	135
Ser Val Asn Val Thr Ala Leu Thr His Arg Pro Asn Ile Leu Ile	140	145	150
Pro Gly Thr Leu Glu Ser Gly Cys Pro Gln Asn Leu Thr Cys Ser	155	160	165
Val Pro Trp Ala Cys Glu Gln Gly Thr Pro Pro Met Ile Ser Trp	170	175	180
Ile Gly Thr Ser Val Ser Pro Leu Asp Pro Ser Thr Thr Arg Ser	185	190	195
Ser Val Leu Thr Leu Ile Pro Gln Pro Gln Asp His Gly Thr Ser	200	205	210
Leu Thr Cys Gln Val Thr Phe Pro Gly Ala Ser Val Thr Thr Asn	215	220	225
Lys Thr Val His Leu Asn Val Ser Tyr Pro Pro Gln Asn Leu Thr	230	235	240
Met Thr Val Phe Gln Gly Asp Gly Thr Val Ser Thr Val Leu Gly	245	250	255
Asn Gly Ser Ser Leu Ser Leu Pro Glu Gly Gln Ser Leu Arg Leu	260	265	270
Val Cys Ala Val Asp Ala Val Asp Ser Asn Pro Pro Ala Arg Leu	275	280	285
Ser Leu Ser Trp Arg Gly Leu Thr Leu Cys Pro Ser Gln Pro Ser	290	295	300
Asn Pro Gly Val Leu Glu Leu Pro Trp Val His Leu Arg Asp Ala	305	310	315
Ala Glu Phe Thr Cys Arg Ala Gln Asn Pro Leu Gly Ser Gln Gln	320	325	330
Val Tyr Leu Asn Val Ser Leu Gln Ser Lys Ala Thr Ser Gly Val	335	340	345
Thr Gln Gly Val Val Gly Gly Ala Gly Ala Thr Ala Leu Val Phe	350	355	360
Leu Ser Phe Cys Val Ile Phe Val Val Val Arg Ser Cys Arg Lys	365	370	375
Lys Ser Ala Arg Pro Ala Ala Gly Val Gly Asp Thr Gly Ile Glu	380	385	390
Asp Ala Asn Ala Val Arg Gly Ser Ala Ser Gln Gly Pro Leu Thr	395	400	405
Glu Pro Trp Ala Glu Asp Ser Pro Pro Asp Gln Pro Pro Pro Ala	410	415	420
Ser Ala Arg Ser Ser Val Gly Glu Gly Glu Leu Gln Tyr Ala Ser	425	430	435

-continued

Leu Ser Phe Gln Met Val Lys Pro Trp Asp Ser Arg Gly Gln Glu
440 445 450

Ala Thr Asp Thr Glu Tyr Ser Glu Ile Lys Ile His Arg
455 460

<210> SEQ ID NO 87

<211> LENGTH: 1176

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 87

agaaagctgc actctgttga gctccagggc gcagtgagg gagggagtga 50
aggagctctc tgtaccaag gaaagtgcag ctgagactca gacaagatta 100
caatgaacca actcagcttc ctgctgtttc tcatagcgac caccagagga 150
tggagtacag atgaggctaa tacttacttc aaggaatgga cctgttcctc 200
gtctccatct ctgcccagaa gctgcaagga aatcaaagac gaatgtccta 250
gtgcatttga tggcctgtat ttctccgca ctgagaatgg tgttatctac 300
cagaccttct gtgacatgac ctctgggggt ggcggctgga ccctggtggc 350
cagcgtgcat gagaatgaca tgcgtgggaa gtgcacggtg ggcgatcgct 400
gggccagcta gcagggcagc aaagcagact acccagaggg ggacggcaac 450
tgggccaact acaacacctt tggatctgca gaggcggcca cgagcgatga 500
ctacaagaac cctggctact acgacatcca ggccaaggac ctgggcatct 550
ggcagctgoc caataagtcc cccatgcagc actggagaaa cagtcctctg 600
ctgaggtacc gcacggacac tggcttcctc cagacactgg gacataatct 650
gtttggcatc taccagaaat atccagtga atatggagaa ggaagtgtt 700
ggactgacaa cggcccgggtg atccctgtgg totatgattt tggcgacgcc 750
cagaaaacag catcttatta ctacacctat ggccagcggg aattcactgc 800
gggatttggt cagttcaggg tatttaataa cgagagagca gccaacgcct 850
tgtgtgctgg aatgagggtc accggatgta aactgagca tcaactgcatt 900
ggtggaggag gatactttcc agaggccagt cccagcagt gtggagattt 950
ttctggtttt gattggagtg gatatggaac tcatgttgggt tacagcagca 1000
gccgtgagat aactgaggca gctgtgcttc tattctatcg ttgagagttt 1050
tgtgggaggg aaccagacc tctcctccca accatgagat cccaaggatg 1100
gagaacaact taccagtag ctagaatgtt aatggcagaa gagaaaacaa 1150
taaatcatat tgactcaaga aaaaaa 1176

<210> SEQ ID NO 88

<211> LENGTH: 313

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 88

Met Asn Gln Leu Ser Phe Leu Leu Phe Leu Ile Ala Thr Thr Arg
1 5 10 15

Gly Trp Ser Thr Asp Glu Ala Asn Thr Tyr Phe Lys Glu Trp Thr
20 25 30

-continued

Cys	Ser	Ser	Ser	Pro	Ser	Leu	Pro	Arg	Ser	Cys	Lys	Glu	Ile	Lys
				35					40					45
Asp	Glu	Cys	Pro	Ser	Ala	Phe	Asp	Gly	Leu	Tyr	Phe	Leu	Arg	Thr
				50					55					60
Glu	Asn	Gly	Val	Ile	Tyr	Gln	Thr	Phe	Cys	Asp	Met	Thr	Ser	Gly
				65					70					75
Gly	Gly	Gly	Trp	Thr	Leu	Val	Ala	Ser	Val	His	Glu	Asn	Asp	Met
				80					85					90
Arg	Gly	Lys	Cys	Thr	Val	Gly	Asp	Arg	Trp	Ser	Ser	Gln	Gln	Gly
				95					100					105
Ser	Lys	Ala	Asp	Tyr	Pro	Glu	Gly	Asp	Gly	Asn	Trp	Ala	Asn	Tyr
				110					115					120
Asn	Thr	Phe	Gly	Ser	Ala	Glu	Ala	Ala	Thr	Ser	Asp	Asp	Tyr	Lys
				125					130					135
Asn	Pro	Gly	Tyr	Tyr	Asp	Ile	Gln	Ala	Lys	Asp	Leu	Gly	Ile	Trp
				140					145					150
His	Val	Pro	Asn	Lys	Ser	Pro	Met	Gln	His	Trp	Arg	Asn	Ser	Ser
				155					160					165
Leu	Leu	Arg	Tyr	Arg	Thr	Asp	Thr	Gly	Phe	Leu	Gln	Thr	Leu	Gly
				170					175					180
His	Asn	Leu	Phe	Gly	Ile	Tyr	Gln	Lys	Tyr	Pro	Val	Lys	Tyr	Gly
				185					190					195
Glu	Gly	Lys	Cys	Trp	Thr	Asp	Asn	Gly	Pro	Val	Ile	Pro	Val	Val
				200					205					210
Tyr	Asp	Phe	Gly	Asp	Ala	Gln	Lys	Thr	Ala	Ser	Tyr	Tyr	Ser	Pro
				215					220					225
Tyr	Gly	Gln	Arg	Glu	Phe	Thr	Ala	Gly	Phe	Val	Gln	Phe	Arg	Val
				230					235					240
Phe	Asn	Asn	Glu	Arg	Ala	Ala	Asn	Ala	Leu	Cys	Ala	Gly	Met	Arg
				245					250					255
Val	Thr	Gly	Cys	Asn	Thr	Glu	His	His	Cys	Ile	Gly	Gly	Gly	Gly
				260					265					270
Tyr	Phe	Pro	Glu	Ala	Ser	Pro	Gln	Gln	Cys	Gly	Asp	Phe	Ser	Gly
				275					280					285
Phe	Asp	Trp	Ser	Gly	Tyr	Gly	Thr	His	Val	Gly	Tyr	Ser	Ser	Ser
				290					295					300
Arg	Glu	Ile	Thr	Glu	Ala	Ala	Val	Leu	Leu	Phe	Tyr	Arg		
				305					310					

<210> SEQ ID NO 89
<211> LENGTH: 759
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 89

ctagatttgt	cggcttg	cg	ggagacttca	ggagtcgctg	tctctgaact	50
tccagcctca	gagaccgccg	cccttg	cccc	cgagggccat	gggccgggtc	100
tcagggcttg	tgcctctcg	cttc	ctgacg	ctcctggcgc	atctggtggt	150
cgctcatcacc	ttattctggt	ccggg	acag	caacatacag	gcctgcctgc	200
ctctcacggt	cacccccgag	gagtatgaca	agcaggacat	tcagctggtg		250
gccgcgctct	ctgtcaccct	gggcctcttt	gcagtgagc	tggccggttt		300

-continued

cctctcagga gtctccatgt tcaacagcac ccagagcctc atctccattg	350
gggctcactg tagtgcatcc gtggccctgt ccttcttcat attcgagcgt	400
tgggagtgca ctacgtattg gtacattttt gtcttctgca gtgcccttcc	450
agctgtcact gaaatggctt tattcgtcac cgtctttggg ctgaaaaaga	500
aacccttctg attaccttca tgacgggaac ctaaggacga agcctacagg	550
ggcaagggcc gcttcgtatt cctggaagaa ggaaggcata ggcttcggtt	600
ttcccctcgg aaactgcttc tgctggagga tatgtgttgg aataattacg	650
tcttgagtct gggattatcc gcattgtatt tagtgctttg taataaaata	700
tgttttgtag taacattaag acttatatac agttttaggg gacaattaa	750
aaaaaaaa	759

<210> SEQ ID NO 90
<211> LENGTH: 140
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 90

Met Gly Arg Val Ser Gly Leu Val Pro Ser Arg Phe Leu Thr Leu	1	5	10	15
Leu Ala His Leu Val Val Val Ile Thr Leu Phe Trp Ser Arg Asp	20	25	30	
Ser Asn Ile Gln Ala Cys Leu Pro Leu Thr Phe Thr Pro Glu Glu	35	40	45	
Tyr Asp Lys Gln Asp Ile Gln Leu Val Ala Ala Leu Ser Val Thr	50	55	60	
Leu Gly Leu Phe Ala Val Glu Leu Ala Gly Phe Leu Ser Gly Val	65	70	75	
Ser Met Phe Asn Ser Thr Gln Ser Leu Ile Ser Ile Gly Ala His	80	85	90	
Cys Ser Ala Ser Val Ala Leu Ser Phe Phe Ile Phe Glu Arg Trp	95	100	105	
Glu Cys Thr Thr Tyr Trp Tyr Ile Phe Val Phe Cys Ser Ala Leu	110	115	120	
Pro Ala Val Thr Glu Met Ala Leu Phe Val Thr Val Phe Gly Leu	125	130	135	
Lys Lys Lys Pro Phe	140			

<210> SEQ ID NO 91
<211> LENGTH: 1871
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 91

ctgggacccc gaaaagagaa ggggagagcg aggggacgag agcggaggag	50
gaagatgcaa ctgactcgct gctgcttcgt gttcctggtg cagggtagcc	100
tctatctggt catctgtggc caggatgatg gtcctcccg ctcagaggac	150
cctgagcgtg atgaccacga gggccagccc cgccccggg tgcctcgga	200
gcggggccac atctcaccta agtcccggcc catggccaat tccactctcc	250

-continued

tagggctgct ggccccgcct ggggaggctt ggggcattct tgggcagccc	300
cccaaccgcc cgaaccacag cccccacccc tcagccaagg tgaagaaaat	350
ctttggctgg ggcgacttct actccaacat caagacggtg gccctgaacc	400
tgctcgtcac agggaagatt gtggaccatg gcaatgggac cttcagcgtc	450
cacttccaac acaatgccac aggccaggga aacatctcca tcagcctcgt	500
gccccccagt aaagctgtag agttccacca ggaacagcag atcttcatcg	550
aagccaaggc ctccaaaatc ttcaactgcc ggatggagtg ggagaaggta	600
gaacgggggc gccggacctc gctttgcacc cacgacccag ccaagatctg	650
ctcccagagc cagcctcaga gctcagccac ctggagctgc tcccagccct	700
tcaaagtctg ctgtgtctac atcgcccttct acagcacgga ctatcggctg	750
gtccagaagg tgtgcccaga ttacaactac catagtgata cccctacta	800
cccatctggy tgacccgggg caggccacag aggccaggcc agggctggaa	850
ggacaggcct gcccatgcag gagaccatct ggacaccggg cagggaaggg	900
gttgggcctc aggcagggag gggggtggag acgaggagat gccaagtggg	950
gccaggggcca agtctcaagt ggcagagaaa ggggcccaag tgctgggtccc	1000
aacctgaagc tgtggagtga ctagatcaca ggagcactgg aggaggagtg	1050
ggctctctgt gcagcctcac agggctttgc cacggagcca cagagagatg	1100
ctgggtcccc gaggcctgtg ggcaggccga tcagtgtggc cccagatcaa	1150
gtcatgggag gaagctaagc ccttggttct tgccatcctg aggaaagata	1200
gcaacaggga gggggagatt tcatcagtgt ggacagcctg tcaacttagg	1250
atggatggct gagagggctt cctagagacc agtcagcagg gtggggtggg	1300
gccagaggag ctctccagcc ctgcctagtg ggcgccctga gcccttgtc	1350
gtgtgctgag catggcatga ggctgaagtg gcaaccctgg ggtctttgat	1400
gtcttgacag attgaccatc tgtctccagc caggccaccc ctttccaaaa	1450
ttccctcttc tgccagtact cccctgtac caccattgc tgatggcaca	1500
cccatcctta agctaagaca ggacattgt ggtcctccca cactaaggcc	1550
acagcccac cgcgtgctgt gtgtccctct tccaccccaa cccctgctgy	1600
ctcctctggg agcatccatg tcccggagag gggccctca acagtcagcc	1650
tcacctgtca gaccggggtt ctcccggatc tggatggcgc cgccctctca	1700
gcagcgggca cgggtggggc ggggcccggc cgcagagcat gtgctggatc	1750
tgttctgtgt gtctgtctgt ggggtggggg aggggaggga agtcttgtga	1800
aaccgctgat tgctgacttt tgtgtgaaga atcgtgttct tggagcagga	1850
aataaagctt gccccggggc a	1871

<210> SEQ ID NO 92
<211> LENGTH: 252
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 92
Met Gln Leu Thr Arg Cys Cys Phe Val Phe Leu Val Gln Gly Ser

-continued

1	5	10	15
Leu Tyr Leu Val	Ile Cys Gly Gln Asp	Asp Gly Pro Pro Gly	Ser
	20	25	30
Glu Asp Pro Glu	Arg Asp Asp His Glu	Gly Gln Pro Arg Pro	Arg
	35	40	45
Val Pro Arg Lys	Arg Gly His Ile Ser	Pro Lys Ser Arg Pro	Met
	50	55	60
Ala Asn Ser Thr	Leu Leu Gly Leu Leu	Ala Pro Pro Gly Glu	Ala
	65	70	75
Trp Gly Ile Leu	Gly Gln Pro Pro Asn	Arg Pro Asn His Ser	Pro
	80	85	90
Pro Pro Ser Ala	Lys Val Lys Lys Ile	Phe Gly Trp Gly Asp	Phe
	95	100	105
Tyr Ser Asn Ile	Lys Thr Val Ala Leu	Asn Leu Leu Val Thr	Gly
	110	115	120
Lys Ile Val Asp	His Gly Asn Gly Thr	Phe Ser Val His Phe	Gln
	125	130	135
His Asn Ala Thr	Gly Gln Gly Asn Ile	Ser Ile Ser Leu Val	Pro
	140	145	150
Pro Ser Lys Ala	Val Glu Phe His Gln	Glu Gln Gln Ile Phe	Ile
	155	160	165
Glu Ala Lys Ala	Ser Lys Ile Phe Asn	Cys Arg Met Glu Trp	Glu
	170	175	180
Lys Val Glu Arg	Gly Arg Arg Thr Ser	Leu Cys Thr His Asp	Pro
	185	190	195
Ala Lys Ile Cys	Ser Arg Asp His Ala	Gln Ser Ser Ala Thr	Trp
	200	205	210
Ser Cys Ser Gln	Pro Phe Lys Val Val	Cys Val Tyr Ile Ala	Phe
	215	220	225
Tyr Ser Thr Asp	Tyr Arg Leu Val Gln	Lys Val Cys Pro Asp	Tyr
	230	235	240
Asn Tyr His Ser	Asp Thr Pro Tyr Tyr	Pro Ser Gly	
	245	250	

<210> SEQ ID NO 93
<211> LENGTH: 902
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 93

cggtggccat gactgcggcc gtgttcttcg gctgcgcctt cattgccttc	50
gggcctgcgc tcgcccttta tgtcttcacc atcgccatcg agccgttgcg	100
tatcatcttc ctcatcgccg gagctttctt ctggttggtg tctctactga	150
tttcgtccct tgtttggttc atggcaagag tcattattga caacaaagat	200
ggaccaaacac agaaatatct gctgatcttt ggagcgtttg tctctgtcta	250
tatccaagaa atgttcgat ttgcatatta taaactctta aaaaaagcca	300
gtgaaggttt gaagagtata aaccaggtg agacagcacc ctctatgcga	350
ctgtggcct atgtttctgg cttgggcttt ggaatcatga gtggagtatt	400
ttcctttgtg aataccctat ctgactcctt ggggccaggc acagtgggca	450

-continued

ttcatggaga ttctcctcaa ttcttccttt attcagcttt catgacgctg	500
gtcattatct tgctgcatgt attctggggc attgtatttt ttgatggctg	550
tgagaagaaa aagtggggca tcctccttat cgttctcctg acccacctgc	600
tggtgtcagc ccagaccttc ataagttctt attatggaat aaacctggcg	650
tcagcattta taatcctggt gctcatgggc acctgggcat tcttagctgc	700
gggaggcagc tgccgaagcc tgaaactctg cctgctctgc caagacaaga	750
actttcttct ttacaaccag cgctccagat aacctcaggg aaccagcact	800
tcccaaaccg cagactacat ctttagagga agcacaactg tgcctttttc	850
tgaaaaatccc tttttctggt ggaattgaga aagaaataaa actatgcaga	900
ta	902

<210> SEQ ID NO 94
<211> LENGTH: 257
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 94

Met Thr Ala Ala Val Phe Phe Gly Cys Ala Phe Ile Ala Phe Gly	1	5	10	15
Pro Ala Leu Ala Leu Tyr Val Phe Thr Ile Ala Ile Glu Pro Leu	20	25	30	
Arg Ile Ile Phe Leu Ile Ala Gly Ala Phe Phe Trp Leu Val Ser	35	40	45	
Leu Leu Ile Ser Ser Leu Val Trp Phe Met Ala Arg Val Ile Ile	50	55	60	
Asp Asn Lys Asp Gly Pro Thr Gln Lys Tyr Leu Leu Ile Phe Gly	65	70	75	
Ala Phe Val Ser Val Tyr Ile Gln Glu Met Phe Arg Phe Ala Tyr	80	85	90	
Tyr Lys Leu Leu Lys Lys Ala Ser Glu Gly Leu Lys Ser Ile Asn	95	100	105	
Pro Gly Glu Thr Ala Pro Ser Met Arg Leu Leu Ala Tyr Val Ser	110	115	120	
Gly Leu Gly Phe Gly Ile Met Ser Gly Val Phe Ser Phe Val Asn	125	130	135	
Thr Leu Ser Asp Ser Leu Gly Pro Gly Thr Val Gly Ile His Gly	140	145	150	
Asp Ser Pro Gln Phe Phe Leu Tyr Ser Ala Phe Met Thr Leu Val	155	160	165	
Ile Ile Leu Leu His Val Phe Trp Gly Ile Val Phe Phe Asp Gly	170	175	180	
Cys Glu Lys Lys Lys Trp Gly Ile Leu Leu Ile Val Leu Leu Thr	185	190	195	
His Leu Leu Val Ser Ala Gln Thr Phe Ile Ser Ser Tyr Tyr Gly	200	205	210	
Ile Asn Leu Ala Ser Ala Phe Ile Ile Leu Val Leu Met Gly Thr	215	220	225	
Trp Ala Phe Leu Ala Ala Gly Gly Ser Cys Arg Ser Leu Lys Leu	230	235	240	

-continued

Cys Leu Leu Cys Gln Asp Lys Asn Phe Leu Leu Tyr Asn Gln Arg
245 250 255

Ser Arg

<210> SEQ ID NO 95
<211> LENGTH: 1073
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 95

aatttttcac cagagtaaacc ttgagaaacc aactggacct tgagtattgt 50
acatttttgcc tcgtggaccac aaaggttagca atctgaaaca tgaggagtac 100
gattctactg ttttgtcttc taggatcaaac tcggtcatta ccacagctca 150
aacctgcttt gggactccct cccacaaaac tggctccgga tcagggaaaca 200
ctaccaaaac aacagcagtc aaatcaggtc tttccttctt taagtctgat 250
accattaaca cagatgctca cactggggcc agatctgcat ctgttaaatc 300
ctgctgcagg aatgacacct ggtaccacaga cccaccatt gaccctggga 350
gggttgaatg tacaacagca actgcaccca catgtgttac caatttttgt 400
cacacaactt ggagcccagg gcactatcct aagctcagag gaattgccac 450
aaatcttcac gagcctcatc atccattcct tgttcccggg aggcatactg 500
cccaccagtc aggcaggggc taatccagat gtccaggatg gaagccttcc 550
agcaggagga gcaggtgtaa atcctgccac ccagggaacc ccagcaggcc 600
gcctcccac tcccagtggc acagatgacg actttgcagt gaccaccct 650
gcaggcatcc aaaggagcac acatgccatc gaggaagcca ccacagaatc 700
agcaaatgga attcagtaag ctgtttcaaa ttttttcaac taagctgcct 750
cgaatttggt gatacatgtg aatctttatc attgattata ttatggaata 800
gattgagaca cattggatag tcttagaaga aattaattct taatttacct 850
gaaaatattc ttgaaatttc agaaaatag ttctatgtag agaatcccaa 900
cttttaaaaa caataattca atggataaat ctgtctttga aatataacat 950
tatgctgcct ggatgatatg catattaaaa catatttggg aaactggaaa 1000
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1050
aaaaaaaaaa aaaaaaaaaa aaa 1073

<210> SEQ ID NO 96
<211> LENGTH: 209
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 96

Met Arg Ser Thr Ile Leu Leu Phe Cys Leu Leu Gly Ser Thr Arg
1 5 10 15
Ser Leu Pro Gln Leu Lys Pro Ala Leu Gly Leu Pro Pro Thr Lys
20 25 30
Leu Ala Pro Asp Gln Gly Thr Leu Pro Asn Gln Gln Gln Ser Asn
35 40 45
Gln Val Phe Pro Ser Leu Ser Leu Ile Pro Leu Thr Gln Met Leu
50 55 60

-continued

Thr	Leu	Gly	Pro	Asp	Leu	His	Leu	Leu	Asn	Pro	Ala	Ala	Gly	Met
			65						70					75
Thr	Pro	Gly	Thr	Gln	Thr	His	Pro	Leu	Thr	Leu	Gly	Gly	Leu	Asn
			80						85					90
Val	Gln	Gln	Gln	Leu	His	Pro	His	Val	Leu	Pro	Ile	Phe	Val	Thr
			95						100					105
Gln	Leu	Gly	Ala	Gln	Gly	Thr	Ile	Leu	Ser	Ser	Glu	Glu	Leu	Pro
			110						115					120
Gln	Ile	Phe	Thr	Ser	Leu	Ile	Ile	His	Ser	Leu	Phe	Pro	Gly	Gly
			125						130					135
Ile	Leu	Pro	Thr	Ser	Gln	Ala	Gly	Ala	Asn	Pro	Asp	Val	Gln	Asp
			140						145					150
Gly	Ser	Leu	Pro	Ala	Gly	Gly	Ala	Gly	Val	Asn	Pro	Ala	Thr	Gln
			155						160					165
Gly	Thr	Pro	Ala	Gly	Arg	Leu	Pro	Thr	Pro	Ser	Gly	Thr	Asp	Asp
			170						175					180
Asp	Phe	Ala	Val	Thr	Thr	Pro	Ala	Gly	Ile	Gln	Arg	Ser	Thr	His
			185						190					195
Ala	Ile	Glu	Glu	Ala	Thr	Thr	Glu	Ser	Ala	Asn	Gly	Ile	Gln	
			200						205					

<210> SEQ ID NO 97
<211> LENGTH: 2848
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 97

gctcaagtgc cctgccttgc cccaccacgc ccagcctggc cagagccccc	50
tggagaagga gctctcttct tgcttggcag ctggaccaag ggagccagtc	100
ttgggcgctg gagggcctgt cctgaccatg gtccctgcct ggctgtggct	150
gctttgtgtc tccgtccccc aggcctctcc caaggcccag cctgcagagc	200
tgtctgtgga agttccagaa aactatgggt gaaatttccc tttataacct	250
accaagtgtc cgctgccccg tgagggggct gaaggccaga tcgtgctgtc	300
aggggactca ggcaaggcaa ctgagggccc atttgctatg gatccagatt	350
ctggcttcct gctggtgacc agggccctgg accgagagga gcaggcagag	400
taccagctac aggtcaccct ggagatgcag gatggacatg tcttgtgggg	450
tccacagcct gtgcttgtgc acgtgaagga tgagaatgac caggtgcccc	500
atttctctca agccatctac agagctcggc tgagccgggg taccaggcct	550
ggcatcccct tcctcttcct tgaggcttca gaccgggatg agccaggcac	600
agccaactcg gatcttcgat tccacatcct gagccaggct ccagcccagc	650
cttccccaga catgttccag ctggagcctc ggctgggggc tctggccctc	700
agccccaagg ggagcaccag ccttgaccac gccctggaga ggacctacca	750
gctgttggtg caggtcaagg acatgggtga ccaggcctca ggccaccag	800
ccactgccac cgtggaagtc tccatcatag agagcacctg ggtgtcccta	850
gagcctatcc acctggcaga gaatctcaaa gtcctatacc cgcaccacat	900
ggcccaggta cactggagtg ggggtgatgt gcactatcac ctggagagcc	950

-continued

atcccccg	accctttgaa	gtgaatgcag	agggaaacct	ctacgtgacc	1000
agagagctgg	acagagaagc	ccaggctgag	tacctgctcc	aggtgcgggc	1050
tcagaattcc	catggcgagg	actatgcggc	ccctctggag	ctgcacgtgc	1100
tggtgatgga	tgagaatgac	aacgtgccta	tctgccctcc	ccgtgacccc	1150
acagtcagca	tccctgagct	cagtccacca	ggtactgaag	tgactagact	1200
gtcagcagag	gatgcagatg	cccccgctc	ccccaattcc	cacgttgtgt	1250
atcagctcct	gagccctgag	cctgaggatg	gggtagaggg	gagagccttc	1300
cagggtgacc	ccacttcagg	cagtgtgacg	ctgggggtgc	tcccaactccg	1350
agcaggccag	aacatcctgc	ttctgggtgct	ggccatggac	ctggcaggcg	1400
cagaggggtg	cttcagcagc	acgtgtgaag	tcgaagtgcg	agtcacagat	1450
atcaatgatc	acgcccctga	gttcactact	tcccagattg	ggcctataag	1500
cctccctgag	gatgtggagc	ccgggactct	ggtggccatg	ctaacagcca	1550
ttgatgtcga	cctcgagccc	gccttcggcc	tcattggattt	tgccattgag	1600
aggggagaca	cagaagggac	ttttggcctg	gattgggagc	cagactctgg	1650
gcatgttaga	ctcagactct	gcaagaacct	cagttatgag	gcagctccaa	1700
gtcatgaggt	ggtggtggtg	gtgcagagtg	tggcgaagct	ggtggggcca	1750
ggcccaggcc	ctggagccac	cgccacggtg	actgtgctag	tggagagagt	1800
gatgccaccc	ccaagttagg	accaggagag	ctacgaggcc	agtgtcccca	1850
tcagtgcacc	agccggctct	ttcctgctga	ccatccagcc	ctccgacccc	1900
atcagccgaa	ccctcagggt	ctccctagtc	aatgactcag	agggtggct	1950
ctgcattgag	aaattctccg	gggaggtgca	caccgccag	tcctgcagg	2000
gcgcccagcc	tggggacacc	tacacggtgc	ttgtggaggc	ccaggataca	2050
gccctgactc	ttgcccctgt	gcctcccaa	tacctctgca	caccccgcca	2100
agaccatggc	ttgatcgtga	gtggaccag	caaggacccc	gatctggcca	2150
gtgggcacgg	tcctacagc	ttcacccttg	gtcccaaccc	cacggtgcaa	2200
cgggattggc	gcctccagac	tctcaatggt	tcccatgcct	acctcacctt	2250
ggccctgcat	tgggtggagc	cacgtgaaca	cataatcccc	gtgggtgtca	2300
gccacaatgc	ccagatgtgg	cagctcctgg	ttcgagtgat	cgtgtgtcgc	2350
tgcaacgtgg	aggggcagtg	catgcgcaag	gtgggccgca	tgaagggcat	2400
ggccacgaag	ctgtcggcag	tgggcatcct	tgtaggcacc	ctggtagcaa	2450
taggaatctt	cctcatcctc	attttcaccc	actggaccat	gtcaaggaag	2500
aaggaccggg	atcaaccagc	agacagcgtg	cccctgaagg	cgactgtctg	2550
aatggcccg	gcagctctag	ctgggagctt	ggcctctggc	tccatctgag	2600
tcccctggga	gagagcccag	caccaagat	ccagcagggg	acaggacaga	2650
gtagaagccc	ctccatctgc	cctgggggtg	aggcaccatc	accatcacca	2700
ggcatgtctg	cagagcctgg	acaccaactt	tatggactgc	ccatgggagt	2750
gtcccaaatg	tcagggtgtt	tgcccaataa	taaagcccca	gagaactggg	2800
ctggggcccta	tgggaaaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaag	2848

-continued

<210> SEQ ID NO 98

<211> LENGTH: 807

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 98

```

Met Val Pro Ala Trp Leu Trp Leu Leu Cys Val Ser Val Pro Gln
 1          5          10          15

Ala Leu Pro Lys Ala Gln Pro Ala Glu Leu Ser Val Glu Val Pro
 20          25          30

Glu Asn Tyr Gly Gly Asn Phe Pro Leu Tyr Leu Thr Lys Leu Pro
 35          40          45

Leu Pro Arg Glu Gly Ala Glu Gly Gln Ile Val Leu Ser Gly Asp
 50          55          60

Ser Gly Lys Ala Thr Glu Gly Pro Phe Ala Met Asp Pro Asp Ser
 65          70          75

Gly Phe Leu Leu Val Thr Arg Ala Leu Asp Arg Glu Glu Gln Ala
 80          85          90

Glu Tyr Gln Leu Gln Val Thr Leu Glu Met Gln Asp Gly His Val
 95          100         105

Leu Trp Gly Pro Gln Pro Val Leu Val His Val Lys Asp Glu Asn
110          115         120

Asp Gln Val Pro His Phe Ser Gln Ala Ile Tyr Arg Ala Arg Leu
125          130         135

Ser Arg Gly Thr Arg Pro Gly Ile Pro Phe Leu Phe Leu Glu Ala
140          145         150

Ser Asp Arg Asp Glu Pro Gly Thr Ala Asn Ser Asp Leu Arg Phe
155          160         165

His Ile Leu Ser Gln Ala Pro Ala Gln Pro Ser Pro Asp Met Phe
170          175         180

Gln Leu Glu Pro Arg Leu Gly Ala Leu Ala Leu Ser Pro Lys Gly
185          190         195

Ser Thr Ser Leu Asp His Ala Leu Glu Arg Thr Tyr Gln Leu Leu
200          205         210

Val Gln Val Lys Asp Met Gly Asp Gln Ala Ser Gly His Gln Ala
215          220         225

Thr Ala Thr Val Glu Val Ser Ile Ile Glu Ser Thr Trp Val Ser
230          235         240

Leu Glu Pro Ile His Leu Ala Glu Asn Leu Lys Val Leu Tyr Pro
245          250         255

His His Met Ala Gln Val His Trp Ser Gly Gly Asp Val His Tyr
260          265         270

His Leu Glu Ser His Pro Pro Gly Pro Phe Glu Val Asn Ala Glu
275          280         285

Gly Asn Leu Tyr Val Thr Arg Glu Leu Asp Arg Glu Ala Gln Ala
290          295         300

Glu Tyr Leu Leu Gln Val Arg Ala Gln Asn Ser His Gly Glu Asp
305          310         315

Tyr Ala Ala Pro Leu Glu Leu His Val Leu Val Met Asp Glu Asn
320          325         330

Asp Asn Val Pro Ile Cys Pro Pro Arg Asp Pro Thr Val Ser Ile

```

-continued

335					340					345				
Pro	Glu	Leu	Ser	Pro	Pro	Gly	Thr	Glu	Val	Thr	Arg	Leu	Ser	Ala
				350					355					360
Glu	Asp	Ala	Asp	Ala	Pro	Gly	Ser	Pro	Asn	Ser	His	Val	Val	Tyr
				365					370					375
Gln	Leu	Leu	Ser	Pro	Glu	Pro	Glu	Asp	Gly	Val	Glu	Gly	Arg	Ala
				380					385					390
Phe	Gln	Val	Asp	Pro	Thr	Ser	Gly	Ser	Val	Thr	Leu	Gly	Val	Leu
				395					400					405
Pro	Leu	Arg	Ala	Gly	Gln	Asn	Ile	Leu	Leu	Leu	Val	Leu	Ala	Met
				410					415					420
Asp	Leu	Ala	Gly	Ala	Glu	Gly	Gly	Phe	Ser	Ser	Thr	Cys	Glu	Val
				425					430					435
Glu	Val	Ala	Val	Thr	Asp	Ile	Asn	Asp	His	Ala	Pro	Glu	Phe	Ile
				440					445					450
Thr	Ser	Gln	Ile	Gly	Pro	Ile	Ser	Leu	Pro	Glu	Asp	Val	Glu	Pro
				455					460					465
Gly	Thr	Leu	Val	Ala	Met	Leu	Thr	Ala	Ile	Asp	Ala	Asp	Leu	Glu
				470					475					480
Pro	Ala	Phe	Arg	Leu	Met	Asp	Phe	Ala	Ile	Glu	Arg	Gly	Asp	Thr
				485					490					495
Glu	Gly	Thr	Phe	Gly	Leu	Asp	Trp	Glu	Pro	Asp	Ser	Gly	His	Val
				500					505					510
Arg	Leu	Arg	Leu	Cys	Lys	Asn	Leu	Ser	Tyr	Glu	Ala	Ala	Pro	Ser
				515					520					525
His	Glu	Val	Val	Val	Val	Val	Gln	Ser	Val	Ala	Lys	Leu	Val	Gly
				530					535					540
Pro	Gly	Pro	Gly	Pro	Gly	Ala	Thr	Ala	Thr	Val	Thr	Val	Leu	Val
				545					550					555
Glu	Arg	Val	Met	Pro	Pro	Pro	Lys	Leu	Asp	Gln	Glu	Ser	Tyr	Glu
				560					565					570
Ala	Ser	Val	Pro	Ile	Ser	Ala	Pro	Ala	Gly	Ser	Phe	Leu	Leu	Thr
				575					580					585
Ile	Gln	Pro	Ser	Asp	Pro	Ile	Ser	Arg	Thr	Leu	Arg	Phe	Ser	Leu
				590					595					600
Val	Asn	Asp	Ser	Glu	Gly	Trp	Leu	Cys	Ile	Glu	Lys	Phe	Ser	Gly
				605					610					615
Glu	Val	His	Thr	Ala	Gln	Ser	Leu	Gln	Gly	Ala	Gln	Pro	Gly	Asp
				620					625					630
Thr	Tyr	Thr	Val	Leu	Val	Glu	Ala	Gln	Asp	Thr	Ala	Leu	Thr	Leu
				635					640					645
Ala	Pro	Val	Pro	Ser	Gln	Tyr	Leu	Cys	Thr	Pro	Arg	Gln	Asp	His
				650					655					660
Gly	Leu	Ile	Val	Ser	Gly	Pro	Ser	Lys	Asp	Pro	Asp	Leu	Ala	Ser
				665					670					675
Gly	His	Gly	Pro	Tyr	Ser	Phe	Thr	Leu	Gly	Pro	Asn	Pro	Thr	Val
				680					685					690
Gln	Arg	Asp	Trp	Arg	Leu	Gln	Thr	Leu	Asn	Gly	Ser	His	Ala	Tyr
				695					700					705
Leu	Thr	Leu	Ala	Leu	His	Trp	Val	Glu	Pro	Arg	Glu	His	Ile	Ile
				710					715					720

-continued

Pro Val Val Val Ser His Asn Ala Gln Met Trp Gln Leu Leu Val
725 730 735
Arg Val Ile Val Cys Arg Cys Asn Val Glu Gly Gln Cys Met Arg
740 745 750
Lys Val Gly Arg Met Lys Gly Met Pro Thr Lys Leu Ser Ala Val
755 760 765
Gly Ile Leu Val Gly Thr Leu Val Ala Ile Gly Ile Phe Leu Ile
770 775 780
Leu Ile Phe Thr His Trp Thr Met Ser Arg Lys Lys Asp Pro Asp
785 790 795
Gln Pro Ala Asp Ser Val Pro Leu Lys Ala Thr Val
800 805

<210> SEQ ID NO 99
<211> LENGTH: 2436
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 99

ggctgaccgt gctacattgc ctggaggaag cctaaggaac ccaggcatcc 50
agctgcccac gcctgagtcc aagattcttc ccaggaacac aaacgtagga 100
gaccacgct cctggaagca ccagccttta tctcttcacc ttcaagtccc 150
ctttctcaag aatcctctgt tctttgccct ctaaagtctt ggtacatcta 200
ggaccacggc atcttgcttt ccagccacaa agagacagat gaagatgcag 250
aaaggaaatg ttctccttat gtttggctta ctattgcatt tagaagctgc 300
aacaaattcc aatgagacta gcacctctgc caaactgga tccagtgtga 350
tctccagtgg agccagcaca gccaccaact ctgggtccag tgtgacctcc 400
agtggggcca gcacagccac catctcaggg tccagcgtga cctccaatgg 450
ggtcagcata gtcaccaact ctgagttcca tacaacctcc agtgggatca 500
gcacagccac caactctgag ttcagcacag cgtccagtgg gatcagcata 550
gccaccaact ctgagtccag cacaacctcc agtggggcca gcacagccac 600
caactctgag tccagcacac cctccagtgg ggccagcaca gtcaccaact 650
ctgggtccag tgtgacctcc agtggagcca gcaactgcca caactctgag 700
tccagcacag tgtccagtag ggccagcact gccaccaact ctgagtctag 750
cacactctcc agtggggcca gcacagccac caactctgac tccagcaca 800
cctccagtgg ggctagcaca gccaccaact ctgagtccag cacaacctcc 850
agtggggcca gcacagccac caactctgag tccagcacag tgtccagtag 900
ggccagcact gccaccaact ctgagtccag cacaacctcc agtggggcca 950
gcacagccac caactctgag tccagaacga cctccaatgg ggctggcaca 1000
gccaccaact ctgagtccag cagcacctcc agtggggcca gcacagccac 1050
caactctgag tccagcacag tgtccagtgg ggccagcact gccaccaact 1100
ctgagtccag cagcacctcc agtggggcca gcacagccac caactctgag 1150
tccagcacga cctccagtgg ggctagcaca gccaccaact ctgactccag 1200
cacaacctcc agtggggcca gcacagccac caactctgag tccagcacag 1250

-continued

tgtccagtgg gatcagcaca gtcaccaatt ctgagtccag cacaccctcc	1300
agtggggcca acacagccac caactctgag tccagtacga cctccagtgg	1350
ggccaacaca gccaccaact ctgagtccag cacagtgtcc agtggggcca	1400
gcactgccac caactctgag tccagcacia cctccagtgg ggtcagcaca	1450
gccaccaact ctgagtccag cacaacctcc agtggggcta gcacagccac	1500
caactctgac tccagcacia cctccagtga ggccagcaca gccaccaact	1550
ctgagtctag cacagtgtcc agtgggatca gcacagtcac caattctgag	1600
tccagcacia cctccagtgg ggccaacaca gccaccaact ctgggtccag	1650
tgtgacctct gcaggctctg gaacagcagc tctgactgga atgcacacia	1700
cttcccatag tgcattctact gcagtgagtg aggcaaaaggc tgggtgggtcc	1750
ctgggtccgtg gggaaatctt cctcatcacc ctggtctcgg ttgtggcggc	1800
cggtggggctc ttgtctgggc tcttcttctg tgtgagaaac agcctgtccc	1850
tgagaaacac ctttaacaca gctgtctacc accctcatgg cctcaacct	1900
ggccttggtc caggccctgg agggaaatcat ggagccccc acaggcccag	1950
gtggagtcct aactggttct ggaggagacc agtatcatcg atagccatgg	2000
agatgagcgg gaggaacagc gggccctgag cagccccga agcaagtgcc	2050
gcattcttca ggaaggaaga gaactgggca cccaagacct ggtttcctt	2100
cattcatccc aggagacccc tcccagcttt gtttgagatc ctgaaaatct	2150
tgaagaaggt attcctcacc tttcttgctt ttaccagaca ctggaaagag	2200
aatactatat tgctcattta gctaagaaat aaatacatct catctaacc	2250
acacgacaaa gagaagctgt gcttgccccg ggggtgggtat ctgctctga	2300
gatgaactca gttataggag aaaacctcca tgctggactc catctggcat	2350
tcaaaatctc cacagtaaaa tccaagacc tcaaaaaaaaa aaaaaaaaaa	2400
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa	2436

<210> SEQ ID NO 100

<211> LENGTH: 596

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 100

Met	Lys	Met	Gln	Lys	Gly	Asn	Val	Leu	Leu	Met	Phe	Gly	Leu	Leu
1				5				10					15	
Leu	His	Leu	Glu	Ala	Ala	Thr	Asn	Ser	Asn	Glu	Thr	Ser	Thr	Ser
			20					25					30	
Ala	Asn	Thr	Gly	Ser	Ser	Val	Ile	Ser	Ser	Gly	Ala	Ser	Thr	Ala
			35					40					45	
Thr	Asn	Ser	Gly	Ser	Ser	Val	Thr	Ser	Ser	Gly	Val	Ser	Thr	Ala
			50					55					60	
Thr	Ile	Ser	Gly	Ser	Ser	Val	Thr	Ser	Asn	Gly	Val	Ser	Ile	Val
			65					70					75	
Thr	Asn	Ser	Glu	Phe	His	Thr	Thr	Ser	Ser	Gly	Ile	Ser	Thr	Ala
			80					85					90	
Thr	Asn	Ser	Glu	Phe	Ser	Thr	Ala	Ser	Ser	Gly	Ile	Ser	Ile	Ala

					95					100					105
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				110					115					120	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Pro	Ser	Ser	Gly	Ala	Ser	Thr	Val	
				125					130					135	
Thr	Asn	Ser	Gly	Ser	Ser	Val	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				140					145					150	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Val	Ser	Ser	Arg	Ala	Ser	Thr	Ala	
				155					160					16	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Leu	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				170					175					180	
Thr	Asn	Ser	Asp	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				185					190					195	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				200					205					210	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Val	Ser	Ser	Arg	Ala	Ser	Thr	Ala	
				215					220					225	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				230					235					240	
Thr	Asn	Ser	Glu	Ser	Arg	Thr	Thr	Ser	Asn	Gly	Ala	Gly	Thr	Ala	
				245					250					255	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				260					265					270	
Thr	Asn	Ser	Asp	Ser	Ser	Thr	Val	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				275					280					285	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				290					295					300	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				305					310					315	
Thr	Asn	Ser	Asp	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Gly	Thr	Ala	
				320					325					330	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Val	Ser	Ser	Gly	Ile	Ser	Thr	Val	
				335					340					345	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Pro	Ser	Ser	Gly	Ala	Asn	Thr	Ala	
				350					355					360	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Asn	Thr	Ala	
				365					370					375	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Val	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				380					385					390	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Val	Ser	Thr	Ala	
				395					400					405	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Ser	Thr	Ala	
				410					415					420	
Thr	Asn	Ser	Asp	Ser	Ser	Thr	Thr	Ser	Ser	Glu	Ala	Ser	Thr	Ala	
				425					430					435	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Val	Ser	Ser	Gly	Ile	Ser	Thr	Val	
				440					445					450	
Thr	Asn	Ser	Glu	Ser	Ser	Thr	Thr	Ser	Ser	Gly	Ala	Asn	Thr	Ala	
				455					460					465	
Thr	Asn	Ser	Gly	Ser	Ser	Val	Thr	Ser	Ala	Gly	Ser	Gly	Thr	Ala	
				470					475					480	

-continued

Ala	Leu	Thr	Gly	Met	His	Thr	Thr	Ser	His	Ser	Ala	Ser	Thr	Ala
				485					490					495
Val	Ser	Glu	Ala	Lys	Pro	Gly	Gly	Ser	Leu	Val	Pro	Trp	Glu	Ile
				500					505					510
Phe	Leu	Ile	Thr	Leu	Val	Ser	Val	Val	Ala	Ala	Val	Gly	Leu	Phe
				515					520					525
Ala	Gly	Leu	Phe	Phe	Cys	Val	Arg	Asn	Ser	Leu	Ser	Leu	Arg	Asn
				530					535					540
Thr	Phe	Asn	Thr	Ala	Val	Tyr	His	Pro	His	Gly	Leu	Asn	His	Gly
				545					550					555
Leu	Gly	Pro	Gly	Pro	Gly	Gly	Asn	His	Gly	Ala	Pro	His	Arg	Pro
				560					565					570
Arg	Trp	Ser	Pro	Asn	Trp	Phe	Trp	Arg	Arg	Pro	Val	Ser	Ser	Ile
				575					580					585
Ala	Met	Glu	Met	Ser	Gly	Arg	Asn	Ser	Gly	Pro				
				590					595					

<210> SEQ ID NO 101
<211> LENGTH: 1728
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 101

ggccggacgc	ctccgcgtta	cgggatgaat	taacggcggg	ttccgcacgg	50
aggttgtgac	ccctacggag	ccccagcttg	cccacgcacc	ccactcggcg	100
tcgcgcggcg	tgccctgctt	gtcacagggtg	ggaggctgga	actatcaggc	150
tgaaaaacag	agtgggtact	ctcttctggg	aagctggcaa	caaatggatg	200
atgtgatata	tgcattccag	gggaagggaa	attgtggtgc	ttctgaaccc	250
atgggtcaatt	aacgaggcag	tttctagcta	ctgcacgtac	ttcataaagc	300
aggactctaa	aagctttgga	atcatgggtgt	catggaaagg	gatttacttt	350
atactgactc	tgttttgggg	aagctttttt	ggaagcattt	tcatgctgag	400
tcccttttta	cctttgatgt	ttgtaaacc	atcttggtat	cgtcggatca	450
acaaccgcct	tgtggcaaca	tggtccacc	tacctgtggc	attattggag	500
accatgtttg	gtgtaaaagt	gattataact	ggggatgcat	ttgttcctgg	550
agaaagaagt	gtcattatca	tgaaccatcg	gacaagaatg	gactggatgt	600
tcctgtggaa	ttgcctgatg	cgatatagct	acctcagatt	ggagaaaatt	650
tgccctcaaag	cgagtctcaa	agggtgtcct	ggatttggtt	gggccatgca	700
ggctgctgcc	tatatcttca	ttcataggaa	atggaaggat	gacaagagcc	750
atttcgaaga	catgattgat	tacttttgtg	atattcacga	accacttcaa	800
ctcctcatat	tcccagaagg	gactgatctc	acagaaaaca	gcaagtctcg	850
aagtaatgca	tttgctgaaa	aaaatggact	tcagaaatat	gaatatgttt	900
tacatccaag	aactacaggc	tttacttttg	tggtagaccg	tctaagagaa	950
ggtaagaacc	ttgatgctgt	ccatgatatc	actgtggcgt	atcctcacia	1000
cattcctcaa	tcagagaagc	acctcctcca	aggagacttt	cccagggaaa	1050
tccactttca	cgtccaccgg	tatccaatag	acaccctccc	cacatccaag	1100

-continued

```
gaggaccttc aactctggtg ccacaaacgg tgggaagaga aagaagagag      1150
gctgcgttcc ttctatcaag gggagaagaa tttttatttt accggacaga      1200
tgttcattcc accttgcaag tctgaactca gggtccttgt ggtcaaatg      1250
ctctctatac tgtattggac cctgttcagc cctgcaatgt gcctactcat      1300
atatattgtac agtccttgta agtggtattt tataatcacc attgtaatct      1350
ttgtgctgca agagagaata tttggtggac tggagatcat agaacttgca      1400
tgttaccgac ttttacacaa acagccacat ttaaattcaa agaaaaatga      1450
gtaagattat aagggtttgcc atgtgaaaac ctagagcata ttttgaaaat      1500
gttctaaacc tttctaagct cagatgcatt tttgcatgac tatgtcgaat      1550
atctcttact gccatcatta ttgtttaaag atattttgca cttaattttg      1600
tgggaaaaat attgctacaa ttttttttaa tctctgaatg taatttcgat      1650
actgtgtaca tagcaggagg tgatcggggg gaaataactt gggccagaat      1700
attattaaac aatcatcagg cttttaaa      1728
```

<210> SEQ ID NO 102

<211> LENGTH: 414

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 102

```
Met His Ser Arg Gly Arg Glu Ile Val Val Leu Leu Asn Pro Trp
 1             5             10             15
Ser Ile Asn Glu Ala Val Ser Ser Tyr Cys Thr Tyr Phe Ile Lys
          20             25             30
Gln Asp Ser Lys Ser Phe Gly Ile Met Val Ser Trp Lys Gly Ile
          35             40             45
Tyr Phe Ile Leu Thr Leu Phe Trp Gly Ser Phe Phe Gly Ser Ile
          50             55             60
Phe Met Leu Ser Pro Phe Leu Pro Leu Met Phe Val Asn Pro Ser
          65             70             75
Trp Tyr Arg Trp Ile Asn Asn Arg Leu Val Ala Thr Trp Leu Thr
          80             85             90
Leu Pro Val Ala Leu Leu Glu Thr Met Phe Gly Val Lys Val Ile
          95             100            105
Ile Thr Gly Asp Ala Phe Val Pro Gly Glu Arg Ser Val Ile Ile
          110            115            120
Met Asn His Arg Thr Arg Met Asp Trp Met Phe Leu Trp Asn Cys
          125            130            135
Leu Met Arg Tyr Ser Tyr Leu Arg Leu Glu Lys Ile Cys Leu Lys
          140            145            150
Ala Ser Leu Lys Gly Val Pro Gly Phe Gly Trp Ala Met Gln Ala
          155            160            165
Ala Ala Tyr Ile Phe Ile His Arg Lys Trp Lys Asp Asp Lys Ser
          170            175            180
His Phe Glu Asp Met Ile Asp Tyr Phe Cys Asp Ile His Glu Pro
          185            190            195
Leu Gln Leu Leu Ile Phe Pro Glu Gly Thr Asp Leu Thr Glu Asn
          200            205            210
```

-continued

Ser	Lys	Ser	Arg	Ser	Asn	Ala	Phe	Ala	Glu	Lys	Asn	Gly	Leu	Gln
				215					220					225
Lys	Tyr	Glu	Tyr	Val	Leu	His	Pro	Arg	Thr	Thr	Gly	Phe	Thr	Phe
				230					235					240
Val	Val	Asp	Arg	Leu	Arg	Glu	Gly	Lys	Asn	Leu	Asp	Ala	Val	His
				245					250					255
Asp	Ile	Thr	Val	Ala	Tyr	Pro	His	Asn	Ile	Pro	Gln	Ser	Glu	Lys
				260					265					270
His	Leu	Leu	Gln	Gly	Asp	Phe	Pro	Arg	Glu	Ile	His	Phe	His	Val
				275					280					285
His	Arg	Tyr	Pro	Ile	Asp	Thr	Leu	Pro	Thr	Ser	Lys	Glu	Asp	Leu
				290					295					300
Gln	Leu	Trp	Cys	His	Lys	Arg	Trp	Glu	Glu	Lys	Glu	Glu	Arg	Leu
				305					310					315
Arg	Ser	Phe	Tyr	Gln	Gly	Glu	Lys	Asn	Phe	Tyr	Phe	Thr	Gly	Gln
				320					325					330
Ser	Val	Ile	Pro	Pro	Cys	Lys	Ser	Glu	Leu	Arg	Val	Leu	Val	Val
				335					340					345
Lys	Leu	Leu	Ser	Ile	Leu	Tyr	Trp	Thr	Leu	Phe	Ser	Pro	Ala	Met
				350					355					360
Cys	Leu	Leu	Ile	Tyr	Leu	Tyr	Ser	Leu	Val	Lys	Trp	Tyr	Phe	Ile
				365					370					375
Ile	Thr	Ile	Val	Ile	Phe	Val	Leu	Gln	Glu	Arg	Ile	Phe	Gly	Gly
				380					385					390
Leu	Glu	Ile	Ile	Glu	Leu	Ala	Cys	Tyr	Arg	Leu	Leu	His	Lys	Gln
				395					400					405
Pro	His	Leu	Asn	Ser	Lys	Lys	Asn	Glu						
				410										

<210> SEQ ID NO 103
<211> LENGTH: 2403
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 103

cggtctcgagc	ggctcgagtg	aagagcctct	ccacggctcc	tgcgctgag	50
acagctggcc	tgacctcaa	atcatccatc	caccctgct	gtcatctgtt	100
ttcatagtgt	gagatcaacc	cacaggaata	tccatggctt	ttgtgctcat	150
tttggttctc	agtttctacg	agctggtgto	aggacagtgg	caagtcactg	200
gaccgggcaa	gtttgtccag	gccttggtgg	gggaggacgc	cgtgttctcc	250
tgctccctct	ttcctgagac	cagtgcagag	gctatggaag	tgcggttctt	300
caggaatcag	ttccatgctg	tggtccacct	ctacagagat	ggggaagact	350
gggaatctaa	gcagatgcca	cagtatcgag	ggagaactga	gtttgtgaag	400
gactccattg	caggggggcg	tgtctctcta	aggctaaaaa	acatcactcc	450
ctcggacatc	ggcctgtatg	ggtgctggtt	cagttcccag	atttacgatg	500
aggaggccac	ctgggagctg	cgggtggcag	cactgggctc	acttcctctc	550
atttccatcg	tgggatatgt	tgacggaggt	atccagttac	tctgcctgtc	600
ctcaggctgg	ttcccccagc	ccacagccaa	gtggaaaggt	ccacaaggac	650

-continued

aggatttgtc ttcagactcc agagcaaatg cagatgggta cagcctgtat	700
gatgtggaga tctccattat agtccaggaa aatgctggga gcatattgtg	750
ttccatccac cttgctgagc agagtcatga ggtggaatcc aaggtattga	800
taggagagac gtttttccag ccctcacctt ggcgctggc ttctatttta	850
ctcgggttac tctgtggtgc cctgtgtggt gttgtcatgg ggatgataat	900
tgttttcttc aaatccaaag ggaaaatcca ggcggaactg gactggagaa	950
gaaagcacgg acaggcagaa ttgagagacg cccggaaaca cgcagtggag	1000
gtgactctgg atccagagac ggctcacccg aagctctgcg tttctgatct	1050
gaaaactgta acccatagaa aagctcccca ggaggtgcct cactctgaga	1100
agagatttac aaggaagagt gtggtggctt ctcaggggtt ccaagcagg	1150
agacattact gggaggtgga cgtgggacaa aatgtagggt ggtatgtgg	1200
agtgtgtcgg gatgacgtag acagggggaa gaacaatgtg actttgtctc	1250
ccaacaatgg gtattgggtc ctcagactga caacagaaca ttgtatttc	1300
acattcaatc cccattttat cagcctcccc cccagcacc ctcctacacg	1350
agttaggggtc ttcttgact atgaggggtg gaccatctcc ttcttcaata	1400
caaatgacca gtcccttatt tataccctgc tgacatgtca gtttgaaggc	1450
ttgttgagac cctatatcca gcatgcgatg tatgacgagg aaaaggggac	1500
tcccatatct atatgtccag tgcctgggg atgagacaga gaagaccctg	1550
cttaaagggc cccacaccac agaccagac acagccaagg gagagtgtc	1600
ccgacaggtg gcccagctt cctctccgga gcctgcgcac agagagtcac	1650
gccccact ctcttttagg gagctgaggt tctctgccc tgagccctgc	1700
agcagcgga gtcacagctt ccagatgagg ggggattggc ctgaccctgt	1750
gggagtcaga agccatggct gcctgaagt ggggacggaa tagactaca	1800
ttaggttag tttgtgaaaa ctccatccag ctaagcgatc ttgaacaagt	1850
cacaacctcc caggctcctc atttgctagt cacggacagt gattcctgcc	1900
tcacagggtg agattaaaga gacaacgaat gtgaatcatg cttgcagggt	1950
tgagggcaca gtgtttgcta atgatgtgtt tttatattat acattttccc	2000
accataaact ctgtttgctt attccacatt aatttacttt tctctatacc	2050
aaatcaccca tggaatagtt attgaacacc tgctttgtga ggctcaaaga	2100
ataaagagga ggtaggattt ttcactgatt ctataagccc agcattacct	2150
gataccaaaa ccaggcaaag aaaacagaag aagaggaagg aaaactacag	2200
gtccatatcc ctcattaaca cagacacaaa aattctaaat aaaatttta	2250
caaattaaac taaacaatat atttaaagat gatataaac tactcagtgt	2300
ggtttgtccc acaaatgcag agttggttta atatttaaat atcaaccagt	2350
gtaattcagc acattaataa agtaaaaaag aaaaccataa aaaaaaaaaa	2400
aaa	2403

<210> SEQ ID NO 104

<211> LENGTH: 466

-continued

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 104

```

Met Ala Phe Val Leu Ile Leu Val Leu Ser Phe Tyr Glu Leu Val
 1             5             10             15

Ser Gly Gln Trp Gln Val Thr Gly Pro Gly Lys Phe Val Gln Ala
          20             25             30

Leu Val Gly Glu Asp Ala Val Phe Ser Cys Ser Leu Phe Pro Glu
          35             40             45

Thr Ser Ala Glu Ala Met Glu Val Arg Phe Phe Arg Asn Gln Phe
          50             55             60

His Ala Val Val His Leu Tyr Arg Asp Gly Glu Asp Trp Glu Ser
          65             70             75

Lys Gln Met Pro Gln Tyr Arg Gly Arg Thr Glu Phe Val Lys Asp
          80             85             90

Ser Ile Ala Gly Gly Arg Val Ser Leu Arg Leu Lys Asn Ile Thr
          95             100            105

Pro Ser Asp Ile Gly Leu Tyr Gly Cys Trp Phe Ser Ser Gln Ile
          110            115            120

Tyr Asp Glu Glu Ala Thr Trp Glu Leu Arg Val Ala Ala Leu Gly
          125            130            135

Ser Leu Pro Leu Ile Ser Ile Val Gly Tyr Val Asp Gly Gly Ile
          140            145            150

Gln Leu Leu Cys Leu Ser Ser Gly Trp Phe Pro Gln Pro Thr Ala
          155            160            165

Lys Trp Lys Gly Pro Gln Gly Gln Asp Leu Ser Ser Asp Ser Arg
          170            175            180

Ala Asn Ala Asp Gly Tyr Ser Leu Tyr Asp Val Glu Ile Ser Ile
          185            190            195

Ile Val Gln Glu Asn Ala Gly Ser Ile Leu Cys Ser Ile His Leu
          200            205            210

Ala Glu Gln Ser His Glu Val Glu Ser Lys Val Leu Ile Gly Glu
          215            220            225

Thr Phe Phe Gln Pro Ser Pro Trp Arg Leu Ala Ser Ile Leu Leu
          230            235            240

Gly Leu Leu Cys Gly Ala Leu Cys Gly Val Val Met Gly Met Ile
          245            250            255

Ile Val Phe Phe Lys Ser Lys Gly Lys Ile Gln Ala Glu Leu Asp
          260            265            270

Trp Arg Arg Lys His Gly Gln Ala Glu Leu Arg Asp Ala Arg Lys
          275            280            285

His Ala Val Glu Val Thr Leu Asp Pro Glu Thr Ala His Pro Lys
          290            295            300

Leu Cys Val Ser Asp Leu Lys Thr Val Thr His Arg Lys Ala Pro
          305            310            315

Gln Glu Val Pro His Ser Glu Lys Arg Phe Thr Arg Lys Ser Val
          320            325            330

Val Ala Ser Gln Gly Phe Gln Ala Gly Arg His Tyr Trp Glu Val
          335            340            345

Asp Val Gly Gln Asn Val Gly Trp Tyr Val Gly Val Cys Arg Asp
          350            355            360

```

-continued

Asp	Val	Asp	Arg	Gly	Lys	Asn	Asn	Val	Thr	Leu	Ser	Pro	Asn	Asn
				365					370					375
Gly	Tyr	Trp	Val	Leu	Arg	Leu	Thr	Thr	Glu	His	Leu	Tyr	Phe	Thr
				380					385					390
Phe	Asn	Pro	His	Phe	Ile	Ser	Leu	Pro	Pro	Ser	Thr	Pro	Pro	Thr
				395					400					405
Arg	Val	Gly	Val	Phe	Leu	Asp	Tyr	Glu	Gly	Gly	Thr	Ile	Ser	Phe
				410					415					420
Phe	Asn	Thr	Asn	Asp	Gln	Ser	Leu	Ile	Tyr	Thr	Leu	Leu	Thr	Cys
				425					430					435
Gln	Phe	Glu	Gly	Leu	Leu	Arg	Pro	Tyr	Ile	Gln	His	Ala	Met	Tyr
				440					445					450
Asp	Glu	Glu	Lys	Gly	Thr	Pro	Ile	Phe	Ile	Cys	Pro	Val	Ser	Trp
				455					460					465

Gly

<210> SEQ ID NO 105
<211> LENGTH: 2103
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 105

ccttcacagg actcttcatt gctggttggc aatgatgtat cggccagatg	50
tggtgagggc taggaaaaga gtttgttggg aaccctgggt tatcggcctc	100
gtcatcttca tatccctgat tgtcctggca gtgtgcattg gactcactgt	150
tcattatgtg agatataatc aaaagaagac ctacaattac tatagcacat	200
tgtcatttac aactgacaaa ctatatgctg agtttggcag agaggcttct	250
aacaatttta cagaaatgag ccagagactt gaatcaatgg tgaaaaatgc	300
attttataaa tctccattaa gggaagaatt tgtcaagtct caggttatca	350
agttcagtea acagaagcat ggagtgttgg ctcatatgct gttgatttgt	400
agatttcact ctactgagga tcctgaaact gtagataaaa ttgttcaact	450
tgttttacat gaaaagctgc aagatgctgt aggacccctt aaagtagatc	500
ctcactcagt taaaattaaa aaaatcaaca agacagaaac agacagctat	550
ctaaaccatt gctgcggaac acgaagaagt aaaactctag gtcagagtct	600
caggatcggt ggtgggacag aagtagaaga gggatgaatg ccctggcagg	650
ctagcctgca gtgggatggg agtcatcgct gtggagcaac cttaattaat	700
gccacatggc ttgtgagtgc tgctcactgt tttacaacat ataagaacct	750
tgccagatgg actgcttcct ttggagtaac aataaaacct tcgaaaatga	800
aacggggtct ccggagaata attgtccatg aaaaatacaa acaccatca	850
catgactatg atatttctct tgcagagctt tctagccctg ttccctacac	900
aatgcagta catagagttt gtctccctga tgcacatctat gagtttcaac	950
caggatgatg gatgtttgtg acaggatttg gagcactgaa aaatgatgg	1000
tacagtcata atcatcttcg acaagcacag gtgactctca tagacgctac	1050
aacttgcaat gaacctcaag cttacaatga cgccataact cctagaatgt	1100

-continued

tatgtgctgg	ctccttagaa	ggaaaaacag	atgcatgcc	gggtgactct	1150
ggaggaccac	tggttagttc	agatgctaga	gatatctggt	accttgctgg	1200
aatagtgagc	tggggagatg	aatgtgcgaa	acccaacaag	cctggtgttt	1250
atactagagt	tacggccttg	cgggactgga	ttacttcaaa	aactggtatc	1300
taagagacaa	aagcctcatg	gaacagataa	catttttttt	tgttttttgg	1350
gtgtggaggc	cattttttaga	gatacagaat	tggagaagac	ttgcaaaaca	1400
gctagatttg	actgatctca	ataaactgtt	tgcttgatgc	atgtattttc	1450
ttcccagctc	tgttccgcac	gtaagcatcc	tgcttctgcc	agatcaactc	1500
tgatcatctg	gagcaatagt	tgaacttta	tgtacataga	gaaatagata	1550
atacaatatt	acattacagc	ctgtattcat	ttgttctcta	gaagttttgt	1600
cagaattttg	acttgttgac	ataaatttgt	aatgcatata	tacaatttga	1650
agcactcctt	ttcttcagtt	cctcagctcc	tctcatttca	gcaaatatcc	1700
attttcaagg	tgagaacaa	ggagtgaag	aaaatataag	aagaaaaaaa	1750
ttccctacat	tttattggca	cagaaaagta	ttaggtgttt	ttcttagtgg	1800
aatattagaa	atgatcatat	tcattatgaa	aggtaagca	aagacagcag	1850
aataccaatc	acttcatcat	ttaggaagta	tgggaactaa	gttaaggaag	1900
tccagaaaga	agccaagata	tatccttatt	ttcatttcca	aacaactact	1950
atgataaatg	tgaagaagat	tctgtttttt	tgtgacctat	aataattata	2000
caaacttcat	gcaatgtact	tggtctaagc	aaattaaagc	aaatatttat	2050
ttaacattgt	tactgaggat	gtcaacatat	aacaataaaa	tataaatcac	2100
cca					2103

<210> SEQ ID NO 106

<211> LENGTH: 423

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 106

Met	Met	Tyr	Arg	Pro	Asp	Val	Val	Arg	Ala	Arg	Lys	Arg	Val	Cys
1				5					10				15	
Trp	Glu	Pro	Trp	Val	Ile	Gly	Leu	Val	Ile	Phe	Ile	Ser	Leu	Ile
				20					25				30	
Val	Leu	Ala	Val	Cys	Ile	Gly	Leu	Thr	Val	His	Tyr	Val	Arg	Tyr
				35					40				45	
Asn	Gln	Lys	Lys	Thr	Tyr	Asn	Tyr	Tyr	Ser	Thr	Leu	Ser	Phe	Thr
				50					55				60	
Thr	Asp	Lys	Leu	Tyr	Ala	Glu	Phe	Gly	Arg	Glu	Ala	Ser	Asn	Asn
				65					70				75	
Phe	Thr	Glu	Met	Ser	Gln	Arg	Leu	Glu	Ser	Met	Val	Lys	Asn	Ala
				80					85				90	
Phe	Tyr	Lys	Ser	Pro	Leu	Arg	Glu	Glu	Phe	Val	Lys	Ser	Gln	Val
				95					100				105	
Ile	Lys	Phe	Ser	Gln	Gln	Lys	His	Gly	Val	Leu	Ala	His	Met	Leu
				110					115				120	
Leu	Ile	Cys	Arg	Phe	His	Ser	Thr	Glu	Asp	Pro	Glu	Thr	Val	Asp
				125					130				135	

-continued

Lys Ile Val Gln Leu Val Leu His Glu Lys Leu Gln Asp Ala Val
140 145 150

Gly Pro Pro Lys Val Asp Pro His Ser Val Lys Ile Lys Lys Ile
155 160 165

Asn Lys Thr Glu Thr Asp Ser Tyr Leu Asn His Cys Cys Gly Thr
170 175 180

Arg Arg Ser Lys Thr Leu Gly Gln Ser Leu Arg Ile Val Gly Gly
185 190 195

Thr Glu Val Glu Glu Gly Glu Trp Pro Trp Gln Ala Ser Leu Gln
200 205 210

Trp Asp Gly Ser His Arg Cys Gly Ala Thr Leu Ile Asn Ala Thr
215 220 225

Trp Leu Val Ser Ala Ala His Cys Phe Thr Thr Tyr Lys Asn Pro
230 235 240

Ala Arg Trp Thr Ala Ser Phe Gly Val Thr Ile Lys Pro Ser Lys
245 250 255

Met Lys Arg Gly Leu Arg Arg Ile Ile Val His Glu Lys Tyr Lys
260 265 270

His Pro Ser His Asp Tyr Asp Ile Ser Leu Ala Glu Leu Ser Ser
275 280 285

Pro Val Pro Tyr Thr Asn Ala Val His Arg Val Cys Leu Pro Asp
290 295 300

Ala Ser Tyr Glu Phe Gln Pro Gly Asp Val Met Phe Val Thr Gly
305 310 315

Phe Gly Ala Leu Lys Asn Asp Gly Tyr Ser Gln Asn His Leu Arg
320 325 330

Gln Ala Gln Val Thr Leu Ile Asp Ala Thr Thr Cys Asn Glu Pro
335 340 345

Gln Ala Tyr Asn Asp Ala Ile Thr Pro Arg Met Leu Cys Ala Gly
350 355 360

Ser Leu Glu Gly Lys Thr Asp Ala Cys Gln Gly Asp Ser Gly Gly
365 370 375

Pro Leu Val Ser Ser Asp Ala Arg Asp Ile Trp Tyr Leu Ala Gly
380 385 390

Ile Val Ser Trp Gly Asp Glu Cys Ala Lys Pro Asn Lys Pro Gly
395 400 405

Val Tyr Thr Arg Val Thr Ala Leu Arg Asp Trp Ile Thr Ser Lys
410 415 420

Thr Gly Ile

<210> SEQ ID NO 107
<211> LENGTH: 2397
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 107

agagaaagaa gcgtctccag ctgaagccaa tgcagccctc cggctctccg 50

cgaagaagtt ccctgccccg atgagcccc gccgtgcgtc cccgactatc 100

cccaggcggg cgtggggcac cgggcccagc gccgacgac gctgccgttt 150

tgcccttggg agtaggatgt ggtgaaagga tggggcttct cccttacggg 200

-continued

gctcacaatg gccagagaag attccgtgaa gtgtctgcgc tgctgtctct	250
acgcctctca tctgctcttt tggtaaatgt ccatcagtgt gttggcagtt	300
tctgcttgga tgagggacta cctaaataat gttctcactt taactgcaga	350
aacgagggta gaggaagcag tcattttgac ttactttcct gtgggtcctc	400
cggatcatgat tgctgtttgc tgtttcccta tcattgtggg gatgttagga	450
tattgtggaa cggtgaaaag aaatctgttg cttcttgcat ggtactttgg	500
aagtttgctt gtcattttct gtgtagaact ggcttgtggc gtttgacat	550
atgaacagga acttatgggt ccagtacaat ggtcagatat ggtcactttg	600
aaagccagga tgacaaatta tggattacct agatatcggg ggcttactca	650
tgcttggaat ttttttcaga gagagttaa gtgctgtgga gtagtatatt	700
tcactgactg gttggaaatg acagagatgg actggccccc agattcctgc	750
tgtgttagag aattcccagg atgttccaaa caggcccacc aggaagatct	800
cagtgcctt tatcaagagg gttgtgggaa gaaaatgtat tcctttttga	850
gaggaaccaa acaactgcag gtgctgaggt ttctgggaat ctccattggg	900
gtgacacaaa tcctggccat gattctcacc attactctgc tctgggctct	950
gtattatgat agaagggagc ctgggacaga ccaaatgatg tccttgaaga	1000
atgacaactc tcagcacctg tcatgtccct cagtagaact gttgaaacca	1050
agcctgtcaa gaatccttga acacacatcc atggcaaaca gctttaatac	1100
acactttgag atggaggagt tataaaaaga aatgtcacag aagaaaacca	1150
caaacttggt ttattggact tgtgaatttt tgagtacata ctatgtgttt	1200
cagaaatatg tagaaataaa aatgttgcca taaaataaca cctaagcata	1250
tactattcta tgctttaaaa tgaggatgga aaagtttcat gtcataagtc	1300
accacctgga caataattga tgcccttaaa atgctgaaga cagatgtcat	1350
accactgtg tagcctgtgt atgactttta ctgaacacag ttatgttttg	1400
aggcagcatg gtttgattag catttccgca tccatgcaa cgagtcacat	1450
atggtgggac tggagccata gtaaagggtg atttacttct accaactagt	1500
atataaagta ctaattaaat gctaacatag gaagttagaa aatactaata	1550
acttttatta ctacgcgac tattcttctg atgctaaata aattatatat	1600
cagaaaactt tcaatattgg tgactacct aatgtgattt ttgctgggta	1650
ctaaaaatatt cttaccactt aaaagagcaa gctaacacat tgtcttaagc	1700
tgatcaggga ttttttgat ataagtctgt gttaaactcg tataaattcag	1750
tcgatttcag ttctgataat gttaagaata accattatga aaaggaaaat	1800
ttgtcctgta tagcatcatt attttttagcc tttcctgtta ataaagcttt	1850
actattctgt cctgggctta tattacacat ataactgtta tttaaatact	1900
taaccactaa ttttgaaaat taccagtgtg atacatagga atcattattc	1950
agaatgtagt ctggtcttta ggaagtatta ataagaaaat ttgcacataa	2000
cttagttgat tcagaaagga cttgtatgct gtttttctcc caaatgaaga	2050
ctctttttga cactaaacac tttttaaaaa gcttatcttt gccttctcca	2100

-continued

aacaagaagc aatagtctcc aagtcaatat aaattctaca gaaaatagtg	2150
ttctttttct ccagaaaaat gcttgtagaga atcattaaaa catgtgacaa	2200
tttagagatt ctttgtttta ttctactgat taatatactg tggcaaatta	2250
cacagattat taaatttttt tacaagagta tagtatatit atttgaatg	2300
ggaaaagtgc attttactgt attttggtga ttttgtttat ttctcagaat	2350
atggaaagaa aattaaaatg tgtcaataaa tattttctag agagtaa	2397

<210> SEQ ID NO 108

<211> LENGTH: 305

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 108

Met Ala Arg Glu Asp Ser Val Lys Cys Leu Arg Cys Leu Leu Tyr	1	5	10	15
Ala Leu Asn Leu Leu Phe Trp Leu Met Ser Ile Ser Val Leu Ala	20	25	30	
Val Ser Ala Trp Met Arg Asp Tyr Leu Asn Asn Val Leu Thr Leu	35	40	45	
Thr Ala Glu Thr Arg Val Glu Glu Ala Val Ile Leu Thr Tyr Phe	50	55	60	
Pro Val Val His Pro Val Met Ile Ala Val Cys Cys Phe Leu Ile	65	70	75	
Ile Val Gly Met Leu Gly Tyr Cys Gly Thr Val Lys Arg Asn Leu	80	85	90	
Leu Leu Leu Ala Trp Tyr Phe Gly Ser Leu Leu Val Ile Phe Cys	95	100	105	
Val Glu Leu Ala Cys Gly Val Trp Thr Tyr Glu Gln Glu Leu Met	110	115	120	
Val Pro Val Gln Trp Ser Asp Met Val Thr Leu Lys Ala Arg Met	125	130	135	
Thr Asn Tyr Gly Leu Pro Arg Tyr Arg Trp Leu Thr His Ala Trp	140	145	150	
Asn Phe Phe Gln Arg Glu Phe Lys Cys Cys Gly Val Val Tyr Phe	155	160	165	
Thr Asp Trp Leu Glu Met Thr Glu Met Asp Trp Pro Pro Asp Ser	170	175	180	
Cys Cys Val Arg Glu Phe Pro Gly Cys Ser Lys Gln Ala His Gln	185	190	195	
Glu Asp Leu Ser Asp Leu Tyr Gln Glu Gly Cys Gly Lys Lys Met	200	205	210	
Tyr Ser Phe Leu Arg Gly Thr Lys Gln Leu Gln Val Leu Arg Phe	215	220	225	
Leu Gly Ile Ser Ile Gly Val Thr Gln Ile Leu Ala Met Ile Leu	230	235	240	
Thr Ile Thr Leu Leu Trp Ala Leu Tyr Tyr Asp Arg Arg Glu Pro	245	250	255	
Gly Thr Asp Gln Met Met Ser Leu Lys Asn Asp Asn Ser Gln His	260	265	270	
Leu Ser Cys Pro Ser Val Glu Leu Leu Lys Pro Ser Leu Ser Arg	275	280	285	

-continued

Ile Phe Glu His Thr Ser Met Ala Asn Ser Phe Asn Thr His Phe
290 295 300
Glu Met Glu Glu Leu
305

<210> SEQ ID NO 109
<211> LENGTH: 2339
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 109

ccaaggccag agctgtggac accttatccc actcatcctc atcctcttcc	50
tctgataaag cccctaccag tgctgataaa gtctttctcg tgagagccta	100
gaggccttaa aaaaaaaagt gcttgaaaga gaaggggaca aaggaacacc	150
agtattaaga ggattttcca gtgtttctgg cagttggtcc agaaggatgc	200
ctccattcct gcttctcacc tgccctcttca tcacaggcac ctccgtgtca	250
cccgtggccc tagatccttg ttctgcttac atcagcctga atgagccctg	300
gaggaacact gaccaccagt tggatgagtc tcaaggctct cctctatgtg	350
acaaccatgt gaatggggag tgggtaccact tcacgggcat ggcgggagat	400
gccatgccta ccttctgcat accagaaaac cactgtggaa cccacgcacc	450
tgtctggctc aatggcagcc accccctaga aggcgacggc attgtgcaac	500
gccaggcttg tgccagcttc aatgggaact gctgtctctg gaacaccacg	550
gtggaagtca aggcttgccc tggaggctac tatgtgtatc gtctgaccaa	600
gccagcgctc tgcttccacg tctactgtgg tcatttttat gacatctgcg	650
acgaggactg ccatggcagc tgctcagata ccagcgagtg cacatgcgct	700
ccaggaactg tgctaggccc tgacaggcag acatgctttg atgaaaatga	750
atgtgagcaa aacaacggtg gctgcagtga gatctgtgtg aacctcaaaa	800
actcctaccg ctgtgagtgt ggggttggcc gtgtgctaag aagtgatggc	850
aagacttgtg aagacgttga aggatgccac aataacaatg gtggctgcag	900
ccactcttgc cttggatctg agaaaggcta ccagtgtgaa tgtccccggg	950
gcctggtgct gtctgaggat aaccacactt gccaagtccc tgtgttgtgc	1000
aaatcaaatg ccattgaagt gaacatcccc agggagctgg ttggtggcct	1050
ggagctcttc ctgaccaaca cctcctgccg aggagtgtcc aacggcacc	1100
atgtcaacat cctcttctct ctcaagacat gtggtacagt ggtcgatgtg	1150
gtgaatgaca agattgtggc cagcaacctc gtgacaggtc tacccaagca	1200
gaccccgggg agcagcgggg acttcatcat ccgaaccagc aagctgctga	1250
tcccggtgac ctgcgagttt ccacgcctgt acaccatttc tgaagatac	1300
gttcccaacc ttcgaaactc cccactggaa atcatgagcc gaaatcatgg	1350
gatcttccca ttcactctgg agatcttcaa ggacaatgag tttgaagagc	1400
cttaccggga agctctgccc accctcaagc ttcgtgactc cctctacttt	1450
ggcattgagc ccgtggtgca cgtgagcggc ttggaaagct tggtgagag	1500
ctgctttgcc accccacct ccaagatcga cgaggctctg aaatactacc	1550

-continued

```
tcatccggga tggctgtgtt tcagatgact cggtaaagca gtacacatcc      1600
cgggatcacc tagcaaagca cttccaggtc cctgtcttca agtttgtagg      1650
caaagaccac aaggaagtgt ttctgcactg ccgggttctt gtctgtggag      1700
tgttggacga gcgttccgc tgtgccagg gttgccaccg gcgaatgcgt      1750
cgtggggcag gaggagagga ctcagccggt ctacagggcc agacgctaac      1800
aggcggcccg atccgcacg actgggagga ctagtctgta gccatacctc      1850
gagtcctgc attggacggc tctgtctttt ggagcttctc cccccaccgc      1900
cctctaagaa catctgcaa cagctgggtt cagacttcac actgtgagtt      1950
cagactccca gcaccaactc actctgatto tggtcattc agtgggcaca      2000
ggtcacagca ctgctgaaca atgtggcctg ggtggggtt catctttcta      2050
gggttgaaaa ctaactgtc caccagaaa gacactcacc ccatttcct      2100
catttcttct ctacacttaa atacctgtg tatggtgcaa tcagaccaca      2150
aatcagaag ctgggtataa tatttcaagt taaaaccct agaaaaatta      2200
aacagt tact gaaattatga cttaaatacc caatgactcc ttaaatatgt      2250
aaattatagt tatacctga aatttcaatt caaatgcaga ctaattatag      2300
ggaatttgga agtgtatcaa taaaacagta tataat      2339
```

<210> SEQ ID NO 110
<211> LENGTH: 545
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 110

```
Met Pro Pro Phe Leu Leu Leu Thr Cys Leu Phe Ile Thr Gly Thr      15
 1          5          10
Ser Val Ser Pro Val Ala Leu Asp Pro Cys Ser Ala Tyr Ile Ser      30
          20          25
Leu Asn Glu Pro Trp Arg Asn Thr Asp His Gln Leu Asp Glu Ser      45
          35          40
Gln Gly Pro Pro Leu Cys Asp Asn His Val Asn Gly Glu Trp Tyr      60
          50          55
His Phe Thr Gly Met Ala Gly Asp Ala Met Pro Thr Phe Cys Ile      75
          65          70
Pro Glu Asn His Cys Gly Thr His Ala Pro Val Trp Leu Asn Gly      90
          80          85
Ser His Pro Leu Glu Gly Asp Gly Ile Val Gln Arg Gln Ala Cys      105
          95          100
Ala Ser Phe Asn Gly Asn Cys Cys Leu Trp Asn Thr Thr Val Glu      120
          110          115
Val Lys Ala Cys Pro Gly Gly Tyr Tyr Val Tyr Arg Leu Thr Lys      135
          125          130
Pro Ser Val Cys Phe His Val Tyr Cys Gly His Phe Tyr Asp Ile      150
          140          145
Cys Asp Glu Asp Cys His Gly Ser Cys Ser Asp Thr Ser Glu Cys      165
          155          160
Thr Cys Ala Pro Gly Thr Val Leu Gly Pro Asp Arg Gln Thr Cys      180
          170          175
```

-continued

Phe	Asp	Glu	Asn	Glu	Cys	Glu	Gln	Asn	Asn	Gly	Gly	Cys	Ser	Glu	185	190	195
Ile	Cys	Val	Asn	Leu	Lys	Asn	Ser	Tyr	Arg	Cys	Glu	Cys	Gly	Val	200	205	210
Gly	Arg	Val	Leu	Arg	Ser	Asp	Gly	Lys	Thr	Cys	Glu	Asp	Val	Glu	215	220	225
Gly	Cys	His	Asn	Asn	Asn	Gly	Gly	Cys	Ser	His	Ser	Cys	Leu	Gly	230	235	240
Ser	Glu	Lys	Gly	Tyr	Gln	Cys	Glu	Cys	Pro	Arg	Gly	Leu	Val	Leu	245	250	255
Ser	Glu	Asp	Asn	His	Thr	Cys	Gln	Val	Pro	Val	Leu	Cys	Lys	Ser	260	265	270
Asn	Ala	Ile	Glu	Val	Asn	Ile	Pro	Arg	Glu	Leu	Val	Gly	Gly	Leu	275	280	285
Glu	Leu	Phe	Leu	Thr	Asn	Thr	Ser	Cys	Arg	Gly	Val	Ser	Asn	Gly	290	295	300
Thr	His	Val	Asn	Ile	Leu	Phe	Ser	Leu	Lys	Thr	Cys	Gly	Thr	Val	305	310	315
Val	Asp	Val	Val	Asn	Asp	Lys	Ile	Val	Ala	Ser	Asn	Leu	Val	Thr	320	325	330
Gly	Leu	Pro	Lys	Gln	Thr	Pro	Gly	Ser	Ser	Gly	Asp	Phe	Ile	Ile	335	340	345
Arg	Thr	Ser	Lys	Leu	Leu	Ile	Pro	Val	Thr	Cys	Glu	Phe	Pro	Arg	350	355	360
Leu	Tyr	Thr	Ile	Ser	Glu	Gly	Tyr	Val	Pro	Asn	Leu	Arg	Asn	Ser	365	370	375
Pro	Leu	Glu	Ile	Met	Ser	Arg	Asn	His	Gly	Ile	Phe	Pro	Phe	Thr	380	385	390
Leu	Glu	Ile	Phe	Lys	Asp	Asn	Glu	Phe	Glu	Glu	Pro	Tyr	Arg	Glu	395	400	405
Ala	Leu	Pro	Thr	Leu	Lys	Leu	Arg	Asp	Ser	Leu	Tyr	Phe	Gly	Ile	410	415	420
Glu	Pro	Val	Val	His	Val	Ser	Gly	Leu	Glu	Ser	Leu	Val	Glu	Ser	425	430	435
Cys	Phe	Ala	Thr	Pro	Thr	Ser	Lys	Ile	Asp	Glu	Val	Leu	Lys	Tyr	440	445	450
Tyr	Leu	Ile	Arg	Asp	Gly	Cys	Val	Ser	Asp	Asp	Ser	Val	Lys	Gln	455	460	465
Tyr	Thr	Ser	Arg	Asp	His	Leu	Ala	Lys	His	Phe	Gln	Val	Pro	Val	470	475	480
Phe	Lys	Phe	Val	Gly	Lys	Asp	His	Lys	Glu	Val	Phe	Leu	His	Cys	485	490	495
Arg	Val	Leu	Val	Cys	Gly	Val	Leu	Asp	Glu	Arg	Ser	Arg	Cys	Ala	500	505	510
Gln	Gly	Cys	His	Arg	Arg	Met	Arg	Arg	Gly	Ala	Gly	Gly	Glu	Asp	515	520	525
Ser	Ala	Gly	Leu	Gln	Gly	Gln	Thr	Leu	Thr	Gly	Gly	Pro	Ile	Arg	530	535	540
Ile	Asp	Trp	Glu	Asp											545		

-continued

<210> SEQ ID NO 111

<211> LENGTH: 2063

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 111

gagagaggca gcagcttgct cagcggacaa ggatgctggg cgtgagggac	50
caaggcctgc cctgcactcg ggcctcctcc agccagtgct gaccagggac	100
ttctgacctg ctggccagcc aggacctgtg tggggaggcc ctctgtctgc	150
cttgggggtga caatctcagc tccaggctac agggagaccg ggaggatcac	200
agagccagca tggtacagga tcctgacagt gatcaacctc tgaacagcct	250
cgatgtcaaa cccctgcgca aaccctgat ccccatggag accttcagaa	300
aggtggggat ccccatcatc atagcactac tgagcctggc gagtatcatc	350
attgtggttg tcctcatcaa ggtgattctg gataaatact acttcctctg	400
cgggcagcct ctccacttca tcccgaggaa gcagctgtgt gacggagagc	450
tggactgtcc cttgggggag gacgaggagc actgtgtcaa gagcttcccc	500
gaagggcctg cagtggcagt ccgcctctcc aaggaccgat ccacactgca	550
ggtgtctggc tcggccacag ggaactggtt ctctgcctgt ttcgacaact	600
tcacagaagc tctcgctgag acagcctgta ggcagatggg ctacagcaga	650
gctgtggaga ttggcccaga ccaggatctg gatgttgttg aaatcacaga	700
aaacagccag gagcttcgca tgcggaactc aagtgggccc tgtctctcag	750
gtcccttggt ctccctgcac tgtcttgctt gtgggaagag cctgaagacc	800
ccccgtgtgg tgggtgggga ggaggcctct gtggattctt ggccttgga	850
ggtcagcatc cagtacgaca aacagcacgt ctgtggaggg agcatcctgg	900
acccccactg ggtcctcacg gcagccact gcttcaggaa acataccgat	950
gtgttcaact ggaagggtcg ggcaggctca gacaaactgg gcagcttccc	1000
atccctggct gtggccaaga tcatcatcat tgaattcaac cccatgtacc	1050
ccaaagacaa tgacatcgcc ctcatgaagc tgcagttccc actcactttc	1100
tcaggcacag tcaggcccat ctgtctgccc ttctttgatg aggagctcac	1150
tccagccacc ccactctgga tcattggatg gggctttacg aagcagaatg	1200
gaggaagat gtctgacata ctgctgcagg cgtcagtcca ggtcattgac	1250
agcacacggt gcaatgcaga cgatgcgtac cagggggaag tcaccgagaa	1300
gatgatgtgt gcaggcatcc cggaaggggg tgtggacacc tgccaggggtg	1350
acagtgggtg gccctgatg taccaatctg accagtggca tgtgggtggc	1400
atcgtagctt ggggctatgg ctgcgggggc ccgagcacc caggagtata	1450
caccaaggtc tcagcctatc tcaactggat ctacaatgtc tggaaggctg	1500
agctgtaatg ctgctgcccc ttgacgtgc tgggagccgc ttccttctg	1550
ccctgcccac ctggggatcc cccaaagtca gacacagagc aagagtcccc	1600
ttgggtacac ccctctgccc acagcctcag catttcttgg agcagcaaa	1650
ggcctcaatt cctgtaagag accctcgag cccagaggcg cccagaggaa	1700

-continued

```

gtcagcagcc ctactctggc cacacttggt gctcccagca tcccaggagg      1750
agacacagcc cactgaacaa ggtctcaggg gtattgctaa gccaagaagg      1800
aactttccca cactactgaa tggaagcagg ctgtcttgta aaagcccaga      1850
tcactgtggg ctggagagga gaaggaaagg gtctgcgcca gccctgtccg      1900
tcttcaccca tccccaagcc tactagagca agaaccagt tgtaataataa      1950
aatgcactgc cctactgttg gtatgactac cgttacctac tgttgtcatt      2000
gttattacag ctatggccac tattattaaa gagctgtgta acatctctgg      2050
caaaaaaaaa aaa                                              2063

```

<210> SEQ ID NO 112

<211> LENGTH: 432

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 112

```

Met Leu Gln Asp Pro Asp Ser Asp Gln Pro Leu Asn Ser Leu Asp
 1             5             10             15
Val Lys Pro Leu Arg Lys Pro Arg Ile Pro Met Glu Thr Phe Arg
 20            25            30
Lys Val Gly Ile Pro Ile Ile Ile Ala Leu Leu Ser Leu Ala Ser
 35            40            45
Ile Ile Ile Val Val Val Leu Ile Lys Val Ile Leu Asp Lys Tyr
 50            55            60
Tyr Phe Leu Cys Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln
 65            70            75
Leu Cys Asp Gly Glu Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu
 80            85            90
His Cys Val Lys Ser Phe Pro Glu Gly Pro Ala Val Ala Val Arg
 95           100           105
Leu Ser Lys Asp Arg Ser Thr Leu Gln Val Leu Asp Ser Ala Thr
110           115           120
Gly Asn Trp Phe Ser Ala Cys Phe Asp Asn Phe Thr Glu Ala Leu
125           130           135
Ala Glu Thr Ala Cys Arg Gln Met Gly Tyr Ser Arg Ala Val Glu
140           145           150
Ile Gly Pro Asp Gln Asp Leu Asp Val Val Glu Ile Thr Glu Asn
155           160           165
Ser Gln Glu Leu Arg Met Arg Asn Ser Ser Gly Pro Cys Leu Ser
170           175           180
Gly Ser Leu Val Ser Leu His Cys Leu Ala Cys Gly Lys Ser Leu
185           190           195
Lys Thr Pro Arg Val Val Gly Gly Glu Glu Ala Ser Val Asp Ser
200           205           210
Trp Pro Trp Gln Val Ser Ile Gln Tyr Asp Lys Gln His Val Cys
215           220           225
Gly Gly Ser Ile Leu Asp Pro His Trp Val Leu Thr Ala Ala His
230           235           240
Cys Phe Arg Lys His Thr Asp Val Phe Asn Trp Lys Val Arg Ala
245           250           255
Gly Ser Asp Lys Leu Gly Ser Phe Pro Ser Leu Ala Val Ala Lys

```

-continued

260	265	270
Ile Ile Ile Ile Glu Phe Asn Pro Met Tyr Pro Lys Asp Asn Asp		
275	280	285
Ile Ala Leu Met Lys Leu Gln Phe Pro Leu Thr Phe Ser Gly Thr		
290	295	300
Val Arg Pro Ile Cys Leu Pro Phe Phe Asp Glu Glu Leu Thr Pro		
305	310	315
Ala Thr Pro Leu Trp Ile Ile Gly Trp Gly Phe Thr Lys Gln Asn		
320	325	330
Gly Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser Val Gln Val		
335	340	345
Ile Asp Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln Gly Glu		
350	355	360
Val Thr Glu Lys Met Met Cys Ala Gly Ile Pro Glu Gly Gly Val		
365	370	375
Asp Thr Cys Gln Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser		
380	385	390
Asp Gln Trp His Val Val Gly Ile Val Ser Trp Gly Tyr Gly Cys		
395	400	405
Gly Gly Pro Ser Thr Pro Gly Val Tyr Thr Lys Val Ser Ala Tyr		
410	415	420
Leu Asn Trp Ile Tyr Asn Val Trp Lys Ala Glu Leu		
425	430	

<210> SEQ ID NO 113
<211> LENGTH: 1768
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 113

ggctggactg gaactcctgg tcccaagtga tccacccgcc tcagcctccc	50
aaggtgctgt gattataggt gtaagccacc gtgtctggcc tctgaacaac	100
tttttcagca actaaaaaag ccacaggagt tgaactgcta ggattctgac	150
tatgctgtgg tggctagtgc tcctactcct acctacatta aaatctgttt	200
tttgttctct tgtaactagc ctttaccttc ctaacacaga ggatctgtca	250
ctgtggctct ggcccaaacc tgaccttcac tctggaacga gaacagaggt	300
ttctaccac accgtcccct cgaagccggg gacagcctca cttgctggc	350
ctctcgctgg agcagtgcc tcaccaactg tctcacgtct ggaggcactg	400
actcgggcag tgcaggtagc tgagcctctt ggtagctgcg gctttcaagg	450
tgggccttgc cctggccgta gaagggattg acaagcccga agatttcata	500
ggcgatggct cccactgccc aggcacagc cttgctgtag tcaatcactg	550
ccctggggcc aggacgggcc gtggacacct gctcagaagc agtgggtgag	600
acatcacgct gcccgcccat ctaacctttt catgtcctgc acatcacctg	650
atccatgggc taatctgaac tctgtcccaa ggaaccaga gcttgagtga	700
gctgtggctc agaccagaa ggggtctgct tagaccacct ggtttatgtg	750
acaggacttg cattctcctg gaacatgagg gaacgccga ggaaagcaaa	800
gtggcaggga aggaacttgt gccaaattat gggtcagaaa agatggaggt	850

-continued

gttgggttat cacaaggcat cgagtctcct gcattcagtg gacatgtggg	900
ggaagggtctg ccgatggcgc atgacacact cgggactcac ctctggggcc	950
atcacagacgc cgtttccgcc ccgatccacg taccagctgc tgaagggcaa	1000
ctgcaggcgc atgctctcat cagccaggca gcagccaaaa tctgcgatca	1050
ccagccaggg gcagccgtct gggaaggagc aagcaaagt accatttctc	1100
ctccctcct tccctctgag aggccctcct atgtccctac taaagccacc	1150
agcaagacat agctgacagg ggctaattgg tcagtgttgg ccaggagggt	1200
cagcaaggcc tgagagctga tcagaagggc ctgctgtgcg aacacggaaa	1250
tgcctccagt aagcacaggc tgcaaatcc ccaggcaaag gactgtgtgg	1300
ctcaatttaa atcatgttct agtaattgga gctgtcccca agaccaaagg	1350
agctagagct tggttcaaat gatctccaag ggcccttata cccaggaga	1400
ctttgatttg aatttgaac cccaatcca aacctaagaa ccagggtgcat	1450
taagaatcag ttattgccgg gtgtggtggc ctgtaatgcc aacattttgg	1500
gaggccgagg cgggtagatc acctgaggtc aggagttcaa gaccagcctg	1550
gccaacatgg tgaaaccct gtcttacta aaaatacaaa aaaactagcc	1600
aggcatggtg gtgtgtgcct gtatcccagc tactcgggag gctgagacag	1650
gagaattact tgaacctggg aggtgaagga ggctgagaca ggagaatcac	1700
ttcagcctga gcaacacagc gagactctgt ctcagaaaaa ataaaaaaag	1750
aattatgggtt atttgtaa	1768

<210> SEQ ID NO 114
<211> LENGTH: 109
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 114

Met Leu Trp Trp Leu Val Leu Leu Leu Pro Thr Leu Lys Ser	
1 5 10 15	
Val Phe Cys Ser Leu Val Thr Ser Leu Tyr Leu Pro Asn Thr Glu	
20 25 30	
Asp Leu Ser Leu Trp Leu Trp Pro Lys Pro Asp Leu His Ser Gly	
35 40 45	
Thr Arg Thr Glu Val Ser Thr His Thr Val Pro Ser Lys Pro Gly	
50 55 60	
Thr Ala Ser Pro Cys Trp Pro Leu Ala Gly Ala Val Pro Ser Pro	
65 70 75	
Thr Val Ser Arg Leu Glu Ala Leu Thr Arg Ala Val Gln Val Ala	
80 85 90	
Glu Pro Leu Gly Ser Cys Gly Phe Gln Gly Gly Pro Cys Pro Gly	
95 100 105	
Arg Arg Arg Asp	

<210> SEQ ID NO 115
<211> LENGTH: 1197
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 115

```
cagcagtggg ctctcagtc tctcaaagca aggaaagagt actgtgtgct      50
gagagaccat ggcaaagaat cctccagaga attgtgaaga ctgtcacatt      100
ctaaatgcag aagcttttaa atccaagaaa atatgtaaat cacttaagat      150
ttgtggactg gtgtttggta tcctggccct aactctaatt gtcctgtttt      200
gggggagcaa gcacttctgg ccggaggtac ccaaaaaagc ctatgacatg      250
gagcacactt tctacagcaa tggagagaag aagaagattt acatggaaat      300
tgatcctgtg accagaactg aaatatctag aagcggaaat ggactgatg      350
aaacattgga agtgcacgac tttaaaaacg gatacactgg catctacttc      400
gtgggtcttc aaaaatgttt tatcaaaact cagattaaag tgattcctga      450
atthttctgaa ccagaagagg aaatagatga gaatgaagaa attaccacaa      500
ctttctttga acagtcatg atttgggtcc cagcagaaaa gcctattgaa      550
aaccgagatt ttcttaaaaa ttccaaaatt ctggagattt gtgataacgt      600
gaccatgtat tggatcaatc ccactcta atcagtttct gagttacaag      650
actttgagga ggaggagaa gatcttctact ttctgcca cgaaaaaaaaa      700
gggattgaac aaaaatgaaca gtgggtggtc cctcaagtga aagtagagaa      750
gaccctgcac gccagacaag caagtgagga agaacttcca ataaatgact      800
atactgaaaa tggaaatagaa ttgatccca tgctggatga gagaggttat      850
tgttgtatth actgccgtcg aggcaaccgc tattgccgcc gcgtctgtga      900
acctttacta ggctactacc catatccata ctgtaccaa ggaggacgag      950
tcatctgtcg tgtcatcatg ccttgtaact ggtgggtggc ccgcatgctg     1000
gggaggggtc aataggaggt ttgagctcaa atgcttaaac tgctggcaac     1050
atataataaa tgcatgctat tcaatgaatt tctgcctatg aggcactctg     1100
cccctggtag ccagctctcc agaattactt gtaggtaatt cctctcttca     1150
tgttctaata aacttctaca ttatcaccaa aaaaaaaaaa aaaaaaa      1197
```

<210> SEQ ID NO 116

<211> LENGTH: 317

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 116

```
Met Ala Lys Asn Pro Pro Glu Asn Cys Glu Asp Cys His Ile Leu
 1             5             10            15
Asn Ala Glu Ala Phe Lys Ser Lys Lys Ile Cys Lys Ser Leu Lys
20            25            30
Ile Cys Gly Leu Val Phe Gly Ile Leu Ala Leu Thr Leu Ile Val
35            40            45
Leu Phe Trp Gly Ser Lys His Phe Trp Pro Glu Val Pro Lys Lys
50            55            60
Ala Tyr Asp Met Glu His Thr Phe Tyr Ser Asn Gly Glu Lys Lys
65            70            75
Lys Ile Tyr Met Glu Ile Asp Pro Val Thr Arg Thr Glu Ile Phe
80            85            90
```

-continued

Arg	Ser	Gly	Asn	Gly	Thr	Asp	Glu	Thr	Leu	Glu	Val	His	Asp	Phe
			95						100					105
Lys	Asn	Gly	Tyr	Thr	Gly	Ile	Tyr	Phe	Val	Gly	Leu	Gln	Lys	Cys
			110						115					120
Phe	Ile	Lys	Thr	Gln	Ile	Lys	Val	Ile	Pro	Glu	Phe	Ser	Glu	Pro
			125						130					135
Glu	Glu	Glu	Ile	Asp	Glu	Asn	Glu	Glu	Ile	Thr	Thr	Thr	Phe	Phe
			140						145					150
Glu	Gln	Ser	Val	Ile	Trp	Val	Pro	Ala	Glu	Lys	Pro	Ile	Glu	Asn
			155						160					165
Arg	Asp	Phe	Leu	Lys	Asn	Ser	Lys	Ile	Leu	Glu	Ile	Cys	Asp	Asn
			170						175					180
Val	Thr	Met	Tyr	Trp	Ile	Asn	Pro	Thr	Leu	Ile	Ser	Val	Ser	Glu
			185						190					195
Leu	Gln	Asp	Phe	Glu	Glu	Glu	Gly	Glu	Asp	Leu	His	Phe	Pro	Ala
			200						205					210
Asn	Glu	Lys	Lys	Gly	Ile	Glu	Gln	Asn	Glu	Gln	Trp	Val	Val	Pro
			215						220					225
Gln	Val	Lys	Val	Glu	Lys	Thr	Arg	His	Ala	Arg	Gln	Ala	Ser	Glu
			230						235					240
Glu	Glu	Leu	Pro	Ile	Asn	Asp	Tyr	Thr	Glu	Asn	Gly	Ile	Glu	Phe
			245						250					255
Asp	Pro	Met	Leu	Asp	Glu	Arg	Gly	Tyr	Cys	Cys	Ile	Tyr	Cys	Arg
			260						265					270
Arg	Gly	Asn	Arg	Tyr	Cys	Arg	Arg	Val	Cys	Glu	Pro	Leu	Leu	Gly
			275						280					285
Tyr	Tyr	Pro	Tyr	Pro	Tyr	Cys	Tyr	Gln	Gly	Gly	Arg	Val	Ile	Cys
			290						295					300
Arg	Val	Ile	Met	Pro	Cys	Asn	Trp	Trp	Val	Ala	Arg	Met	Leu	Gly
			305						310					315

Arg Val

<210> SEQ ID NO 117

<211> LENGTH: 2121

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 117

gagctcccct caggagcgcg ttagcttcac accttcggca gcaggagggc	50
ggcagcttct cgcaggcggc agggcggcg gccaggatca tgtccaccac	100
cacatgccaa gtgttgccgt tcctcctgtc catcctgggg ctggccggct	150
gcacgcggc caccgggatg gacatgtgga gcaccagga cctgtacgac	200
aaccccgta cctccgtgtt ccagtacgaa gggctctgga ggagctgcgt	250
gaggcagagt tcaggcttca ccgaatgcag gccctatttc accatcctgg	300
gacttcacgc catgctgcag gcagtgcgag ccctgatgat cgtaggcac	350
gtcctgggtg ccattggcct cctggtatcc atctttgccc tgaatgcac	400
ccgcattggc agcatggagg actctgccaa agccaacatg aactgacct	450
ccgggatcat gttcattgtc tcaggtcttt gtgcaattgc tggagtgtct	500
gtgtttgccaa acatgctggt gactaacttc tggatgtcca cagctaacat	550

-continued

gtacaccggc atgggtggga tgggtgcagac tgttcagacc aggtacacat	600
ttggtgcggc tctgttcgtg ggctgggtcg ctggaggcct cacactaatt	650
gggggtgtga tgatgtgcat cgctgccgg ggcctggcac cagaagaaac	700
caactacaaa gccgtttctt atcatgcctc aggccacagt gttgcctaca	750
agcctggagg cttcaaggcc agcactggct ttgggtccaa caccaaaaac	800
aagaagatat acgatggagg tgcccgcaca gaggacgagg tacaatctta	850
tccttccaa gacgactatg tgtaatgctc taagacctct cagcacgggc	900
ggaagaaact cccggagagc tcacccaaaa aacaaggaga tcccatctag	950
atcttctctt gcttttgact cacagctgga agttagaaaa gcctcgattt	1000
catcttttga gagccaaat ggtcttagcc tcagtctctg tctctaataa	1050
ttccaccata aaacagctga gttatattatg aattagaggc tatagctcac	1100
atcttcaatc ctctatttct ttttttaaat ataactttct actctgatga	1150
gagaatgtgg ttttaatctc tctctcacat tttgatgatt tagacagact	1200
ccccctcttc ctctagtca ataaacccat tgatgatcta tttccagct	1250
tatccccaa gaaaactttt aaaggaaaga gtagaccaa agatgttatt	1300
ttctgctgtt tgaattttgt ctcccacccc ccaacttggc tagtaataaa	1350
cacttactga agaagaagca ataagagaaa gatatttgta atctctccag	1400
cccatgatct cggttttctt acactgtgat cttaaaagt accaaaccaa	1450
agtcattttc agtttgaggc aaccaaaccct ttctactgct gttgacatct	1500
tcttattaca gcaacacccat tctaggagtt tctgagctc tccactggag	1550
tcctctttct gtcgcggtc agaaattgtc cctagatgaa tgagaaaatt	1600
atctttttta atttaagtcc taaatatagt taaaaataat aatgttttag	1650
taaaatgata cactatctct gtgaaatagc ctacccccta catgtggata	1700
gaaggaaatg aaaaaataat tgctttgaca ttgtctatat ggtactttgt	1750
aaagtcatgc ttaagtacaa attccatgaa aagctcacac ctgtaatcct	1800
agcacttttg gaggtgagg aggaaggatc acttgagccc agaagttcga	1850
gactagcctg ggcaacatgg agaagccctg tctctacaaa atacagagag	1900
aaaaaatcag ccagtcatgg tggcatacac ctgtagtccc agcattccgg	1950
gaggctgagg tgggaggatc acttgagccc agggagggtt gggctgcagt	2000
gagccatgat cacaccactg cactccagcc aggtgacata gcgagatcct	2050
gtctaaaaaa ataaaaata aataatggaa cacagcaagt cctaggaagt	2100
aggttaaaac taattcttta a	2121

<210> SEQ ID NO 118
<211> LENGTH: 261
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 118

Met Ser Thr Thr Thr Cys Gln Val Val Ala Phe Leu Leu Ser Ile
1 5 10 15

-continued

Leu	Gly	Leu	Ala	Gly	Cys	Ile	Ala	Ala	Thr	Gly	Met	Asp	Met	Trp	
				20					25					30	
Ser	Thr	Gln	Asp	Leu	Tyr	Asp	Asn	Pro	Val	Thr	Ser	Val	Phe	Gln	
				35					40					45	
Tyr	Glu	Gly	Leu	Trp	Arg	Ser	Cys	Val	Arg	Gln	Ser	Ser	Gly	Phe	
				50					55					60	
Thr	Glu	Cys	Arg	Pro	Tyr	Phe	Thr	Ile	Leu	Gly	Leu	Pro	Ala	Met	
				65					70					75	
Leu	Gln	Ala	Val	Arg	Ala	Leu	Met	Ile	Val	Gly	Ile	Val	Leu	Gly	
				80					85					90	
Ala	Ile	Gly	Leu	Leu	Val	Ser	Ile	Phe	Ala	Leu	Lys	Cys	Ile	Arg	
				95					100					105	
Ile	Gly	Ser	Met	Glu	Asp	Ser	Ala	Lys	Ala	Asn	Met	Thr	Leu	Thr	
				110					115					120	
Ser	Gly	Ile	Met	Phe	Ile	Val	Ser	Gly	Leu	Cys	Ala	Ile	Ala	Gly	
				125					130					135	
Val	Ser	Val	Phe	Ala	Asn	Met	Leu	Val	Thr	Asn	Phe	Trp	Met	Ser	
				140					145					150	
Thr	Ala	Asn	Met	Tyr	Thr	Gly	Met	Gly	Gly	Met	Val	Gln	Thr	Val	
				155					160					165	
Gln	Thr	Arg	Tyr	Thr	Phe	Gly	Ala	Ala	Leu	Phe	Val	Gly	Trp	Val	
				170					175					180	
Ala	Gly	Gly	Leu	Thr	Leu	Ile	Gly	Gly	Val	Met	Met	Cys	Ile	Ala	
				185					190					195	
Cys	Arg	Gly	Leu	Ala	Pro	Glu	Glu	Thr	Asn	Tyr	Lys	Ala	Val	Ser	
				200					205					210	
Tyr	His	Ala	Ser	Gly	His	Ser	Val	Ala	Tyr	Lys	Pro	Gly	Gly	Phe	
				215					220					225	
Lys	Ala	Ser	Thr	Gly	Phe	Gly	Ser	Asn	Thr	Lys	Asn	Lys	Lys	Ile	
				230					235					240	
Tyr	Asp	Gly	Gly	Ala	Arg	Thr	Glu	Asp	Glu	Val	Gln	Ser	Tyr	Pro	
				245					250					255	
Ser	Lys	His	Asp	Tyr	Val										
				260											

<210> SEQ ID NO 119

<211> LENGTH: 2010

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 119

ggaaaaactg ttctcttctg tggcacagag aacctgctt caaagcagaa	50
gtagcagttc cggagtccag ctggctaaaa ctcattccag aggataatgg	100
caacccatgc ctagaaaatc gctgggctgt ttcttggtgg tgttggaatg	150
gtgggcacag tggctgtcac tgtcatgcct cagtggagag tgtcgccctt	200
cattgaaaac aacatcgtgg tttttgaaaa cttctgggaa ggactgtgga	250
tgaattgcgt gaggcaggct aacatcagga tgcagtgcaa aatctatgat	300
ttcctgctgg ctctttctcc ggacctacag gcagccagag gactgatgtg	350
tgctgcttcc gtgatgtcct tcttggtttt catgatggcc atccttgcca	400
tgaaatgcac caggtgcacg ggggacaatg agaaggtgaa ggctcacatt	450

-continued

ctgctgacgg ctggaatcat cttcatcatc acgggcatgg tgggtgctcat	500
ccctgtgagc tgggttgcca atgccatcat cagagatttc tataactcaa	550
tagtgaatgt tgcccaaaaa cgtgagcttg gagaagctct ctacttagga	600
tggaccacgg cactggtgct gattgttga ggagctctgt tctgctgcgt	650
tttttgttg c aacgaaaaga gcagtagcta cagatactcg ataccttccc	700
atcgcacaa ccaaaaaagt tatcacaccg gaaagaagtc accgagcgtc	750
tactccagaa gtcagtatgt gtagttgtgt atgttttttt aactttacta	800
taaagccatg caaatgacaa aaatctatat tactttctca aaatggaccc	850
caaagaaact ttgatttact gttcttaact gcctaactctt aattacagga	900
actgtgcac agctattttat gattctataa gctatttcag cagaatgaga	950
tattaaaccc aatgctttga ttgttctaga aagtatagta atttgttttc	1000
taaggtgggt caagcatcta ctctttttat catttacttc aaaatgacat	1050
tgctaaagac tgcattattt tactactgta atttctccac gacatagcat	1100
tatgtacata gatgagtgtg acatttatat ctacacataga gacatgctta	1150
tatgggttta tttaaaatga aatgccagtc cattacactg aataaataga	1200
actcaactat tgcttttcag ggaaatcatg gatagggttg aagaaggtta	1250
ctattaattg tttaaaaaca gcttagggat taatgtcctc cattttataat	1300
gaagattaaa atgaaggctt taatcagcat tgtaaaggaa attgaatggc	1350
tttctgatat gctgtttttt agcctaggag ttagaaatcc taacttcott	1400
atcctcttct cccagaggct ttttttttct tgtgtattaa attaacattt	1450
ttaaaacgca gatattttgt caaggggctt tgcattcaaa ctgcttttcc	1500
agggtatata tcagaagaaa gataaaagtg tgatctaaga aaaagtgatg	1550
gttttaggaa agtgaaaata tttttgtttt tgtatttgaa gaagaatgat	1600
gcattttgac aagaaatcat atatgtatgg atatatttta ataagtattt	1650
gagtacagac tttgaggttt catcaatata aataaaagag cagaaaaata	1700
tgtcttggtt ttcatttgct taccaaaaaa acaacaacaa aaaaagttgt	1750
cctttgagaa cttcacctgc tcctatgtgg gtacctgagt caaaattgtc	1800
atttttgttc tgtgaaaaat aaatttcctt cttgtacat tctgttttag	1850
ttttactaaa atctgtaaat actgtatttt tctgtttatt ccaaatttga	1900
tgaaactgac aatccaattt gaaagtttgt gtcgacgtct gtctagctta	1950
aatgaatgtg ttctatttgc ttatataatt tatattaata aattgtacat	2000
ttttctaatt	2010

<210> SEQ ID NO 120
<211> LENGTH: 225
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 120

Met Ala Thr His Ala Leu Glu Ile Ala Gly Leu Phe Leu Gly Gly
1 5 10 15

-continued

Val	Gly	Met	Val	Gly	Thr	Val	Ala	Val	Thr	Val	Met	Pro	Gln	Trp
			20						25					30
Arg	Val	Ser	Ala	Phe	Ile	Glu	Asn	Asn	Ile	Val	Val	Phe	Glu	Asn
			35						40					45
Phe	Trp	Glu	Gly	Leu	Trp	Met	Asn	Cys	Val	Arg	Gln	Ala	Asn	Ile
			50						55					60
Arg	Met	Gln	Cys	Lys	Ile	Tyr	Asp	Ser	Leu	Leu	Ala	Leu	Ser	Pro
			65						70					75
Asp	Leu	Gln	Ala	Ala	Arg	Gly	Leu	Met	Cys	Ala	Ala	Ser	Val	Met
			80						85					90
Ser	Phe	Leu	Ala	Phe	Met	Met	Ala	Ile	Leu	Gly	Met	Lys	Cys	Thr
			95						100					105
Arg	Cys	Thr	Gly	Asp	Asn	Glu	Lys	Val	Lys	Ala	His	Ile	Leu	Leu
			110						115					120
Thr	Ala	Gly	Ile	Ile	Phe	Ile	Ile	Thr	Gly	Met	Val	Val	Leu	Ile
			125						130					135
Pro	Val	Ser	Trp	Val	Ala	Asn	Ala	Ile	Ile	Arg	Asp	Phe	Tyr	Asn
			140						145					150
Ser	Ile	Val	Asn	Val	Ala	Gln	Lys	Arg	Glu	Leu	Gly	Glu	Ala	Leu
			155						160					165
Tyr	Leu	Gly	Trp	Thr	Thr	Ala	Leu	Val	Leu	Ile	Val	Gly	Gly	Ala
			170						175					180
Leu	Phe	Cys	Cys	Val	Phe	Cys	Cys	Asn	Glu	Lys	Ser	Ser	Ser	Tyr
			185						190					195
Arg	Tyr	Ser	Ile	Pro	Ser	His	Arg	Thr	Thr	Gln	Lys	Ser	Tyr	His
			200						205					210
Thr	Gly	Lys	Lys	Ser	Pro	Ser	Val	Tyr	Ser	Arg	Ser	Gln	Tyr	Val
			215						220					225

<210> SEQ ID NO 121
<211> LENGTH: 1257
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 121

ggagagaggc gcgcggtga aaggcgcat gatgcagcct gcggcggcct	50
cgagagcgcg cgagccaga cgctgaccac gttcctctcc tcggtctcct	100
ccgcctccag ctccgcgctg ccggcgagcc gggagccatg cgaccccagg	150
gccccgccgc ctccccgcag cggctccgcg gcctcctgct gctcctgctg	200
ctgcagctgc ccgcgccgtc gagcgccctc gagatcccca aggggaagca	250
aaaggcgcag ctccggcaga gggaggtggt ggacctgtat aatggaatgt	300
gcttacaagg gccagcagga gtgcctggtc gagacgggag ccctggggcc	350
aatgttattc cgggtacacc tgggatccca ggtcgggatg gattcaaagg	400
agaaaagggg gaatgtctga gggaaagctt tgaggagtcc tggacacca	450
actacaagca gtgttcattg agttcattga attatggcat agatcttggg	500
aaaattgctg agtgtacatt tacaaagatg cgttcaaata gtgctctaag	550
agttttgttc agtggtctac ttcggctaaa atgcagaaat gcatgctgtc	600
agcgttggtg ttccacattc aatggagctg aatgttcagg acctcttccc	650

-continued

attgaagcta taatttat	700
ttt ggaccaagga agccctgaaa tgaattcaac	
aattaatatt catcgactt cttctgtgga aggactttgt gaaggaattg	750
gtgctggatt agtggatggt gctatctggg ttggcacttg ttcagattac	800
ccaaaaggag atgcttctac tggatggaat tcagtttctc gcatcattat	850
tgaagaacta ccaaaataaa tgctttaatt ttcatttgct acctcttttt	900
ttattatgcc ttggaatggt tcacttaaat gacattttaa ataagtttat	950
gtatacatct gaatgaaaag caaagctaaa tatgtttaca gaccaaagtg	1000
tgatttcaca ctgtttttta atctagcatt attcattttg cttcaatcaa	1050
aagtggtttc aatatttttt ttagttggtt agaatacttt cttcatagtc	1100
acattctctc aacctataat ttggaatatt gttgtgtct tttgttttt	1150
ctcttagtat agcattttta aaaaaatata aaagctacca atctttgtac	1200
aatttgtaaa tgttaagaat tttttttata tctgttaa	1250
at aaaaattatt	
tccaaca	1257

<210> SEQ ID NO 122
<211> LENGTH: 243
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 122

Met Arg Pro Gln Gly Pro Ala Ala Ser Pro Gln Arg Leu Arg Gly	
1 5 10 15	
Leu Leu Leu Leu Leu Leu Gln Leu Pro Ala Pro Ser Ser Ala	
20 25 30	
Ser Glu Ile Pro Lys Gly Lys Gln Lys Ala Gln Leu Arg Gln Arg	
35 40 45	
Glu Val Val Asp Leu Tyr Asn Gly Met Cys Leu Gln Gly Pro Ala	
50 55 60	
Gly Val Pro Gly Arg Asp Gly Ser Pro Gly Ala Asn Val Ile Pro	
65 70 75	
Gly Thr Pro Gly Ile Pro Gly Arg Asp Gly Phe Lys Gly Glu Lys	
80 85 90	
Gly Glu Cys Leu Arg Glu Ser Phe Glu Glu Ser Trp Thr Pro Asn	
95 100 105	
Tyr Lys Gln Cys Ser Trp Ser Ser Leu Asn Tyr Gly Ile Asp Leu	
110 115 120	
Gly Lys Ile Ala Glu Cys Thr Phe Thr Lys Met Arg Ser Asn Ser	
125 130 135	
Ala Leu Arg Val Leu Phe Ser Gly Ser Leu Arg Leu Lys Cys Arg	
140 145 150	
Asn Ala Cys Cys Gln Arg Trp Tyr Phe Thr Phe Asn Gly Ala Glu	
155 160 165	
Cys Ser Gly Pro Leu Pro Ile Glu Ala Ile Ile Tyr Leu Asp Gln	
170 175 180	
Gly Ser Pro Glu Met Asn Ser Thr Ile Asn Ile His Arg Thr Ser	
185 190 195	
Ser Val Glu Gly Leu Cys Glu Gly Ile Gly Ala Gly Leu Val Asp	
200 205 210	

-continued

Val Ala Ile Trp Val Gly Thr Cys Ser Asp Tyr Pro Lys Gly Asp
215 220 225
Ala Ser Thr Gly Trp Asn Ser Val Ser Arg Ile Ile Ile Glu Glu
230 235 240
Leu Pro Lys

<210> SEQ ID NO 123
<211> LENGTH: 2379
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 123

gctgagcgtg tgcgcggtac ggggctctcc tgccttctgg gctccaacgc	50
agctctgttg ctgaactggg tgctcatcac gggaaactgct gggctatgga	100
atacagatgt ggcagctcag gtagcccaa attgcctgga agaatacatc	150
atgtttttcg ataagaagaa attgtaggat ccagtttttt ttttaaccgc	200
ccccctccca cccccaaaa aaactgtaaa gatgcaaaaa cgtaatatcc	250
atgaagatcc tattacctag gaagattttg atgttttgct gcgaatgcgg	300
tgttgggatt tatttgttct tggagtgttc tgcgtggctg gcaaagaata	350
atgttccaaa atcgggtccat ctccaagggt gtccaatttt tcttcctggg	400
tgtcagcgag ccctgactca ctacagtgca gctgacaggg gctgtcatgc	450
aactggcccc taagccaaag caaaagacct aaggacgacc tttgaacaat	500
acaaaggatg ggtttcaatg taattaggct actgagcgga tcagctgtag	550
cactggttat agccccact gtcttactga caatgctttc tctgccgaa	600
cgaggatgcc ctaagggctg taggtgtgaa ggcaaatgg tatattgtga	650
atctcagaaa ttacaggaga taccctcaag tatactctgct ggttgcttag	700
gtttgtccct tcgctataac agccttcaaa aacttaagta taatcaattt	750
aaagggctca accagctcac ctggctatac cttgaccata accatatcag	800
caatattgac gaaaatgctt ttaatggaat acgcagactc aaagagctga	850
ttcttagttc caatagaatc tcctattttc ttaacaatac cttcagacct	900
gtgacaaatt tacggaactt ggatctgtcc tataatcagc tgcattctct	950
gggatctgaa cagtttcggg gcttgcgga gctgctgagt ttacatttac	1000
ggtctaactc cctgagaacc atccctgtgc gaatattcca agactgccgc	1050
aacctggaac ttttggaact gggatataac cggatccgaa gtttagccag	1100
gaatgtcttt gctggcatga tcagactcaa agaacttcac ctggagcaca	1150
atcaattttc caagctcaac ctggcccttt ttccaaggtt ggtcagcctt	1200
cagaaccttt acttgcagtg gaataaaatc agtgtcatag gacagaccat	1250
gtcctggacc tggagctcct tacaaaggct tgatttatca ggcaatgaga	1300
tcgaagcttt cagtggacct agtgttttcc agtgtgtccc gaatctgcag	1350
cgctcaacc tggattccaa caagctcaca tttattggtc aagagatttt	1400
ggattcttgg atatccctca atgacatcag tottgctggg aatatatggg	1450
aatgcagcag aaatatttgc tccttgttaa actggctgaa aagttttaaa	1500

-continued

ggtctaaggg agaatacaat tatctgtgcc agtcccaaag agctgcaagg	1550
agtaaatgtg atcgatgcag tgaagaacta cagcatctgt ggcaaaagta	1600
ctacagagag gtttgatctg gccagggctc tcccaaagcc gacgtttaag	1650
cccaagctcc ccaggccgaa gcatgagagc aaacccctt tgcctccgac	1700
ggtgggagcc acagagcccg gccagagac cgatgctgac gccgagcaca	1750
tctctttcca taaaatcadc gcgggcagcg tggcgctttt cctgtccgtg	1800
ctcgtcatcc tgctggttat ctacgtgtca tggaagcggc accctgcgag	1850
catgaagcag ctgcagcagc gctccctcat gcgaaggcac aggaaaaaga	1900
aaagacagtc cctaaagcaa atgactccca gcaccagga attttatgta	1950
gattataaac ccaccaacac ggagaccagc gagatgctgc tgaatgggac	2000
gggaccctgc acctataaca aatcgggctc cagggagtgt gaggtatgaa	2050
ccattgtgat aaaaagagct cttaaagct gggaaataag tggtgcttta	2100
ttgaactctg gtgactatca agggaacgcg atgcccccc tcccttccc	2150
tctccctctc actttggtgg caagatcctt cctgtccgt tttagtgcac	2200
tcataatact ggtcattttc ctctcataca taatcaaccc attgaaattt	2250
aaataccaca atcaatgtga agcttgaact ccggtttaat ataataccta	2300
ttgtataaga ccctttactg attccattaa tgcgcattt gttttaagat	2350
aaaacttctt tcataggtaa aaaaaaaaa	2379

<210> SEQ ID NO 124
<211> LENGTH: 513
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 124

Met Gly Phe Asn Val Ile Arg Leu Leu Ser Gly Ser Ala Val Ala	1	5	10	15
Leu Val Ile Ala Pro Thr Val Leu Leu Thr Met Leu Ser Ser Ala	20	25	30	
Glu Arg Gly Cys Pro Lys Gly Cys Arg Cys Glu Gly Lys Met Val	35	40	45	
Tyr Cys Glu Ser Gln Lys Leu Gln Glu Ile Pro Ser Ser Ile Ser	50	55	60	
Ala Gly Cys Leu Gly Leu Ser Leu Arg Tyr Asn Ser Leu Gln Lys	65	70	75	
Leu Lys Tyr Asn Gln Phe Lys Gly Leu Asn Gln Leu Thr Trp Leu	80	85	90	
Tyr Leu Asp His Asn His Ile Ser Asn Ile Asp Glu Asn Ala Phe	95	100	105	
Asn Gly Ile Arg Arg Leu Lys Glu Leu Ile Leu Ser Ser Asn Arg	110	115	120	
Ile Ser Tyr Phe Leu Asn Asn Thr Phe Arg Pro Val Thr Asn Leu	125	130	135	
Arg Asn Leu Asp Leu Ser Tyr Asn Gln Leu His Ser Leu Gly Ser	140	145	150	
Glu Gln Phe Arg Gly Leu Arg Lys Leu Leu Ser Leu His Leu Arg	155	160	165	

-continued

Ser	Asn	Ser	Leu	Arg	Thr	Ile	Pro	Val	Arg	Ile	Phe	Gln	Asp	Cys
				170					175					180
Arg	Asn	Leu	Glu	Leu	Leu	Asp	Leu	Gly	Tyr	Asn	Arg	Ile	Arg	Ser
				185					190					195
Leu	Ala	Arg	Asn	Val	Phe	Ala	Gly	Met	Ile	Arg	Leu	Lys	Glu	Leu
				200					205					210
His	Leu	Glu	His	Asn	Gln	Phe	Ser	Lys	Leu	Asn	Leu	Ala	Leu	Phe
				215					220					225
Pro	Arg	Leu	Val	Ser	Leu	Gln	Asn	Leu	Tyr	Leu	Gln	Trp	Asn	Lys
				230					235					240
Ile	Ser	Val	Ile	Gly	Gln	Thr	Met	Ser	Trp	Thr	Trp	Ser	Ser	Leu
				245					250					255
Gln	Arg	Leu	Asp	Leu	Ser	Gly	Asn	Glu	Ile	Glu	Ala	Phe	Ser	Gly
				260					265					270
Pro	Ser	Val	Phe	Gln	Cys	Val	Pro	Asn	Leu	Gln	Arg	Leu	Asn	Leu
				275					280					285
Asp	Ser	Asn	Lys	Leu	Thr	Phe	Ile	Gly	Gln	Glu	Ile	Leu	Asp	Ser
				290					295					300
Trp	Ile	Ser	Leu	Asn	Asp	Ile	Ser	Leu	Ala	Gly	Asn	Ile	Trp	Glu
				305					310					315
Cys	Ser	Arg	Asn	Ile	Cys	Ser	Leu	Val	Asn	Trp	Leu	Lys	Ser	Phe
				320					325					330
Lys	Gly	Leu	Arg	Glu	Asn	Thr	Ile	Ile	Cys	Ala	Ser	Pro	Lys	Glu
				335					340					345
Leu	Gln	Gly	Val	Asn	Val	Ile	Asp	Ala	Val	Lys	Asn	Tyr	Ser	Ile
				350					355					360
Cys	Gly	Lys	Ser	Thr	Thr	Glu	Arg	Phe	Asp	Leu	Ala	Arg	Ala	Leu
				365					370					375
Pro	Lys	Pro	Thr	Phe	Lys	Pro	Lys	Leu	Pro	Arg	Pro	Lys	His	Glu
				380					385					390
Ser	Lys	Pro	Pro	Leu	Pro	Pro	Thr	Val	Gly	Ala	Thr	Glu	Pro	Gly
				395					400					405
Pro	Glu	Thr	Asp	Ala	Asp	Ala	Glu	His	Ile	Ser	Phe	His	Lys	Ile
				410					415					420
Ile	Ala	Gly	Ser	Val	Ala	Leu	Phe	Leu	Ser	Val	Leu	Val	Ile	Leu
				425					430					435
Leu	Val	Ile	Tyr	Val	Ser	Trp	Lys	Arg	Tyr	Pro	Ala	Ser	Met	Lys
				440					445					450
Gln	Leu	Gln	Gln	Arg	Ser	Leu	Met	Arg	Arg	His	Arg	Lys	Lys	Lys
				455					460					465
Arg	Gln	Ser	Leu	Lys	Gln	Met	Thr	Pro	Ser	Thr	Gln	Glu	Phe	Tyr
				470					475					480
Val	Asp	Tyr	Lys	Pro	Thr	Asn	Thr	Glu	Thr	Ser	Glu	Met	Leu	Leu
				485					490					495
Asn	Gly	Thr	Gly	Pro	Cys	Thr	Tyr	Asn	Lys	Ser	Gly	Ser	Arg	Glu
				500					505					510
Cys	Glu	Val												

<210> SEQ ID NO 125
<211> LENGTH: 998
<212> TYPE: DNA

-continued

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 125

ccgttatcgt cttgcgctac tgctgaatgt ccgtcccgga ggaggaggag	50
aggcttttgc cgctgaccca gagatggccc cgagcgagca aattcctact	100
gtccggctgc gcggctaccg tggccgagct agcaaccttt ccctcggtac	150
tcacaaaaac tcgactccaa atgcaaggag aagcagctct tgctcggttg	200
ggagacggtg caagagaatc tgccccctat aggggaatgg tgcgcacagc	250
cctagggtac attgaagagg aaggctttct aaagctttgg caaggagtga	300
caccgcgcac ttacagacac gtagtgtatt ctggaggctc aatggtcaca	350
tatgaacatc tccgagaggt tgtgtttggc aaaagtgaag atgagcatta	400
ttccctttgg aaatcagtc ttggagggat gatggctggt gttattggcc	450
agtttttagc caatccaact gacctagtga aggttcagat gcaaatggaa	500
ggaaaaagga aactggaagg aaaaccattg cgatttcgtg gtgtacatca	550
tgcatattga aaaatcttag ctgaaggagg aatacgaggg ctttgggcag	600
gctgggtacc caatatacaa agagcagcac tggatgaatat gggagattta	650
accacttatg atacagtga aactacttgg gtattgaata caccacttga	700
ggacaatatc atgactcacg gtttatcaag tttatgttct ggactggtag	750
cttctattct gggaacacca gccgatgtca tcaaaagcag aataatgaat	800
caaccacgag ataaacaagg aaggggactt ttgtataaat catcgactga	850
ctgcttgatt caggctgttc aaggtgaagg attcatgagt ctatataaag	900
gctttttacc atcttggtg agaatgaccc cttggtcaat ggtgttctgg	950
cttacttatg aaaaaatcag agagatgagt ggagtcagtc cattttta	998

<210> SEQ ID NO 126

<211> LENGTH: 323

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 126

Met Ser Val Pro Glu Glu Glu Arg Leu Leu Pro Leu Thr Gln	
1 5 10 15	
Arg Trp Pro Arg Ala Ser Lys Phe Leu Leu Ser Gly Cys Ala Ala	
20 25 30	
Thr Val Ala Glu Leu Ala Thr Phe Pro Leu Asp Leu Thr Lys Thr	
35 40 45	
Arg Leu Gln Met Gln Gly Glu Ala Ala Leu Ala Arg Leu Gly Asp	
50 55 60	
Gly Ala Arg Glu Ser Ala Pro Tyr Arg Gly Met Val Arg Thr Ala	
65 70 75	
Leu Gly Ile Ile Glu Glu Glu Gly Phe Leu Lys Leu Trp Gln Gly	
80 85 90	
Val Thr Pro Ala Ile Tyr Arg His Val Val Tyr Ser Gly Gly Arg	
95 100 105	
Met Val Thr Tyr Glu His Leu Arg Glu Val Val Phe Gly Lys Ser	
110 115 120	

-continued

Glu	Asp	Glu	His	Tyr	Pro	Leu	Trp	Lys	Ser	Val	Ile	Gly	Gly	Met	
				125					130					135	
Met	Ala	Gly	Val	Ile	Gly	Gln	Phe	Leu	Ala	Asn	Pro	Thr	Asp	Leu	
				140					145					150	
Val	Lys	Val	Gln	Met	Gln	Met	Glu	Gly	Lys	Arg	Lys	Leu	Glu	Gly	
				155					160					165	
Lys	Pro	Leu	Arg	Phe	Arg	Gly	Val	His	His	Ala	Phe	Ala	Lys	Ile	
				170					175					180	
Leu	Ala	Glu	Gly	Gly	Ile	Arg	Gly	Leu	Trp	Ala	Gly	Trp	Val	Pro	
				185					190					195	
Asn	Ile	Gln	Arg	Ala	Ala	Leu	Val	Asn	Met	Gly	Asp	Leu	Thr	Thr	
				200					205					210	
Tyr	Asp	Thr	Val	Lys	His	Tyr	Leu	Val	Leu	Asn	Thr	Pro	Leu	Glu	
				215					220					225	
Asp	Asn	Ile	Met	Thr	His	Gly	Leu	Ser	Ser	Leu	Cys	Ser	Gly	Leu	
				230					235					240	
Val	Ala	Ser	Ile	Leu	Gly	Thr	Pro	Ala	Asp	Val	Ile	Lys	Ser	Arg	
				245					250					255	
Ile	Met	Asn	Gln	Pro	Arg	Asp	Lys	Gln	Gly	Arg	Gly	Leu	Leu	Tyr	
				260					265					270	
Lys	Ser	Ser	Thr	Asp	Cys	Leu	Ile	Gln	Ala	Val	Gln	Gly	Glu	Gly	
				275					280					285	
Phe	Met	Ser	Leu	Tyr	Lys	Gly	Phe	Leu	Pro	Ser	Trp	Leu	Arg	Met	
				290					295					300	
Thr	Pro	Trp	Ser	Met	Val	Phe	Trp	Leu	Thr	Tyr	Glu	Lys	Ile	Arg	
				305					310					315	
Glu	Met	Ser	Gly	Val	Ser	Pro	Phe								
				320											

<210> SEQ ID NO 127
<211> LENGTH: 1505
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 127

cgcggaatcgg acccaagcag gtcggcggcg gcggcaggag agcggccggg	50
cgctcagctcc tcgacccccg tgtcgggcta gtccagcgag gcggacgggc	100
ggcgtgggcc catggccagg cccggcatgg agcggtggcg cgaccggctg	150
gcgctggtga cgggggcctc ggggggcctc ggcgcggccg tggcccgggc	200
cctggtccag cagggaactga aggtggtggg ctgcgccgcg actgtgggca	250
acatcgagga gctggctgct gaatgtaaga gtgcaggcta ccccgggact	300
ttgatcccct acagatgtga cctatcaaat gaagaggaca tcctctccat	350
gttctcagct atccgttctc agcacagcgg tgtagacatc tgcataaca	400
atgctggctt ggcccgccct gacacccctg totcaggcag caccagtgtt	450
tggaaggaca tgttcaatgt gaacgtgctg gccctcagca tctgcacacg	500
ggaagcctac cagtccatga aggagcggaa tgtggacgat gggcacatca	550
ttaacatcaa tagcatgtct ggccaccgag tgttaccctt gtctgtgacc	600
cacttctata gtgccaccaa gtatgccgtc actgcgctga cagagggact	650

-continued

gaggcaagag cttcgggagg cccagaccca catccgagcc acgtgcatct	700
ctccaggtgt ggtggagaca caattcgctt tcaaactcca cgacaaggac	750
cctgagaagg cagctgccac ctatgagcaa atgaagtgtc tcaaaccgga	800
ggatgtggcg gaggtgttta tctacgtcct cagcaccccc gcacacatcc	850
agattggaga catccagatg aggccacgg agcagggtgac ctagtgactg	900
tgggagctcc tccttccttc cccacccttc atggcttgcc tcctgcctct	950
ggatttttagg tgttgatttc tggatcacgg gataccactt cctgtccaca	1000
ccccgaccag gggctagaaa atttgtttga gatttttata tcatcttgct	1050
aaattgcttc agttgtaaat gtgaaaaatg ggctggggaa aggaggtggt	1100
gtccctaatt gttttacttg ttaacttggt cttgtgcccc tgggcacttg	1150
gcctttgtct gctctcagtg tcttcctttt gacatgggaa aggagttgtg	1200
gccaaaaatc ccatcttctt gcacctcaac gtctgtggct cagggctggg	1250
gtggcagagg gaggccttca ccttatatct gtgtgttat ccagggctcc	1300
agacttcctc ctctgcctgc cccactgcac cctctcccc ttatctatct	1350
ccttctcggc tccccagccc agtcttggtt tcttgcccc tcctggggtc	1400
atccctccac tctgactctg actatggcag cagaacacca gggcctggcc	1450
cagtggtatt catggtgatc attaaaaag aaaaatcgca accaaaaaaa	1500
aaaaa	1505

<210> SEQ ID NO 128

<211> LENGTH: 260

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 128

Met	Ala	Arg	Pro	Gly	Met	Glu	Arg	Trp	Arg	Asp	Arg	Leu	Ala	Leu
1				5					10					15
Val	Thr	Gly	Ala	Ser	Gly	Gly	Ile	Gly	Ala	Ala	Val	Ala	Arg	Ala
				20					25					30
Leu	Val	Gln	Gln	Gly	Leu	Lys	Val	Val	Gly	Cys	Ala	Arg	Thr	Val
				35					40					45
Gly	Asn	Ile	Glu	Glu	Leu	Ala	Ala	Glu	Cys	Lys	Ser	Ala	Gly	Tyr
				50					55					60
Pro	Gly	Thr	Leu	Ile	Pro	Tyr	Arg	Cys	Asp	Leu	Ser	Asn	Glu	Glu
				65					70					75
Asp	Ile	Leu	Ser	Met	Phe	Ser	Ala	Ile	Arg	Ser	Gln	His	Ser	Gly
				80					85					90
Val	Asp	Ile	Cys	Ile	Asn	Asn	Ala	Gly	Leu	Ala	Arg	Pro	Asp	Thr
				95					100					105
Leu	Leu	Ser	Gly	Ser	Thr	Ser	Gly	Trp	Lys	Asp	Met	Phe	Asn	Val
				110					115					120
Asn	Val	Leu	Ala	Leu	Ser	Ile	Cys	Thr	Arg	Glu	Ala	Tyr	Gln	Ser
				125					130					135
Met	Lys	Glu	Arg	Asn	Val	Asp	Asp	Gly	His	Ile	Ile	Asn	Ile	Asn
				140					145					150
Ser	Met	Ser	Gly	His	Arg	Val	Leu	Pro	Leu	Ser	Val	Thr	His	Phe
				155					160					165

-continued

Tyr	Ser	Ala	Thr	Lys	Tyr	Ala	Val	Thr	Ala	Leu	Thr	Glu	Gly	Leu
				170					175					180
Arg	Gln	Glu	Leu	Arg	Glu	Ala	Gln	Thr	His	Ile	Arg	Ala	Thr	Cys
				185					190					195
Ile	Ser	Pro	Gly	Val	Val	Glu	Thr	Gln	Phe	Ala	Phe	Lys	Leu	His
				200					205					210
Asp	Lys	Asp	Pro	Glu	Lys	Ala	Ala	Ala	Thr	Tyr	Glu	Gln	Met	Lys
				215					220					225
Cys	Leu	Lys	Pro	Glu	Asp	Val	Ala	Glu	Ala	Val	Ile	Tyr	Val	Leu
				230					235					240
Ser	Thr	Pro	Ala	His	Ile	Gln	Ile	Gly	Asp	Ile	Gln	Met	Arg	Pro
				245					250					255
Thr	Glu	Gln	Val	Thr										
				260										

<210> SEQ ID NO 129
<211> LENGTH: 1177
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 129

aactttctaca tgggcctcct gctgctgggtg ctcttcctca gcctcctgcc	50
ggtggcctac accatcatgt ccctcccacc ctcccttgac tgcggggcgt	100
tcaggtgcag agtctcagtt gcccgggagc acctcccctc ccgaggcagt	150
ctgctcagag ggctcggcc cagaattcca gttctggttt catgccagcc	200
tgtaaaaggc catggaactt tgggtgaatc accgatgcca tttaagaggg	250
ttttctgcc ggaatgaaat gttaggtcgt tctgtgtctg cgtgttcat	300
ttcagtagcc accagccacc tgtggccgtt gagtgttga aatgaggaa	350
tgagaaaatt aatttctcat gtatttttct cattatttta ttaattttta	400
actgatagtt gtacatatat ggggttacat gtgatatttg gatacatgta	450
tacaatatat aatgatcaaa tcagggtaac tgggatatcc atcacatcaa	500
acatttattt ttatttcttt ttagacagag tctcactctg taccagggc	550
tggagtgcag tgggtgccatc tcagcttact gcaacctctg cctgccaggt	600
tcaagcgatt ctcatgcctc cacctcccaa gtagctggga ctacaggcat	650
gcaccacaat gcccaactaa tttttgtatt tttagtagag acggggtttt	700
gccatgttgc ccaggctggc ctggaactcc tggcctcaaa caatccactt	750
gcctcgccct cccaaagtgt tatgattaca ggcgtgagcc accgtgcctg	800
gcctaaacat ttatcttttc ttgtgttgg gaactttgaa attatacaat	850
gaattattgt taactgtcat ctccctgctg tgctatggaa cactgggact	900
tcttccctct atctaactgt atatttgtac cagttaacca accgtacttc	950
atccccactc ctctctatcc ttcccaacct ctgatcacct cattctactc	1000
tctacctcca tgagatccac ttttttagct cccacatgtg agtaagaaaa	1050
tgcaatatatt gtctttctgt gcctggctta tttcacttaa cataatgact	1100
tcctgttcca tccatgttgc tgcaaatgac aggatttcgt tcttaatttc	1150

-continued

aattaaata accacacatg gcaaaaa 1177

<210> SEQ ID NO 130
<211> LENGTH: 111
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 130

Met Gly Leu Leu Leu Val Leu Phe Leu Ser Leu Leu Pro Val
1 5 10 15
Ala Tyr Thr Ile Met Ser Leu Pro Pro Ser Phe Asp Cys Gly Pro
20 25 30
Phe Arg Cys Arg Val Ser Val Ala Arg Glu His Leu Pro Ser Arg
35 40 45
Gly Ser Leu Leu Arg Gly Pro Arg Pro Arg Ile Pro Val Leu Val
50 55 60
Ser Cys Gln Pro Val Lys Gly His Gly Thr Leu Gly Glu Ser Pro
65 70 75
Met Pro Phe Lys Arg Val Phe Cys Gln Asp Gly Asn Val Arg Ser
80 85 90
Phe Cys Val Cys Ala Val His Phe Ser Ser His Gln Pro Pro Val
95 100 105
Ala Val Glu Cys Leu Lys
110

<210> SEQ ID NO 131
<211> LENGTH: 2061
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 131

ttctgaagta acggaagcta ccttgataaa agacctcaac actgctgacc 50
atgatcagcg cagcctggag catcttcctc atcgggacta aaattgggct 100
gttccttcaa gtagcacctc tatcagttat ggctaaatcc tgtccatctg 150
tgtgtcgctg cgatgcgggt ttcatttact gtaatgatcg ctttctgaca 200
tccattccaa caggaatacc agaggatgct acaactctct accttcagaa 250
caaccaaata aataatgctg ggattccttc agatttgaaa aacttgctga 300
aagtagaaag aatataccta taccacaaca gttagatga atttcctacc 350
aacctcccaa agtatgtaaa agagttacat ttgcaagaaa ataacataag 400
gactatcact tatgattcac tttcaaaaat tccctatctg gaagaattac 450
atttagatga caactctgtc tctgcagtta gcatagaaga gggagcattc 500
cgagacagca actatctccg actgcttttc ctgtcccgta atcaccttag 550
cacaattccc tggggtttgc ccaggactat agaagaacta cgcttgatg 600
ataatcgcat atccactatt tcatcccat ctcttcaagg tctcactagt 650
ctaaaacgcc tggttctaga tggaacctg ttgaacaatc atggtttagg 700
tgacaaagtt ttcttcaacc tagttaattt gacagagctg tccctgggtg 750
ggaattccct gactgctgca ccagtaaacc ttccaggcac aaacctgagg 800
aagctttatc ttcaagataa ccacatcaat cgggtgcccc caaatgcttt 850

-continued

ttcttatcta aggcagctct atcgactgga tatgtccaat aataacctaa	900
gtaatttacc tcagggtatc ttgatgatt tggacaatat aacacaactg	950
attcttcgca acaatccctg gtattgcggg tgcaagatga aatgggtacg	1000
tgactggtta caatcactac ctgtgaaggt caacgtgcgt gggctcatgt	1050
gccaagcccc agaaaagggt cgtgggatgg ctattaagga tctcaatgca	1100
gaactgtttg attgtaagga cagtgggatt gtaagcacca ttcagataac	1150
cactgcaata cccaacacag tgtatcctgc ccaaggacag tggccagctc	1200
cagtgcacaa acagccagat attaagaacc ccaagctcac taaggatcaa	1250
caaaccacag ggagtcctc aagaaaaaca attacaatta ctgtgaagtc	1300
tgtcacctct gataccattc atatctcttg gaaacttgct ctacctatga	1350
ctgctttgag actcagctgg cttaactgg gccatagccc ggcatttgga	1400
tctataacag aaacaattgt aacaggggaa cgcagtgagt acttggtcac	1450
agccctggag cctgattcac cctataaagt atgcatgggt cccatggaaa	1500
ccagcaacct ctacctattt gatgaaactc ctgtttgtat tgagactgaa	1550
actgcacccc ttcgaatgta caaccctaca accaccctca atcgagagca	1600
agagaaagaa ccttacaaaa accccaattt acctttggct gccatcattg	1650
gtggggctgt ggccctgggt accattgccc ttcttgcttt agtgtgttg	1700
tatgttcata ggaatggatc gctcttctca aggaactgtg catatagcaa	1750
agggaggaga agaaaggatg actatgcaga agctggcact aagaaggaca	1800
actctatcct ggaaatcagg gaaacttctt ttcagatggt accaataagc	1850
aatgaaccca tctcgaagga ggagtttgta atacacacca tatttcctcc	1900
taatggaatg aatctgtaca aaaacaatca cagtgaaagc agtagtaacc	1950
gaagctacag agacagtggg attccagact cagatcactc acactcatga	2000
tgctgaagga ctcacagcag acttgtgttt tgggtttttt aaacctaagg	2050
gaggtgatgg t	2061

<210> SEQ ID NO 132

<211> LENGTH: 649

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 132

Met	Ile	Ser	Ala	Ala	Trp	Ser	Ile	Phe	Leu	Ile	Gly	Thr	Lys	Ile
1				5					10					15
Gly	Leu	Phe	Leu	Gln	Val	Ala	Pro	Leu	Ser	Val	Met	Ala	Lys	Ser
			20						25					30
Cys	Pro	Ser	Val	Cys	Arg	Cys	Asp	Ala	Gly	Phe	Ile	Tyr	Cys	Asn
			35						40					45
Asp	Arg	Phe	Leu	Thr	Ser	Ile	Pro	Thr	Gly	Ile	Pro	Glu	Asp	Ala
			50						55					60
Thr	Thr	Leu	Tyr	Leu	Gln	Asn	Asn	Gln	Ile	Asn	Asn	Ala	Gly	Ile
			65						70					75
Pro	Ser	Asp	Leu	Lys	Asn	Leu	Leu	Lys	Val	Glu	Arg	Ile	Tyr	Leu
			80						85					90

-continued

Tyr	His	Asn	Ser	Leu	Asp	Glu	Phe	Pro	Thr	Asn	Leu	Pro	Lys	Tyr
				95					100					105
Val	Lys	Glu	Leu	His	Leu	Gln	Glu	Asn	Asn	Ile	Arg	Thr	Ile	Thr
				110					115					120
Tyr	Asp	Ser	Leu	Ser	Lys	Ile	Pro	Tyr	Leu	Glu	Glu	Leu	His	Leu
				125					130					135
Asp	Asp	Asn	Ser	Val	Ser	Ala	Val	Ser	Ile	Glu	Glu	Gly	Ala	Phe
				140					145					150
Arg	Asp	Ser	Asn	Tyr	Leu	Arg	Leu	Leu	Phe	Leu	Ser	Arg	Asn	His
				155					160					165
Leu	Ser	Thr	Ile	Pro	Trp	Gly	Leu	Pro	Arg	Thr	Ile	Glu	Glu	Leu
				170					175					180
Arg	Leu	Asp	Asp	Asn	Arg	Ile	Ser	Thr	Ile	Ser	Ser	Pro	Ser	Leu
				185					190					195
Gln	Gly	Leu	Thr	Ser	Leu	Lys	Arg	Leu	Val	Leu	Asp	Gly	Asn	Leu
				200					205					210
Leu	Asn	Asn	His	Gly	Leu	Gly	Asp	Lys	Val	Phe	Phe	Asn	Leu	Val
				215					220					225
Asn	Leu	Thr	Glu	Leu	Ser	Leu	Val	Arg	Asn	Ser	Leu	Thr	Ala	Ala
				230					235					240
Pro	Val	Asn	Leu	Pro	Gly	Thr	Asn	Leu	Arg	Lys	Leu	Tyr	Leu	Gln
				245					250					255
Asp	Asn	His	Ile	Asn	Arg	Val	Pro	Pro	Asn	Ala	Phe	Ser	Tyr	Leu
				260					265					270
Arg	Gln	Leu	Tyr	Arg	Leu	Asp	Met	Ser	Asn	Asn	Asn	Leu	Ser	Asn
				275					280					285
Leu	Pro	Gln	Gly	Ile	Phe	Asp	Asp	Leu	Asp	Asn	Ile	Thr	Gln	Leu
				290					295					300
Ile	Leu	Arg	Asn	Asn	Pro	Trp	Tyr	Cys	Gly	Cys	Lys	Met	Lys	Trp
				305					310					315
Val	Arg	Asp	Trp	Leu	Gln	Ser	Leu	Pro	Val	Lys	Val	Asn	Val	Arg
				320					325					330
Gly	Leu	Met	Cys	Gln	Ala	Pro	Glu	Lys	Val	Arg	Gly	Met	Ala	Ile
				335					340					345
Lys	Asp	Leu	Asn	Ala	Glu	Leu	Phe	Asp	Cys	Lys	Asp	Ser	Gly	Ile
				350					355					360
Val	Ser	Thr	Ile	Gln	Ile	Thr	Thr	Ala	Ile	Pro	Asn	Thr	Val	Tyr
				365					370					375
Pro	Ala	Gln	Gly	Gln	Trp	Pro	Ala	Pro	Val	Thr	Lys	Gln	Pro	Asp
				380					385					390
Ile	Lys	Asn	Pro	Lys	Leu	Thr	Lys	Asp	Gln	Gln	Thr	Thr	Gly	Ser
				395					400					405
Pro	Ser	Arg	Lys	Thr	Ile	Thr	Ile	Thr	Val	Lys	Ser	Val	Thr	Ser
				410					415					420
Asp	Thr	Ile	His	Ile	Ser	Trp	Lys	Leu	Ala	Leu	Pro	Met	Thr	Ala
				425					430					435
Leu	Arg	Leu	Ser	Trp	Leu	Lys	Leu	Gly	His	Ser	Pro	Ala	Phe	Gly
				440					445					450
Ser	Ile	Thr	Glu	Thr	Ile	Val	Thr	Gly	Glu	Arg	Ser	Glu	Tyr	Leu
				455					460					465
Val	Thr	Ala	Leu	Glu	Pro	Asp	Ser	Pro	Tyr	Lys	Val	Cys	Met	Val

-continued

470										475										480									
Pro	Met	Glu	Thr	Ser	Asn	Leu	Tyr	Leu	Phe	Asp	Glu	Thr	Pro	Val															
				485						490				495															
Cys	Ile	Glu	Thr	Glu	Thr	Ala	Pro	Leu	Arg	Met	Tyr	Asn	Pro	Thr															
				500						505				510															
Thr	Thr	Leu	Asn	Arg	Glu	Gln	Glu	Lys	Glu	Pro	Tyr	Lys	Asn	Pro															
				515						520				525															
Asn	Leu	Pro	Leu	Ala	Ala	Ile	Ile	Gly	Gly	Ala	Val	Ala	Leu	Val															
				530						535				540															
Thr	Ile	Ala	Leu	Leu	Ala	Leu	Val	Cys	Trp	Tyr	Val	His	Arg	Asn															
				545						550				555															
Gly	Ser	Leu	Phe	Ser	Arg	Asn	Cys	Ala	Tyr	Ser	Lys	Gly	Arg	Arg															
				560						565				570															
Arg	Lys	Asp	Asp	Tyr	Ala	Glu	Ala	Gly	Thr	Lys	Lys	Asp	Asn	Ser															
				575						580				585															
Ile	Leu	Glu	Ile	Arg	Glu	Thr	Ser	Phe	Gln	Met	Leu	Pro	Ile	Ser															
				590						595				600															
Asn	Glu	Pro	Ile	Ser	Lys	Glu	Glu	Phe	Val	Ile	His	Thr	Ile	Phe															
				605						610				615															
Pro	Pro	Asn	Gly	Met	Asn	Leu	Tyr	Lys	Asn	Asn	His	Ser	Glu	Ser															
				620						625				630															
Ser	Ser	Asn	Arg	Ser	Tyr	Arg	Asp	Ser	Gly	Ile	Pro	Asp	Ser	Asp															
				635						640				645															
His	Ser	His	Ser																										

<210> SEQ ID NO 133
<211> LENGTH: 1882
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 133

ccgtcatccc cctgcagcca ccctcccg agtcctttgc ccaggccacc	50
ccaggcttct tggcagccct gccgggccac ttgtcttcat gtctgccagg	100
gggaggtggg aaggaggtgg gaggagggcg tgcagaggca gtctgggctt	150
ggccagagct cagggtgctg agcgtgtgac cagcagtgag cagaggccgg	200
ccatggccag cctggggctg ctgctcctgc tcttactgac agcactgccca	250
ccgctgtggt cctcctcact gcctgggctg gacactgctg aaagtaaagc	300
caccattgca gacctgatcc tgtctgcgct ggagagagcc accgtcttcc	350
tagaacagag gctgcctgaa atcaacctgg atggcatggt ggggggccga	400
gtgctggaag agcagctaaa aagtgtccgg gagaagtggg cccaggagcc	450
cctgctgacg ccgctgagcc tgcgcgtggg gatgctgggg gagaagtctg	500
aggctgccat ccagagatcc ctccactacc tcaagtctgag tgatcccaag	550
tacctaagag agttccagct gaccctccag cccgggtttt ggaagctccc	600
acatgcctgg atccacactg atgcctcctt ggtgtacccc acgttcgggc	650
cccaggactc attctcagag gagagaagtg acgtgtgcct ggtgcagctg	700
ctgggaacog ggacggacag cagcgagccc tgcggcctct cagacctctg	750
caggagcctc atgaccaagc ccggctgctc aggtactctg ctgtcccacc	800

-continued

```
aactgctctt cttcctctgg gccagaatga ggggatgcac acagggacca      850
ctccaacaga gccaggacta tatcaacctc ttctgcgcca acatgatgga      900
cttgaaccgc agagctgagg ccatcggata cgcctaccct acccgggaca      950
tcttcatgga aaacatcatg ttctgtgga tggcgggctt ctccgacttc     1000
tacaagctcc ggtggctgga ggccattctc agctggcaga aacagcagga     1050
aggatgcttc ggggagcctg atgctgaaga tgaagaatta tctaaagcta     1100
ttcaatatca gcagcathtt tcgaggagag tgaagaggcg agaaaaacaa     1150
tttccagatt ctgctctgt tgctcaggct ggagtacagt ggcgcaatct     1200
cggctcactg caacctttgc ctctggggtt caagcaattc tcttgctca     1250
tcctcccag tagctgggac tacaggagcg tgccaccata cctggctaata     1300
ttttatatatt ttttagtaga gacagggttt catcatgttg ctcatgctgg     1350
tctcgaactc ctgatctcaa gagatccgco cacctcaggc tcccaaagt     1400
tggtgattata ggtgtgagcc accgtgtctg gctgaaaagc actttcaaag     1450
agactgtggt gaataaaggg ccaaggttct tgccaccag cactcatggg     1500
ggctctctcc cctagatggc tgctcctccc acaacacagc cacagcagt     1550
gcagccctgg gtggcttcct atacatcctg gcagaatacc cccagcaaaa     1600
cagagagcca caccatcca caccgccacc accaagcagc cgctgagacg     1650
gacggttcca tgccagctgc ctggaggagg aacagacccc tttagtctc     1700
atcccttaga tcctggaggg cacggatcac atcctgggaa gaaggcatct     1750
ggaggataag caaagccacc ccgacacca atcttggaag ccctgagtag     1800
gcagggccag ggtagggtgg ggccgggagg gacccagtg tgaacggatg     1850
aataaagttc aactgcaact gaaaaaaaaa aa                        1882
```

<210> SEQ ID NO 134

<211> LENGTH: 440

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 134

```
Met Ser Ala Arg Gly Arg Trp Glu Gly Gly Gly Arg Arg Ala Cys
 1             5             10            15
Arg Gly Ser Leu Gly Leu Ala Arg Ala Gln Gly Ala Glu Arg Val
          20             25            30
Thr Ser Ser Glu Gln Arg Pro Ala Met Ala Ser Leu Gly Leu Leu
          35             40            45
Leu Leu Leu Leu Leu Thr Ala Leu Pro Pro Leu Trp Ser Ser Ser
          50             55            60
Leu Pro Gly Leu Asp Thr Ala Glu Ser Lys Ala Thr Ile Ala Asp
          65             70            75
Leu Ile Leu Ser Ala Leu Glu Arg Ala Thr Val Phe Leu Glu Gln
          80             85            90
Arg Leu Pro Glu Ile Asn Leu Asp Gly Met Val Gly Val Arg Val
          95            100           105
Leu Glu Glu Gln Leu Lys Ser Val Arg Glu Lys Trp Ala Gln Glu
          110           115           120
```

-continued

Pro Leu Leu Gln Pro Leu Ser Leu Arg Val Gly Met Leu Gly Glu
125 130 135

Lys Leu Glu Ala Ala Ile Gln Arg Ser Leu His Tyr Leu Lys Leu
140 145 150

Ser Asp Pro Lys Tyr Leu Arg Glu Phe Gln Leu Thr Leu Gln Pro
155 160 165

Gly Phe Trp Lys Leu Pro His Ala Trp Ile His Thr Asp Ala Ser
170 175 180

Leu Val Tyr Pro Thr Phe Gly Pro Gln Asp Ser Phe Ser Glu Glu
185 190 195

Arg Ser Asp Val Cys Leu Val Gln Leu Leu Gly Thr Gly Thr Asp
200 205 210

Ser Ser Glu Pro Cys Gly Leu Ser Asp Leu Cys Arg Ser Leu Met
215 220 225

Thr Lys Pro Gly Cys Ser Gly Tyr Cys Leu Ser His Gln Leu Leu
230 235 240

Phe Phe Leu Trp Ala Arg Met Arg Gly Cys Thr Gln Gly Pro Leu
245 250 255

Gln Gln Ser Gln Asp Tyr Ile Asn Leu Phe Cys Ala Asn Met Met
260 265 270

Asp Leu Asn Arg Arg Ala Glu Ala Ile Gly Tyr Ala Tyr Pro Thr
275 280 285

Arg Asp Ile Phe Met Glu Asn Ile Met Phe Cys Gly Met Gly Gly
290 295 300

Phe Ser Asp Phe Tyr Lys Leu Arg Trp Leu Glu Ala Ile Leu Ser
305 310 315

Trp Gln Lys Gln Gln Glu Gly Cys Phe Gly Glu Pro Asp Ala Glu
320 325 330

Asp Glu Glu Leu Ser Lys Ala Ile Gln Tyr Gln Gln His Phe Ser
335 340 345

Arg Arg Val Lys Arg Arg Glu Lys Gln Phe Pro Asp Ser Arg Ser
350 355 360

Val Ala Gln Ala Gly Val Gln Trp Arg Asn Leu Gly Ser Leu Gln
365 370 375

Pro Leu Pro Pro Gly Phe Lys Gln Phe Ser Cys Leu Ile Leu Pro
380 385 390

Ser Ser Trp Asp Tyr Arg Ser Val Pro Pro Tyr Leu Ala Asn Phe
395 400 405

Tyr Ile Phe Leu Val Glu Thr Gly Phe His His Val Ala His Ala
410 415 420

Gly Leu Glu Leu Leu Ile Ser Arg Asp Pro Pro Thr Ser Gly Ser
425 430 435

Gln Ser Val Gly Leu
440

<210> SEQ ID NO 135

<211> LENGTH: 884

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 135

gggtctgagtg cagagctgct gtcattggcgg ccgctctgtg gggcttcttt

-continued

```
cccgctcctgc tgctgctgct gctatcgggg gatgtccaga gctcggagggt      100
gccccggggct gctgctgagg gatcggggagg gagtgggggtc ggcataggag      150
atcgcttcaa gattgagggg cgtgcagttg ttccaggggt gaagcctcag      200
gactggatct cggcggcccg agtgcaggta gacggagaag agcacgtcgg      250
tttccttaag acagatggga gttttgtggt tcatgatata ctttctggat      300
cttatgtagt ggaagttgta tctccagctt acagatttga tcccgttcga      350
gtggatatca cttcgaaagg aaaaatgaga gcaagatatg tgaattacat      400
caaaacatca gaggttgtca gactgcccta tcctctccaa atgaaatctt      450
cagggtccacc ttcttacttt attaaaaggg aatcgtgggg ctggacagac      500
tttctaatag acccaatggt tatgatgatg gttcttcctt tattgatatt      550
tgtgtctctg cctaaagtgg tcaacacaag tgatcctgac atgagacggg      600
aaatggagca gtcaatgaat atgctgaatt ccaaccatga gttgcctgat      650
gtttctgagt tcatgacaag actcttctct tcaaaatcat ctggcaaatc      700
tagcagcgcg agcagtaaaa caggcaaaag tggggctggc aaaaggagggt      750
agtcaggccg tccagagctg gcatttcac aaacacggca aactggggtg      800
gcatccaagt cttggaaaac cgtgtgaagc aactactata aacttgagtc      850
atcccgcagt tgatctctta caactgtgta tggt      884
```

<210> SEQ ID NO 136

<211> LENGTH: 242

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 136

```
Met Ala Ala Ala Leu Trp Gly Phe Phe Pro Val Leu Leu Leu Leu      1
1      5      10     15
Leu Leu Ser Gly Asp Val Gln Ser Ser Glu Val Pro Gly Ala Ala      20
20     25     30
Ala Glu Gly Ser Gly Gly Ser Gly Val Gly Ile Gly Asp Arg Phe      35
35     40     45
Lys Ile Glu Gly Arg Ala Val Val Pro Gly Val Lys Pro Gln Asp      50
50     55     60
Trp Ile Ser Ala Ala Arg Val Leu Val Asp Gly Glu Glu His Val      65
65     70     75
Gly Phe Leu Lys Thr Asp Gly Ser Phe Val Val His Asp Ile Pro      80
80     85     90
Ser Gly Ser Tyr Val Val Glu Val Val Ser Pro Ala Tyr Arg Phe      95
95    100    105
Asp Pro Val Arg Val Asp Ile Thr Ser Lys Gly Lys Met Arg Ala     110
110    115    120
Arg Tyr Val Asn Tyr Ile Lys Thr Ser Glu Val Val Arg Leu Pro     125
125    130    135
Tyr Pro Leu Gln Met Lys Ser Ser Gly Pro Pro Ser Tyr Phe Ile     140
140    145    150
Lys Arg Glu Ser Trp Gly Trp Thr Asp Phe Leu Met Asn Pro Met     155
155    160    165
```

-continued

Val	Met	Met	Met	Val	Leu	Pro	Leu	Leu	Ile	Phe	Val	Leu	Leu	Pro
				170					175					180
Lys	Val	Val	Asn	Thr	Ser	Asp	Pro	Asp	Met	Arg	Arg	Glu	Met	Glu
				185					190					195
Gln	Ser	Met	Asn	Met	Leu	Asn	Ser	Asn	His	Glu	Leu	Pro	Asp	Val
				200					205					210
Ser	Glu	Phe	Met	Thr	Arg	Leu	Phe	Ser	Ser	Lys	Ser	Ser	Gly	Lys
				215					220					225
Ser	Ser	Ser	Gly	Ser	Ser	Lys	Thr	Gly	Lys	Ser	Gly	Ala	Gly	Lys
				230					235					240
Arg	Arg													

<210> SEQ ID NO 137
<211> LENGTH: 1571
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 137

gatggcgag ccacagcttc tgtgagattc gatttctccc cagttcccct	50
gtgggtctga ggggaccaga agggtgagct acgttggtt tctggaagg	100
gaggctatat gcgtcaattc cccaaaacaa gttttgacat ttcccctgaa	150
atgtcattct ctatctattc actgcaagtg cctgctgttc caggccttac	200
ctgctgggga ctaacggcgg agccaggatg gggacagaat aaaggagcca	250
cgacctgtgc caccaactcg cactcagact ctgaactcag acctgaaatc	300
ttctcttcac gggaggcttg gcagtttttc ttactcctgt ggtctccaga	350
tttcaggcct aagatgaaag cctctagtct tgccttcagc cttctctctg	400
ctgcgtttta tctcctatgg actccttcca ctggactgaa gacactcaat	450
ttgggaagct gtgtgatcgc cacaacctt caggaaatac gaaatggatt	500
ttctgagata cggggcagtg tgcaagccaa agatggaaac attgacatca	550
gaatcttaag gaggactgag tctttgcaag acacaaagcc tgcgaatcga	600
tgctgcctcc tgcgccattt gctaagactc tatctggaca gggatattta	650
aaactaccag acccctgacc attatactct ccggaagatc agcagcctcg	700
ccaattcctt tcttaccatc aagaaggacc tccggctctc tcatgcccac	750
atgacatgcc attgtgggga ggaagcaatg aagaaataca gccagattct	800
gagtcacttt gaaaagctgg aacctcaggc agcagttgtg aaggctttgg	850
gggaactaga cattcttctg caatggatgg aggagacaga ataggaggaa	900
agtgatgctg ctgctaagaa tattcgaggt caagagctcc agtcttcaat	950
acctgcagag gaggcatgac cccaaaccac catctcttta ctgtactagt	1000
cttgtgtctg tcacagtgtg tcttatttat gcattacttg cttccttgca	1050
tgattgtctt tatgcatccc caatcttaat tgagaccata cttgtataag	1100
atttttgtaa tatctttctg ctattggata tatttattag ttaatatatt	1150
tattttattt ttgctattta atgtatttat ttttttactt ggacatgaaa	1200
ctttaaaaaa attcacagat tatatttata acctgactag agcaggtgat	1250
gtatttttat acagtaaaaa aaaaaaacct tgtaaattct agaagagtgg	1300

-continued

<210> SEQ ID NO 139

<211> LENGTH: 2395

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 139

cctggagcgcg gaagcgcggc tgcagcaggc cgaggctcca ggtggggctcg	50
gttccgcatac cagcctagcg tgtccacgat gcggctgggc tccgggactt	100
tcgctacctg ttgcgtagcg atcgaggctg tagggatcgc ggtcttctct	150
cggggattct tcccggctcc cgttcgttcc tctgccagag cggaacacgg	200
agcgggagccc ccagcgcccc aaccctcggc tggagccagt tctaactgga	250
ccacgctgcc accacctctc ttcagtaaa ttgttattgt tctgatagat	300
gccttgagag atgattttgt gtttgggtca aagggtgtga aatttatgcc	350
ctacacaact taccttgttg aaaaaggagc atctcacagt tttgtggctg	400
aagcaaaagg ccctacagtt actatgccto gaatcaaggc attgatgacg	450
gggagccttc ctggcctttgt cgacgtcatc aggaacctca attctcctgc	500
actgctggaa gacagtgtga taagacaagc aaaagcagct ggaaaaagaa	550
tagtctttta tggagatgaa acctggggtta aattattccc aaagcatttt	600
gtggaatatg atggaacaac ctcatctttc gtgtcagatt acacagaggt	650
ggataataat gtcacgaggc atttgataaa agtattaaaa agaggagatt	700
gggacataatt aatcctccac tacctggggc tggaccacat tggccacatt	750
tcaggggccc acagccccct gattgggcag aagctgagcg agatggacag	800
cgtgctgatg aagatccaca cctcactgca gtcgaaggag agagagacgc	850
ctttaccocaa tttgctgggt ctttgtgtg accatggcat gctgaaaca	900
ggaagtcaag gggcctcctc caccgaggag gtgaatacac ctctgatttt	950
aatcagttct gcgtttgaaa ggaaccocgg tgatatccga catccaaagc	1000
acgtccaata gacggatgtg gctgcgacac tggcgatagc acttggttta	1050
ccgattccaa aagacagtgt agggagcctc ctattcccag ttgtggaagg	1100
aagaccaatg agagagcagt tgagattttt acatttgaat acagtgcagc	1150
ttagtaaaact gttgcaagag aatgtgccgt catatgaaaa agatcctggg	1200
tttgagcagt ttaaaatgtc agaaagattg catgggaact ggatcagact	1250
gtacttgagg gaaaagcatt cagaagtcct attcaacctg ggctccaagg	1300
ttctcaggca gtacctggat gctctgaaga cgctgagctt gtccctgagt	1350
gcacaagtgg ccagttctc accctgctcc tgctcagcgt cccacaggca	1400
ctgcacagaa aggctgagct ggaagtccca ctgtcatctc ctgggttttc	1450
tctgctcttt tatttggtga tcctggttct ttccggcgtt cactgcattg	1500
tgtgcacctc agctgaaagt tcgtgctact tctgtggcct ctcgtggctg	1550
gcggcaggct gcctttcggt taccagactc tgggtgaaca cctgggtgtg	1600
gccaaagtgt ggcagtgcc tggacagggg gcctcagga aggacgtgga	1650
gcagccttat ccagggcctc tgggtgtccc gacacaggtg ttcacatctg	1700
tgctgtcagg tcagatgcct cagttcttgg aaagctaggt tcctgcgact	1750

-continued

gttaccaagg tgattgtaaa gagctggcgg tcacagagga acaagcccc 1800
cagctgaggg ggtgtgtgaa tcggacagcc tcccagcaga ggtgtgggag 1850
ctgcagctga gggaagaaga gacaatcggc ctggacactc aggaggggtca 1900
aaaggagact tggtcgcacc actcatcctg ccacccccag aatgcacacct 1950
gcctcatcag gtccagattt ctttccaagg cggacgtttt ctgttggaat 2000
tcttagtcct tggcctcgga caccttcatt cgtagctgg ggagtgggtg 2050
tgaggcagtg aagaagaggc ggatggtcac actcagatcc acagagccca 2100
ggatcaaggg acccactgca gtggcagcag gactgttggg cccccacccc 2150
aaccctgcac agccctcatc ccctcttggc ttgagccgtc agaggccctg 2200
tgctgagtg ctgaccgaga cactcacagc tttgtcatca gggcacaggc 2250
ttcctcggag ccaggatgat ctgtgccacg cttgcacctc gggcccatct 2300
gggctcatgc tctctctcct gctattgaat tagtacctag ctgcacacag 2350
tatgtagtta ccaaagaat aaacggcaat aattgagaaa aaaaa 2395

<210> SEQ ID NO 140

<211> LENGTH: 310

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 140

Met Arg Leu Gly Ser Gly Thr Phe Ala Thr Cys Cys Val Ala Ile
1 5 10 15
Glu Val Leu Gly Ile Ala Val Phe Leu Arg Gly Phe Phe Pro Ala
20 25 30
Pro Val Arg Ser Ser Ala Arg Ala Glu His Gly Ala Glu Pro Pro
35 40 45
Ala Pro Glu Pro Ser Ala Gly Ala Ser Ser Asn Trp Thr Thr Leu
50 55 60
Pro Pro Pro Leu Phe Ser Lys Val Val Ile Val Leu Ile Asp Ala
65 70 75
Leu Arg Asp Asp Phe Val Phe Gly Ser Lys Gly Val Lys Phe Met
80 85 90
Pro Tyr Thr Thr Tyr Leu Val Glu Lys Gly Ala Ser His Ser Phe
95 100 105
Val Ala Glu Ala Lys Pro Pro Thr Val Thr Met Pro Arg Ile Lys
110 115 120
Ala Leu Met Thr Gly Ser Leu Pro Gly Phe Val Asp Val Ile Arg
125 130 135
Asn Leu Asn Ser Pro Ala Leu Leu Glu Asp Ser Val Ile Arg Gln
140 145 150
Ala Lys Ala Ala Gly Lys Arg Ile Val Phe Tyr Gly Asp Glu Thr
155 160 165
Trp Val Lys Leu Phe Pro Lys His Phe Val Glu Tyr Asp Gly Thr
170 175 180
Thr Ser Phe Phe Val Ser Asp Tyr Thr Glu Val Asp Asn Asn Val
185 190 195
Thr Arg His Leu Asp Lys Val Leu Lys Arg Gly Asp Trp Asp Ile
200 205 210

-continued

Leu Ile Leu His Tyr Leu Gly Leu Asp His Ile Gly His Ile Ser
215 220 225
Gly Pro Asn Ser Pro Leu Ile Gly Gln Lys Leu Ser Glu Met Asp
230 235 240
Ser Val Leu Met Lys Ile His Thr Ser Leu Gln Ser Lys Glu Arg
245 250 255
Glu Thr Pro Leu Pro Asn Leu Leu Val Leu Cys Gly Asp His Gly
260 265 270
Met Ser Glu Thr Gly Ser His Gly Ala Ser Ser Thr Glu Glu Val
275 280 285
Asn Thr Pro Leu Ile Leu Ile Ser Ser Ala Phe Glu Arg Lys Pro
290 295 300
Gly Asp Ile Arg His Pro Lys His Val Gln
305 310

<210> SEQ ID NO 141
<211> LENGTH: 754
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 141

ggcacgaggc aagccttcca ggttatcgtg acgcaccttg aaagtctgag 50
agctactgcc ctacagaaag ttactagtgc cctaaagctg gcgctggcac 100
tgatgttact gctgctgttg gagtacaact tccctataga aaacaactgc 150
cagcacctta agaccactca caccttcaga gtgaagaact taaacccgaa 200
gaaattcagc attcatgacc aggatcacaa agtactggtc ctggactctg 250
ggaatctcat agcagttcca gataaaaact acatacgccc agagatcttc 300
tttgcattag cctcatcctt gagctcagcc totgcggaga aaggaagtcc 350
gattctcctg ggggtctcta aaggggagtt ttgtctctac tgtgacaagg 400
ataaaggaca aagtcattcca tcccttcagc tgaagaagga gaaactgatg 450
aagctggctg cccaaaagga atcagcacgc cggcccttca tcttttatag 500
ggctcaggty ggctcctgga acatgctgga gtcggcggct caccocggat 550
ggttcatctg cacctcctgc aattgtaatg agcctgttgg ggtgacagat 600
aaatttgaga acaggaaaca cattgaattt tcatttcaac cagtttgcaa 650
agctgaaatg agccccagtg aggtcagcga ttaggaaact gccccattga 700
acgccttcct cgctaatttg aactaattgt ataaaaacac caaacctgct 750
cact 754

<210> SEQ ID NO 142
<211> LENGTH: 193
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 142

Met Leu Leu Leu Leu Glu Tyr Asn Phe Pro Ile Glu Asn Asn
1 5 10 15
Cys Gln His Leu Lys Thr Thr His Thr Phe Arg Val Lys Asn Leu
20 25 30

-continued

Asn	Pro	Lys	Lys	Phe	Ser	Ile	His	Asp	Gln	Asp	His	Lys	Val	Leu
			35						40					45
Val	Leu	Asp	Ser	Gly	Asn	Leu	Ile	Ala	Val	Pro	Asp	Lys	Asn	Tyr
			50						55					60
Ile	Arg	Pro	Glu	Ile	Phe	Phe	Ala	Leu	Ala	Ser	Ser	Leu	Ser	Ser
			65						70					75
Ala	Ser	Ala	Glu	Lys	Gly	Ser	Pro	Ile	Leu	Leu	Gly	Val	Ser	Lys
			80						85					90
Gly	Glu	Phe	Cys	Leu	Tyr	Cys	Asp	Lys	Asp	Lys	Gly	Gln	Ser	His
			95						100					105
Pro	Ser	Leu	Gln	Leu	Lys	Lys	Glu	Lys	Leu	Met	Lys	Leu	Ala	Ala
			110						115					120
Gln	Lys	Glu	Ser	Ala	Arg	Arg	Pro	Phe	Ile	Phe	Tyr	Arg	Ala	Gln
			125						130					135
Val	Gly	Ser	Trp	Asn	Met	Leu	Glu	Ser	Ala	Ala	His	Pro	Gly	Trp
			140						145					150
Phe	Ile	Cys	Thr	Ser	Cys	Asn	Cys	Asn	Glu	Pro	Val	Gly	Val	Thr
			155						160					165
Asp	Lys	Phe	Glu	Asn	Arg	Lys	His	Ile	Glu	Phe	Ser	Phe	Gln	Pro
			170						175					180
Val	Cys	Lys	Ala	Glu	Met	Ser	Pro	Ser	Glu	Val	Ser	Asp		
			185						190					

<210> SEQ ID NO 143
<211> LENGTH: 961
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 143

ctagagagta tagggcagaa ggatggcaga tgagtgactc cacatccaga	50
gctgcctccc tttaatccag gatcctgtcc ttctgtcct gtaggagtgc	100
ctgttgccag tgtgggtga gacaagtttg tcccacaggg ctgtctgagc	150
agataagatt aagggtggg tctgtgctca attaactcct gtgggcacgg	200
gggctgggaa gagcaaagtc agcggtcct acagtcagca ccatgctggg	250
cctgccgtgg aagggaggtc tgtcctgggc gctgctgctg cttctcttag	300
gtccccagat cctgctgata tatgcctggc attccacga gcaaaggac	350
tgtgatgaac acaatgtcat ggctcggtac ctccctgcc cagtggagtt	400
tgctgtccac acattcaacc aacagagcaa ggactactat gcctacagac	450
tggggcacat cttgaattcc tggaaggagc aggtggagtc caagactgta	500
ttctcaatgg agctactgct ggggagaact aggtgtggga aatttgaaga	550
cgacattgac aactgccatt tccaagaaag cacagagctg aacaatactt	600
tcacctgctt cttcaccatc agcaccagc cctggatgac tcagttcagc	650
ctcctgaaca agacctgctt ggagggatc cactgagtga aaccactca	700
caggcttgct catgtgctgc tcccacattc cgtggacatc agcactactc	750
tcctgaggac tcttcagtgg ctgagcagct ttggacttgt ttgttatcct	800
atcttgcatg tgtttgagat ctcagatcag tgttttagaa aatccacaca	850
tccttgagcct aatcatgtag tgtagatcat taaacatcag cattttaaga	900

-continued

aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	950
aaaaaaaaa a	961
<210> SEQ ID NO 144	
<211> LENGTH: 147	
<212> TYPE: PRT	
<213> ORGANISM: Homo Sapien	
<400> SEQUENCE: 144	
Met Leu Gly Leu Pro Trp Lys Gly Gly Leu Ser Trp Ala Leu Leu	
1 5 10 15	
Leu Leu Leu Leu Gly Ser Gln Ile Leu Leu Ile Tyr Ala Trp His	
20 25 30	
Phe His Glu Gln Arg Asp Cys Asp Glu His Asn Val Met Ala Arg	
35 40 45	
Tyr Leu Pro Ala Thr Val Glu Phe Ala Val His Thr Phe Asn Gln	
50 55 60	
Gln Ser Lys Asp Tyr Tyr Ala Tyr Arg Leu Gly His Ile Leu Asn	
65 70 75	
Ser Trp Lys Glu Gln Val Glu Ser Lys Thr Val Phe Ser Met Glu	
80 85 90	
Leu Leu Leu Gly Arg Thr Arg Cys Gly Lys Phe Glu Asp Asp Ile	
95 100 105	
Asp Asn Cys His Phe Gln Glu Ser Thr Glu Leu Asn Asn Thr Phe	
110 115 120	
Thr Cys Phe Phe Thr Ile Ser Thr Arg Pro Trp Met Thr Gln Phe	
125 130 135	
Ser Leu Leu Asn Lys Thr Cys Leu Glu Gly Phe His	
140 145	
<210> SEQ ID NO 145	
<211> LENGTH: 1157	
<212> TYPE: DNA	
<213> ORGANISM: Homo Sapien	
<400> SEQUENCE: 145	
ctgtgcagct cgaggctcca gaggcacact ccagagagag ccaaggttct	50
gacgcgatga ggaagcacct gagctggtgg tggctggcca ctgtctgcat	100
gctgctcttc agccacctct ctgcggtcca gacgaggggc atcaagcaca	150
gaatcaagtg gaaccggaag gccctgccca gcaactgccca gatcaactgag	200
gcccaggtgg ctgagaaccg cccgggagcc ttcatcaagc aaggccgcaa	250
gctcgacatt gacttcggag ccgagggcaa caggtactac gaggccaact	300
actggcagtt ccccgatggc atccactaca acggtgctc tgaggctaatt	350
gtgaccaagg aggcatttgt caccgctgc atcaatgccca ccaggcgggc	400
gaaccagggg gagtccaga agccagacaa caagctccac cagcaggtgc	450
tctggcggct ggtccaggag ctctgctccc toaagcattg cgagtttttg	500
ttggagaggg gcgcaggact tcgggtcacc atgcaccagc cagtgtcct	550
ctgccttctg gctttgatct ggctcatggt gaaataagct tgccaggagg	600
ctggcagtag agagcgcagc agcgagcaaa toctggcaag tgaccagct	650

-continued

cttctccccc aaaccacgc gtgttctgaa ggtgcccagg agcggcgatg	700
cactcgcact gcaaatgccg ctcccacgta tgcgccctgg tatgtgcttg	750
cgttctgata gatgggggac tgttgcttct ccgtcactcc attctcagcc	800
cctagcagag cgtctggcac actagattag tagtaaatgc ttgatgagaa	850
gaacacatca ggcactgcgc cacctgcttc acagtacttc ccaacaactc	900
ttagaggtag gtgtattccc gttttacaga taaggaaact gagggccaga	950
gagctgaagt actgcaccca gcatcaccag ctagaaagtg gcagagccag	1000
gattcaaccc tggcttgtct aaccccaggt tttctgctct gtccaattcc	1050
agagctgtct ggtgatcact ttatgtctca cagggaccca catccaaaca	1100
tgtatctcta atgaaattgt gaaagctcca tgtttagaaa taaatgaaaa	1150
cacctga	1157

<210> SEQ ID NO 146
<211> LENGTH: 176
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 146

Met Arg Lys His Leu Ser Trp Trp Trp Leu Ala Thr Val Cys Met	1 5 10 15
Leu Leu Phe Ser His Leu Ser Ala Val Gln Thr Arg Gly Ile Lys	20 25 30
His Arg Ile Lys Trp Asn Arg Lys Ala Leu Pro Ser Thr Ala Gln	35 40 45
Ile Thr Glu Ala Gln Val Ala Glu Asn Arg Pro Gly Ala Phe Ile	50 55 60
Lys Gln Gly Arg Lys Leu Asp Ile Asp Phe Gly Ala Glu Gly Asn	65 70 75
Arg Tyr Tyr Glu Ala Asn Tyr Trp Gln Phe Pro Asp Gly Ile His	80 85 90
Tyr Asn Gly Cys Ser Glu Ala Asn Val Thr Lys Glu Ala Phe Val	95 100 105
Thr Gly Cys Ile Asn Ala Thr Gln Ala Ala Asn Gln Gly Glu Phe	110 115 120
Gln Lys Pro Asp Asn Lys Leu His Gln Gln Val Leu Trp Arg Leu	125 130 135
Val Gln Glu Leu Cys Ser Leu Lys His Cys Glu Phe Trp Leu Glu	140 145 150
Arg Gly Ala Gly Leu Arg Val Thr Met His Gln Pro Val Leu Leu	155 160 165
Cys Leu Leu Ala Leu Ile Trp Leu Met Val Lys	170 175

<210> SEQ ID NO 147
<211> LENGTH: 333
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 147

gccttggcct cccaaagggc tgggattata ggcgtgacca ccatgtctgg	50
--	----

-continued

tccagagtct catttcctga tgatttatag actcaaagaa aactcatgtt	100
cagaagctct cttctcttct ggctcctct ctgtcttctt tccctctttc	150
ttcttatttt aattagtagc atctactcag agtcatgcaa gctggaaatc	200
tttcattttg cttgtcagtg gggtaggtca ctgagtctta gtttttattt	250
tttgaaattt caactttcag attcaggggg tacatgtgaa ggtttgtttt	300
atgagtatat tgcatgatgc tgaggtttgg ggt	333

<210> SEQ ID NO 148
<211> LENGTH: 73
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 148

Met	Phe	Arg	Ser	Ser	Leu	Leu	Phe	Trp	Pro	Pro	Leu	Cys	Leu	Leu
1				5					10				15	
Ser	Leu	Phe	Leu	Leu	Ile	Leu	Ile	Ser	Ser	Ile	Tyr	Ser	Glu	Ser
			20					25					30	
Cys	Lys	Leu	Glu	Ile	Phe	His	Phe	Ala	Cys	Gln	Trp	Gly	Arg	Ser
			35					40					45	
Leu	Ser	Leu	Ser	Phe	Tyr	Phe	Leu	Lys	Phe	Gln	Leu	Ser	Asp	Ser
			50					55					60	
Gly	Gly	Thr	Cys	Glu	Gly	Leu	Phe	Tyr	Glu	Tyr	Ile	Ala		
			65					70						

<210> SEQ ID NO 149
<211> LENGTH: 1893
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 149

gtctccgcgt cacaggaact tcagcaccca cagggcggac agcgctcccc	50
tctacctgga gacttgactc ccgcgcgccc caaccctgct tatcccttga	100
ccgtcgagtg tcagagatcc tgcagccgcc cagtcccgcc ccctctcccg	150
ccccacacc accctcctgg ctcttcctgt ttttactcct ccttttcatt	200
cataacaaaa gctacagctc caggagccca gcgccgggct gtgacccaag	250
ccgagcgtgg aagaatgggg ttctctcgga ccggcacttg gattctggtg	300
ttagtgctcc cgattcaagc tttcccaaaa cctggaggaa gccaaagaaa	350
atctctacat aatagagaat taagtgcaga aagaccttg aatgaacaga	400
ttgctgaagc agaagaagac aagattaaaa aaacatatcc tccagaaaac	450
aagccaggtc agagcaacta ttcttttgtt gataactga acctgctaaa	500
ggcaataaca gaaaaggaaa aaattgagaa agaaagacaa tctataagaa	550
gtccccact tgataataag ttgaatgtgg aagatgttga ttcaaccaag	600
aatcgaaaa tgatcgatga ttatgactct actaagagtg gattggatca	650
taaatttcaa gatgatccag atggctctta tcaactagac gggactcctt	700
taaccgctga agacattgtc cataaaatcg ctgccaggat ttatgaagaa	750
aatgacagag ccgtgtttga caagattgtt tctaaactac ttaatctcgg	800

```

<210> SEQ ID NO 150
<211> LENGTH: 468
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 150

Met Gly Phe Leu Gly Thr Gly Thr Trp Ile Leu Val Leu Val Leu
  1           5           10           15

Pro Ile Gln Ala Phe Pro Lys Pro Gly Gly Ser Gln Asp Lys Ser
          20           25           30

Leu His Asn Arg Glu Leu Ser Ala Glu Arg Pro Leu Asn Glu Gln
          35           40           45

Ile Ala Glu Ala Glu Glu Asp Lys Ile Lys Lys Thr Tyr Pro Pro
          50           55           60

Glu Asn Lys Pro Gly Gln Ser Asn Tyr Ser Phe Val Asp Asn Leu
          65           70           75

Asn Leu Leu Lys Ala Ile Thr Glu Lys Glu Lys Ile Glu Lys Glu
          80           85           90

Arg Gln Ser Ile Arg Ser Ser Pro Leu Asp Asn Lys Leu Asn Val
          95           100          105

Glu Asp Val Asp Ser Thr Lys Asn Arg Lys Leu Ile Asp Asp Tyr
          110          115          120

```

-continued

Asp	Ser	Thr	Lys	Ser	Gly	Leu	Asp	His	Lys	Phe	Gln	Asp	Asp	Pro	125	130	135
Asp	Gly	Leu	His	Gln	Leu	Asp	Gly	Thr	Pro	Leu	Thr	Ala	Glu	Asp	140	145	150
Ile	Val	His	Lys	Ile	Ala	Ala	Arg	Ile	Tyr	Glu	Glu	Asn	Asp	Arg	155	160	165
Ala	Val	Phe	Asp	Lys	Ile	Val	Ser	Lys	Leu	Leu	Asn	Leu	Gly	Leu	170	175	180
Ile	Thr	Glu	Ser	Gln	Ala	His	Thr	Leu	Glu	Asp	Glu	Val	Ala	Glu	185	190	195
Val	Leu	Gln	Lys	Leu	Ile	Ser	Lys	Glu	Ala	Asn	Asn	Tyr	Glu	Glu	200	205	210
Asp	Pro	Asn	Lys	Pro	Thr	Ser	Trp	Thr	Glu	Asn	Gln	Ala	Gly	Lys	215	220	225
Ile	Pro	Glu	Lys	Val	Thr	Pro	Met	Ala	Ala	Ile	Gln	Asp	Gly	Leu	230	235	240
Ala	Lys	Gly	Glu	Asn	Asp	Glu	Thr	Val	Ser	Asn	Thr	Leu	Thr	Leu	245	250	255
Thr	Asn	Gly	Leu	Glu	Arg	Arg	Thr	Lys	Thr	Tyr	Ser	Glu	Asp	Asn	260	265	270
Phe	Glu	Glu	Leu	Gln	Tyr	Phe	Pro	Asn	Phe	Tyr	Ala	Leu	Leu	Lys	275	280	285
Ser	Ile	Asp	Ser	Glu	Lys	Glu	Ala	Lys	Glu	Lys	Glu	Thr	Leu	Ile	290	295	300
Thr	Ile	Met	Lys	Thr	Leu	Ile	Asp	Phe	Val	Lys	Met	Met	Val	Lys	305	310	315
Tyr	Gly	Thr	Ile	Ser	Pro	Glu	Glu	Gly	Val	Ser	Tyr	Leu	Glu	Asn	320	325	330
Leu	Asp	Glu	Met	Ile	Ala	Leu	Gln	Thr	Lys	Asn	Lys	Leu	Glu	Lys	335	340	345
Asn	Ala	Thr	Asp	Asn	Ile	Ser	Lys	Leu	Phe	Pro	Ala	Pro	Ser	Glu	350	355	360
Lys	Ser	His	Glu	Glu	Thr	Asp	Ser	Thr	Lys	Glu	Glu	Ala	Ala	Lys	365	370	375
Met	Glu	Lys	Glu	Tyr	Gly	Ser	Leu	Lys	Asp	Ser	Thr	Lys	Asp	Asp	380	385	390
Asn	Ser	Asn	Pro	Gly	Gly	Lys	Thr	Asp	Glu	Pro	Lys	Gly	Lys	Thr	395	400	405
Glu	Ala	Tyr	Leu	Glu	Ala	Ile	Arg	Lys	Asn	Ile	Glu	Trp	Leu	Lys	410	415	420
Lys	His	Asp	Lys	Lys	Gly	Asn	Lys	Glu	Asp	Tyr	Asp	Leu	Ser	Lys	425	430	435
Met	Arg	Asp	Phe	Ile	Asn	Lys	Gln	Ala	Asp	Ala	Tyr	Val	Glu	Lys	440	445	450
Gly	Ile	Leu	Asp	Lys	Glu	Glu	Ala	Glu	Ala	Ile	Lys	Arg	Ile	Tyr	455	460	465
Ser	Ser	Leu															

<210> SEQ ID NO 151

<211> LENGTH: 2598

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

-continued

<400> SEQUENCE: 151

cggtctcgagg ctcccgccag gagaaaggaa cattctgagg ggagtctaca	50
ccctgtggag ctcaagatgg tcctgagtg ggctgtgc ttccgaatga	100
aggactcggc attgaaggtg ctttatctgc ataataacca gcttctagct	150
ggagggctgc atgcaggga ggtcattaaa ggtgaagaga tcagcgtggt	200
ccccaatcgg tggctggatg ccagcctgtc ccccgctcgc ctgggtgtcc	250
agggtggaag ccagtgcctg tcatgtggg tggggcagga gccgactcta	300
acactagagc cagtgaacat catggagctc tatcttggtg ccaaggaatc	350
caagagcttc accttctacc ggccggacat ggggtcacc tccagcttcg	400
agtcggctgc ctacccgggc tggttcctgt gcacgggtgc tgaagccgat	450
cagcctgtca gactcaccca gcttcccgag aatggtggct ggaatgcccc	500
catcacagac ttctacttcc agcagtgtga ctaggccaac gtgccccca	550
gaactccctg ggcagagcca gctcgggtga ggggtgagtg gaggagacct	600
atggcggaac atcactctct ctgctctcag gacccccacg tctgacttag	650
tgggcacctg accactttgt cttctgggtc ccagtgtgga taaattctga	700
gatttggagc tcagtccacg gtccctcccc actggatggt gctactgctg	750
tggaaacctg taaaaacat gtggggtaaa ctgggaataa catgaaaaga	800
tttctgtggg ggtgggtggt gggagtgtg ggaatcattc ctgcttaatg	850
gtaactgaca agtgttaccc tgagccccgc aggccaaacc atccccagtt	900
gagccttata gggtcagtag ctctccacat gaagtcctgt cactcaccac	950
tgtgcaggag agggaggtg tcatagatgc agggatctat ggccttggtc	1000
ccagccccac ccccttccct ttaatcctgc cactgtcata tgetaccttt	1050
cctatctctt ccctcatcat ctgtgtgtg goatgaggag gtggtgatgt	1100
cagaagaaat ggctcgagct cagaagataa aagataagta gggtagctg	1150
atcctctttt aaaaacccaa gatacaatca aaatcccaga tgetgtgtct	1200
tattcccatg aaaaagtgt catgacatat tgagaagacc tacttacaaa	1250
gtggcatata ttgcaattta ttttaattaa aagataccta tttatatatt	1300
tctttataga aaaaagtctg gaagagtta cttcaattgt agcaatgtca	1350
gggtggtggt agtatagggt atttttcttt taattctgtt aatttatctg	1400
tatttcctaa tttttctaca atgaagatga attccttgta taaaaataag	1450
aaaagaaatt aatcttgagg taagcagagc agacatcgc tctgattgtc	1500
ctcagcctcc acttccccag agtaaatca aattgaatcg agctctgctg	1550
ctctggttgg ttgtagtagt gatcaggaag cagatctcag caaagccact	1600
gaggaggagg ctgtgctgag ttgtgtgtgc tggaatctct gggtaaggaa	1650
cttaagaac aaaaatcgc tggtaattct ttcctagaag gatcacagcc	1700
cctgggattc caaggcattg gatccagtct ctaagaaggc tgctgtactg	1750
gttgaattgt gtccccctca aattcacatc cttcttgtaa tctcagtctg	1800
tgagtttatt tggagataag gtctctgcag atgtagttag ttaagacaag	1850

-continued

gtcatgctgg atgaaggtag acctaaattc aatatgactg gtttccttgt 1900
atgaaaagga gaggacacag agacagagga gacgcgggga agactatgta 1950
aagatgaagg cagagatcgg agttttgcag ccacaagcta agaaacacca 2000
aggattgtgg caaccatcag aagcttgga gaggcaaaga agaattcttc 2050
cctagagggt ttagagggat aacggctctg ctgaaacctt aatctcagac 2100
ttccagcctc ctgaacgaag aaagaataaa tttcggctgt ttttaagccac 2150
caagataat tggttacagc agctctagga aactaataca gctgctaaaa 2200
tgatccctgt ctctcgtgt ttacattctg tgtgtgtccc ctcccacaat 2250
gtaccaaagt tgtctttgtg accaatagaa tatggcagaa gtgatggcat 2300
gccacttcca agattaggtt ataaaagaca ctgcagcttc tacttgagcc 2350
ctctctctct gccaccacc gccccaatc tatcttggt cactcgtctt 2400
gggggaagct agctgccatg ctatgagcag gcctataaag agacttacgt 2450
ggtaaaaaat gaagtctcct gccacagcc acattagtga acctagaagc 2500
agagactctg tgagataatc gatgtttgtt gttttaagtt gctcagtttt 2550
ggtctaactt gttatgcagc aatagataaa taatatgcag agaaagag 2598

<210> SEQ ID NO 152
<211> LENGTH: 155
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 152

Met Val Leu Ser Gly Ala Leu Cys Phe Arg Met Lys Asp Ser Ala
1 5 10 15
Leu Lys Val Leu Tyr Leu His Asn Asn Gln Leu Leu Ala Gly Gly
20 25 30
Leu His Ala Gly Lys Val Ile Lys Gly Glu Glu Ile Ser Val Val
35 40 45
Pro Asn Arg Trp Leu Asp Ala Ser Leu Ser Pro Val Ile Leu Gly
50 55 60
Val Gln Gly Gly Ser Gln Cys Leu Ser Cys Gly Val Gly Gln Glu
65 70 75
Pro Thr Leu Thr Leu Glu Pro Val Asn Ile Met Glu Leu Tyr Leu
80 85 90
Gly Ala Lys Glu Ser Lys Ser Phe Thr Phe Tyr Arg Arg Asp Met
95 100 105
Gly Leu Thr Ser Ser Phe Glu Ser Ala Ala Tyr Pro Gly Trp Phe
110 115 120
Leu Cys Thr Val Pro Glu Ala Asp Gln Pro Val Arg Leu Thr Gln
125 130 135
Leu Pro Glu Asn Gly Gly Trp Asn Ala Pro Ile Thr Asp Phe Tyr
140 145 150
Phe Gln Gln Cys Asp
155

<210> SEQ ID NO 153
<211> LENGTH: 1152
<212> TYPE: DNA

-continued

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 153

```
cttcagaaca ggttctcctt ccccgagtcac cagttgctcg agttagaatt      50
gtctgcaatg gccgccctgc agaaatctgt gagctctttc cttatgggga      100
ccctggccac cagctgcctc cttctcttgg ccctcttggg acagggagga      150
gcagctgcgc ccatcagctc cactgcagg cttgacaagt ccaacttcca      200
gcagccctat atcaccaacc gcaccttcac gctggctaag gaggctagct      250
tggctgataa caacacagac gttcgtctca ttggggagaa actgttcac      300
ggagtcagta tgagtgagcg ctgctatctg atgaagcagg tgctgaactt      350
cacccttgaa gaagtgtgtg tccctcaatc tgataggttc cagccttata      400
tgcaggaggt ggtgcccttc ctggccaggc tcagcaacag gctaagcaca      450
tgtcatattg aaggtgatga cctgcatatc cagaggaaatg tgcaaaagct      500
gaagcacaca gtgaaaaagc ttggagagag tggagagatc aaagcaattg      550
gagaactgga tttgctgttt atgtctctga gaaatgcctg catttgacca      600
gagcaaaagct gaaaaatgaa taactaacc cctttccctg ctagaaataa      650
caattagatg ccccaaagcg atttttttta accaaaagga agatgggaag      700
ccaaactcca tcatgatggg tggattccaa atgaaccctt gcgttagtta      750
caaaggaaac caatgccact tttgtttata agaccagaag gtagactttc      800
taagcataga tatttattga taacatttca ttgtaactgg tgttctatac      850
acagaaaaca atttattttt taaataattg tctttttcca taaaaaagat      900
tactttccat tccttttaggg gaaaaaaccc ctaaatagct tcatgtttcc      950
ataatcagta ctttatattt ataatgtat ttattattat tataagactg     1000
cattttattt atatcatttt attaatatgg atttatttat agaaacatca     1050
ttcgatattg ctacttgagt gtaaggctaa tattgatatt tatgacaata     1100
attatagagc tataacatgt ttatttgacc tcaataaaca cttggatatc     1150
cc                                                                1152
```

<210> SEQ ID NO 154

<211> LENGTH: 179

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 154

```
Met Ala Ala Leu Gln Lys Ser Val Ser Ser Phe Leu Met Gly Thr
  1             5             10             15

Leu Ala Thr Ser Cys Leu Leu Leu Leu Ala Leu Leu Val Gln Gly
 20             25             30

Gly Ala Ala Ala Pro Ile Ser Ser His Cys Arg Leu Asp Lys Ser
 35             40             45

Asn Phe Gln Gln Pro Tyr Ile Thr Asn Arg Thr Phe Met Leu Ala
 50             55             60

Lys Glu Ala Ser Leu Ala Asp Asn Asn Thr Asp Val Arg Leu Ile
 65             70             75

Gly Glu Lys Leu Phe His Gly Val Ser Met Ser Glu Arg Cys Tyr
```

-continued

80										85					90				
Leu	Met	Lys	Gln	Val	Leu	Asn	Phe	Thr	Leu	Glu	Glu	Val	Leu	Phe					
				95					100					105					
Pro	Gln	Ser	Asp	Arg	Phe	Gln	Pro	Tyr	Met	Gln	Glu	Val	Val	Pro					
				110					115					120					
Phe	Leu	Ala	Arg	Leu	Ser	Asn	Arg	Leu	Ser	Thr	Cys	His	Ile	Glu					
				125					130					135					
Gly	Asp	Asp	Leu	His	Ile	Gln	Arg	Asn	Val	Gln	Lys	Leu	Lys	Asp					
				140					145					150					
Thr	Val	Lys	Lys	Leu	Gly	Glu	Ser	Gly	Glu	Ile	Lys	Ala	Ile	Gly					
				155					160					165					
Glu	Leu	Asp	Leu	Leu	Phe	Met	Ser	Leu	Arg	Asn	Ala	Cys	Ile						
				170					175										

<210> SEQ ID NO 155
<211> LENGTH: 1320
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 155

ggcttgctga aaataaaatc aggactccta acctgctcca gtcagcctgc	50
ttccacgagg cctgtcagtc agtgcccgac ttgtgactga gtgtgcagtg	100
cccagcatgt accaggtcag tgcagagggc tgcctgaggg ctgtgctgag	150
agggagagga gcagagatgc tgctgagggg ggagggaggg caagctgcca	200
ggtttggggg tggggggcaa gtggagttag aaactgggat cccaggggga	250
gggtgcagat gagggagcga cccagattag gtgaggacag ttctctcatt	300
agccttttcc tacaggtggt tgcattcttg gcaatggtca tgggaaccca	350
cacctacagc cactggccca gctgtgccc cagcaaaggg caggacacct	400
ctgaggagct gctgaggtgg agcactgtgc ctgtgcctcc cctagagcct	450
gctaggccca accgccaccc agagtccctgt agggccagtg aagatggacc	500
cctcaacagc agggccatct cccctggag atatgagttg gacagagact	550
tgaaccggct cccccaaggac ctgtaccagc ccggttgctt gtgccgcac	600
tgcgtcagcc tacagacagg ctcccacatg gacccccggg gcaactcgga	650
gctgctctac cacaaccaga ctgtcttcta caggcggcca tgccatggcg	700
agaagggcac ccacaagggc tactgcctgg agcgcaggct gtaccgtgtt	750
tccttagctt gtgtgtgtgt gcggccccgt gtgatgggct agccggacct	800
gctggaggct ggtccctttt tgggaaacct ggagccaggt gtacaaccac	850
ttgccatgaa gggccaggat gccagatgc ttggccctg tgaagtgtg	900
tctggagcag caggatcccc ggacaggatg gggggctttg gggaaaacct	950
gcacttctgc acattttgaa aagagcagct gctgcttagg gccgccgga	1000
gctggtgtcc tgtcattttc tctcaggaaa ggttttcaa gttctgccca	1050
tttctggagg ccaccactcc tgtctcttcc tcttttccca tccccgcta	1100
ccctggccca gcacaggcac tttctagata tttccccctt gctggagaag	1150
aaagagcccc tggttttatt tgtttgttta ctcatcactc agtgagcatc	1200

-continued

tactttgggt gcattctagt gtagttacta gtcttttgac atggatgatt	1250
ctgaggagga agctgttatt gaatgtatag agatttatcc aaataaatat	1300
ctttatttaa aaatgaaaaa	1320

<210> SEQ ID NO 156
<211> LENGTH: 177
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 156

Met Arg Glu Arg Pro Arg Leu Gly Glu Asp Ser Ser Leu Ile Ser	
1 5 10 15	
Leu Phe Leu Gln Val Val Ala Phe Leu Ala Met Val Met Gly Thr	
20 25 30	
His Thr Tyr Ser His Trp Pro Ser Cys Cys Pro Ser Lys Gly Gln	
35 40 45	
Asp Thr Ser Glu Glu Leu Leu Arg Trp Ser Thr Val Pro Val Pro	
50 55 60	
Pro Leu Glu Pro Ala Arg Pro Asn Arg His Pro Glu Ser Cys Arg	
65 70 75	
Ala Ser Glu Asp Gly Pro Leu Asn Ser Arg Ala Ile Ser Pro Trp	
80 85 90	
Arg Tyr Glu Leu Asp Arg Asp Leu Asn Arg Leu Pro Gln Asp Leu	
95 100 105	
Tyr His Ala Arg Cys Leu Cys Pro His Cys Val Ser Leu Gln Thr	
110 115 120	
Gly Ser His Met Asp Pro Arg Gly Asn Ser Glu Leu Leu Tyr His	
125 130 135	
Asn Gln Thr Val Phe Tyr Arg Arg Pro Cys His Gly Glu Lys Gly	
140 145 150	
Thr His Lys Gly Tyr Cys Leu Glu Arg Arg Leu Tyr Arg Val Ser	
155 160 165	
Leu Ala Cys Val Cys Val Arg Pro Arg Val Met Gly	
170 175	

<210> SEQ ID NO 157
<211> LENGTH: 1515
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 157

ccggcgatgt cgctcgtgct gctaagcctg gccgcgctgt gcaggagcgc	50
cgtaccccca gagccgaccg ttcaatgtgg ctctgaaact gggccatctc	100
cagagtggat gctacaacat gatctaatacc ccggagactt gagggacctc	150
cgagtagaac ctgttacaac tagtggtgca acaggggact attcaatttt	200
gatgaatgta agctgggtac tccgggcaga tgccagcatc cgcttggtga	250
aggccaccaaa gatttgtgtg acgggcacaaa gcaacttcca gtcctacagc	300
tgtgtgaggt gcaattacac agaggccttc cagactcaga ccagaccctc	350
tggtggtgtaa tggacatttt cctacatcgg cttccctgta gagctgaaca	400
cagtctatth cattggggcc cataatatcc ctaatgcaaa tatgaatgaa	450

-continued

gatggccctt ccatgtctgt gaatttcacc tcaccaggct gcctagacca	500
cataatgaaa tataaaaaaa agtgtgtcaa ggccggaagc ctgtgggatc	550
cgaacatcac tgcttctaag aagaatgagg agacagtaga agtgaacttc	600
acaaccactc ccctgggaaa cagatacatg gctcttatcc aacacagcac	650
tatcatcggg ttttctcagg tgtttgagcc acaccagaag aaacaaacgc	700
gagcttcagt ggtgattcca gtgactggg atagtgaagg tgctacggtg	750
cagctgactc catattttcc tacttgtggc agcgactgca tccgacataa	800
aggaacagtt gtgctctgcc cacaaacagg cgtccctttc cctctggata	850
acaacaaaaa caagccggga ggctggctgc ctctcctcct gctgtctctg	900
ctgggtggcca catgggtgct ggtggcagg atctatctaa tgtggaggca	950
cgaaggatc aagaagactt ccttttctac caccacacta ctgcccccca	1000
ttaaggttct tgtggtttac ccatctgaaa tatgtttcca tcacacaatt	1050
tgttacttca ctgaatttct tcaaaacat tgcagaagtg aggtcatcct	1100
tgaaaagtgg cagaaaaaga aaatagcaga gatgggtcca gtgcagtggc	1150
ttgccactca aaagaaggca gcagacaaag tcgtcttcct tctttccaat	1200
gacgtcaaca gtgtgtgcga tggtagctgt ggcaagagcg agggcagtcc	1250
cagtgagaac tctcaagacc tcttccccct tgcctttaac cttttctgca	1300
gtgatctaa aagccagatt catctgcaca aatacgtggt ggtctacttt	1350
agagagattg atacaaaaga cgattacaat gctctcagtg tctgccccaa	1400
gtaccacctc atgaaggatg ccactgcttt ctgtgcagaa cttctccatg	1450
tcaagcagca ggtgtcagca ggaaaaagat cacaagcctg ccacgatggc	1500
tgctgctcct ttag	1515

<210> SEQ ID NO 158

<211> LENGTH: 502

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 158

Met	Ser	Leu	Val	Leu	Leu	Ser	Leu	Ala	Ala	Leu	Cys	Arg	Ser	Ala
1				5						10				15
Val	Pro	Arg	Glu	Pro	Thr	Val	Gln	Cys	Gly	Ser	Glu	Thr	Gly	Pro
			20						25					30
Ser	Pro	Glu	Trp	Met	Leu	Gln	His	Asp	Leu	Ile	Pro	Gly	Asp	Leu
			35						40					45
Arg	Asp	Leu	Arg	Val	Glu	Pro	Val	Thr	Thr	Ser	Val	Ala	Thr	Gly
			50						55					60
Asp	Tyr	Ser	Ile	Leu	Met	Asn	Val	Ser	Trp	Val	Leu	Arg	Ala	Asp
			65						70					75
Ala	Ser	Ile	Arg	Leu	Leu	Lys	Ala	Thr	Lys	Ile	Cys	Val	Thr	Gly
			80						85					90
Lys	Ser	Asn	Phe	Gln	Ser	Tyr	Ser	Cys	Val	Arg	Cys	Asn	Tyr	Thr
			95						100					105
Glu	Ala	Phe	Gln	Thr	Gln	Thr	Arg	Pro	Ser	Gly	Gly	Lys	Trp	Thr
			110						115					120

-continued

Phe	Ser	Tyr	Ile	Gly	Phe	Pro	Val	Glu	Leu	Asn	Thr	Val	Tyr	Phe	125	130	135
Ile	Gly	Ala	His	Asn	Ile	Pro	Asn	Ala	Asn	Met	Asn	Glu	Asp	Gly	140	145	150
Pro	Ser	Met	Ser	Val	Asn	Phe	Thr	Ser	Pro	Gly	Cys	Leu	Asp	His	155	160	165
Ile	Met	Lys	Tyr	Lys	Lys	Lys	Cys	Val	Lys	Ala	Gly	Ser	Leu	Trp	170	175	180
Asp	Pro	Asn	Ile	Thr	Ala	Cys	Lys	Lys	Asn	Glu	Glu	Thr	Val	Glu	185	190	195
Val	Asn	Phe	Thr	Thr	Thr	Pro	Leu	Gly	Asn	Arg	Tyr	Met	Ala	Leu	200	205	210
Ile	Gln	His	Ser	Thr	Ile	Ile	Gly	Phe	Ser	Gln	Val	Phe	Glu	Pro	215	220	225
His	Gln	Lys	Lys	Gln	Thr	Arg	Ala	Ser	Val	Val	Ile	Pro	Val	Thr	230	235	240
Gly	Asp	Ser	Glu	Gly	Ala	Thr	Val	Gln	Leu	Thr	Pro	Tyr	Phe	Pro	245	250	255
Thr	Cys	Gly	Ser	Asp	Cys	Ile	Arg	His	Lys	Gly	Thr	Val	Val	Leu	260	265	270
Cys	Pro	Gln	Thr	Gly	Val	Pro	Phe	Pro	Leu	Asp	Asn	Asn	Lys	Ser	275	280	285
Lys	Pro	Gly	Gly	Trp	Leu	Pro	Leu	Leu	Leu	Leu	Ser	Leu	Leu	Val	290	295	300
Ala	Thr	Trp	Val	Leu	Val	Ala	Gly	Ile	Tyr	Leu	Met	Trp	Arg	His	305	310	315
Glu	Arg	Ile	Lys	Lys	Thr	Ser	Phe	Ser	Thr	Thr	Thr	Leu	Leu	Pro	320	325	330
Pro	Ile	Lys	Val	Leu	Val	Val	Tyr	Pro	Ser	Glu	Ile	Cys	Phe	His	335	340	345
His	Thr	Ile	Cys	Tyr	Phe	Thr	Glu	Phe	Leu	Gln	Asn	His	Cys	Arg	350	355	360
Ser	Glu	Val	Ile	Leu	Glu	Lys	Trp	Gln	Lys	Lys	Lys	Ile	Ala	Glu	365	370	375
Met	Gly	Pro	Val	Gln	Trp	Leu	Ala	Thr	Gln	Lys	Lys	Ala	Ala	Asp	380	385	390
Lys	Val	Val	Phe	Leu	Leu	Ser	Asn	Asp	Val	Asn	Ser	Val	Cys	Asp	395	400	405
Gly	Thr	Cys	Gly	Lys	Ser	Glu	Gly	Ser	Pro	Ser	Glu	Asn	Ser	Gln	410	415	420
Asp	Leu	Phe	Pro	Leu	Ala	Phe	Asn	Leu	Phe	Cys	Ser	Asp	Leu	Arg	425	430	435
Ser	Gln	Ile	His	Leu	His	Lys	Tyr	Val	Val	Val	Tyr	Phe	Arg	Glu	440	445	450
Ile	Asp	Thr	Lys	Asp	Asp	Tyr	Asn	Ala	Leu	Ser	Val	Cys	Pro	Lys	455	460	465
Tyr	His	Leu	Met	Lys	Asp	Ala	Thr	Ala	Phe	Cys	Ala	Glu	Leu	Leu	470	475	480
His	Val	Lys	Gln	Gln	Val	Ser	Ala	Gly	Lys	Arg	Ser	Gln	Ala	Cys	485	490	495
His	Asp	Gly	Cys	Cys	Ser	Leu											

-continued

500

<210> SEQ ID NO 159
<211> LENGTH: 535
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 159

agccaccagc gcaacatgac agtgaagacc ctgcatggcc cagccatggt	50
caagtacttg ctgtgtcgca tattggggct tgcctttctg agtgaggcgg	100
cagctcggaa aatcccaaa gtaggacata cttttttcca aaagcctgag	150
agttgccgc ctgtgccagg aggtagtatg aagcttgaca ttggcatcat	200
caatgaaaac cagcgcgttt ccatgtcacg taacatcgag agccgcgtcca	250
cctccccctg gaattacact gtcacttggg accccaaccg gtaccctctg	300
gaagttgtac aggcccagtg taggaacttg ggctgcatca atgctcaagg	350
aaaggaagac atctccatga attccgttcc catccagcaa gagaccctgg	400
tcgtccggag gaagcaccaa ggctgctctg tttctttcca gttggagaag	450
gtgtgtgtga ctgttggtg cacctgcgtc acccctgtca tccacatgt	500
gcagtaagag gtgcatatcc actcagctga agaag	535

<210> SEQ ID NO 160
<211> LENGTH: 163
<212> TYPE: PRT
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 160

Met Thr Val Lys Thr Leu His Gly Pro Ala Met Val Lys Tyr Leu	
1 5 10 15	
Leu Leu Ser Ile Leu Gly Leu Ala Phe Leu Ser Glu Ala Ala Ala	
20 25 30	
Arg Lys Ile Pro Lys Val Gly His Thr Phe Phe Gln Lys Pro Glu	
35 40 45	
Ser Cys Pro Pro Val Pro Gly Gly Ser Met Lys Leu Asp Ile Gly	
50 55 60	
Ile Ile Asn Glu Asn Gln Arg Val Ser Met Ser Arg Asn Ile Glu	
65 70 75	
Ser Arg Ser Thr Ser Pro Trp Asn Tyr Thr Val Thr Trp Asp Pro	
80 85 90	
Asn Arg Tyr Pro Ser Glu Val Val Gln Ala Gln Cys Arg Asn Leu	
95 100 105	
Gly Cys Ile Asn Ala Gln Gly Lys Glu Asp Ile Ser Met Asn Ser	
110 115 120	
Val Pro Ile Gln Gln Glu Thr Leu Val Val Arg Arg Lys His Gln	
125 130 135	
Gly Cys Ser Val Ser Phe Gln Leu Glu Lys Val Leu Val Thr Val	
140 145 150	
Gly Cys Thr Cys Val Thr Pro Val Ile His His Val Gln	
155 160	

<210> SEQ ID NO 161
<211> LENGTH: 2380

-continued

<212> TYPE: DNA

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 161

acactggcca acaaaaaacg aaagcactcc gtgctggaag taggaggaga	50
gtcaggactc ccaggacaga gagtgcacaa actaccacgc acagccccct	100
ccgccccctc tggaggctga agagggattc cagcccctgc caccacaga	150
cacgggctga ctgggggtgc tgccccctt gggggggggc agcacagggc	200
ctcaggcctg ggtgccacct ggcacctaga agatgcctgt gccctggttc	250
ttgtgtctct tggcactggg ccgaagccca gtggtccttt ctctggagag	300
gcttggtggg cctcaggacg ctaccactg ctctccgggc ctctcctgcc	350
gcctctggga cagtgcata ctctgcctgc ctggggacat cgtgcctgct	400
ccggggcccg tgctggcgcc tacgcacctg cagacagagc tgggtctgag	450
gtgccagaag gagaccgact gtgacctctg tctgcgtgtg gctgtccact	500
tggccgtgca tgggcactgg gaagagcctg aagatgagga aaagtttgga	550
ggagcagctg actcaggggt ggaggagcct aggaatgcct ctctccaggc	600
ccaagtctgt ctctccttcc aggcctaccc tactgcccgc tgcgtcctgc	650
tggagggtgca agtgcctgct gcccttgtgc agtttggtca gtctgtgggc	700
tctgtggtat atgactgctt cgaggctgcc ctagggagtg aggtacgaat	750
ctggtcctat actcagccca ggtacgagaa ggaactcaac cacacacagc	800
agctgcctgc cctgccctgg ctcaacgtgt cagcagatgg tgacaacgtg	850
catctggttc tgaatgtctc tgaggagcag cacttcggcc tctccctgta	900
ctggaatcag gtccagggcc ccccaaaacc ccggtggcac aaaaacctga	950
ctggaccgca gatcattacc ttgaaccaca cagacctggt tccttgcttc	1000
tgtattcagg tgtggcctct ggaacctgac tccgttagga cgaacatctg	1050
ccccttcagg gaggaccccc gcgcacacca gaacctctgg caagccgccc	1100
gactgcgact gctgacctg cagagctggc tgctggacgc accgtgctcg	1150
ctgcccgcag aagcggcact gtgctggcgg gctccgggtg gggacctg	1200
ccagccactg gtcccaccgc ttctctggga gaacgtcact gtggacaagg	1250
ttctcgagtt ccatttctg aaaggccacc ctaacctctg tgttcagggtg	1300
aacagctcgg agaagctgca gctgcaggag tgcttggtgg ctgactccct	1350
ggggcctctc aaagacgatg tgctactgtt ggagacacga ggccccag	1400
acaacagatc cctctgtgcc ttggaaccga gtggctgtac ttactaccc	1450
agcaaaacct ccacgagggc agctcgctt ggagagtact tactacaaga	1500
cctgcagtca ggccagtgtc tgcagctatg ggacgatgac ttgggagcgc	1550
tatgggcctg ccccatggac aaatacatcc acaagcgctg ggccctcgtg	1600
tggctggcct gcctactctt tgccgtgcg ctttcctca tcctccttct	1650
caaaaaggat cagcgaaaag ggtggctgag gctcttgaaa caggacgtcc	1700
gctcgggggc ggccgccagg ggccgcgcgg ctctgtcct ctactcagcc	1750
gatgactcgg gtttcgagcg cctggtgggc gccctggcgt cggccctgtg	1800

-continued

```

ccagctgccg ctgcgcgtgg ccgtagacct gtggagccgt cgtgaactga      1850
gcgcgcaggg gcccgtaggt tggtttcacg cgcagcggcg ccagaccctg      1900
caggagggcg gcgtggtggt cttgtctctc tctcccggtg cggtggcgct      1950
gtgcagcgag tggctacagg atggggtgtc cgggcccggg gcgcacggcc      2000
cgcacgacgc cttccgcgcc tcgtcagct gcgtgctgcc cgacttcttg      2050
cagggccggg cgcccggcag ctacgtgggg gcctgcttcg acaggctgct      2100
ccacccggac gccgtaccgg ccttttccg caccgtgccc gtcttcacac      2150
tgccctccca actgccagac ttctggggg cctgcagca gcctcgcgcc      2200
ccgcgttcgg ggcggctcca agagagagcg gagcaagtgt cccgggcctt      2250
tcagccagcc ctgatagct acttccatcc cccggggact cccgcgccgg      2300
gacgcggggg gggaccaggg gcgggacctg gggcggggga cgggacttaa      2350
ataaaggcag acgctgtttt tctaaaaaaa      2380

```

<210> SEQ ID NO 162

<211> LENGTH: 705

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 162

```

Met Pro Val Pro Trp Phe Leu Leu Ser Leu Ala Leu Gly Arg Ser
 1             5             10            15

Pro Val Val Leu Ser Leu Glu Arg Leu Val Gly Pro Gln Asp Ala
 20            25            30

Thr His Cys Ser Pro Gly Leu Ser Cys Arg Leu Trp Asp Ser Asp
 35            40            45

Ile Leu Cys Leu Pro Gly Asp Ile Val Pro Ala Pro Gly Pro Val
 50            55            60

Leu Ala Pro Thr His Leu Gln Thr Glu Leu Val Leu Arg Cys Gln
 65            70            75

Lys Glu Thr Asp Cys Asp Leu Cys Leu Arg Val Ala Val His Leu
 80            85            90

Ala Val His Gly His Trp Glu Glu Pro Glu Asp Glu Glu Lys Phe
 95           100           105

Gly Gly Ala Ala Asp Ser Gly Val Glu Glu Pro Arg Asn Ala Ser
110           115           120

Leu Gln Ala Gln Val Val Leu Ser Phe Gln Ala Tyr Pro Thr Ala
125           130           135

Arg Cys Val Leu Leu Glu Val Gln Val Pro Ala Ala Leu Val Gln
140           145           150

Phe Gly Gln Ser Val Gly Ser Val Val Tyr Asp Cys Phe Glu Ala
155           160           165

Ala Leu Gly Ser Glu Val Arg Ile Trp Ser Tyr Thr Gln Pro Arg
170           175           180

Tyr Glu Lys Glu Leu Asn His Thr Gln Gln Leu Pro Ala Leu Pro
185           190           195

Trp Leu Asn Val Ser Ala Asp Gly Asp Asn Val His Leu Val Leu
200           205           210

Asn Val Ser Glu Glu Gln His Phe Gly Leu Ser Leu Tyr Trp Asn

```

-continued

215					220					225				
Gln	Val	Gln	Gly	Pro	Pro	Lys	Pro	Arg	Trp	His	Lys	Asn	Leu	Thr
				230					235					240
Gly	Pro	Gln	Ile	Ile	Thr	Leu	Asn	His	Thr	Asp	Leu	Val	Pro	Cys
				245					250					255
Leu	Cys	Ile	Gln	Val	Trp	Pro	Leu	Glu	Pro	Asp	Ser	Val	Arg	Thr
				260					265					270
Asn	Ile	Cys	Pro	Phe	Arg	Glu	Asp	Pro	Arg	Ala	His	Gln	Asn	Leu
				275					280					285
Trp	Gln	Ala	Ala	Arg	Leu	Arg	Leu	Leu	Thr	Leu	Gln	Ser	Trp	Leu
				290					295					300
Leu	Asp	Ala	Pro	Cys	Ser	Leu	Pro	Ala	Glu	Ala	Ala	Leu	Cys	Trp
				305					310					315
Arg	Ala	Pro	Gly	Gly	Asp	Pro	Cys	Gln	Pro	Leu	Val	Pro	Pro	Leu
				320					325					330
Ser	Trp	Glu	Asn	Val	Thr	Val	Asp	Lys	Val	Leu	Glu	Phe	Pro	Leu
				335					340					345
Leu	Lys	Gly	His	Pro	Asn	Leu	Cys	Val	Gln	Val	Asn	Ser	Ser	Glu
				350					355					360
Lys	Leu	Gln	Leu	Gln	Glu	Cys	Leu	Trp	Ala	Asp	Ser	Leu	Gly	Pro
				365					370					375
Leu	Lys	Asp	Asp	Val	Leu	Leu	Leu	Glu	Thr	Arg	Gly	Pro	Gln	Asp
				380					385					390
Asn	Arg	Ser	Leu	Cys	Ala	Leu	Glu	Pro	Ser	Gly	Cys	Thr	Ser	Leu
				395					400					405
Pro	Ser	Lys	Ala	Ser	Thr	Arg	Ala	Ala	Arg	Leu	Gly	Glu	Tyr	Leu
				410					415					420
Leu	Gln	Asp	Leu	Gln	Ser	Gly	Gln	Cys	Leu	Gln	Leu	Trp	Asp	Asp
				425					430					435
Asp	Leu	Gly	Ala	Leu	Trp	Ala	Cys	Pro	Met	Asp	Lys	Tyr	Ile	His
				440					445					450
Lys	Arg	Trp	Ala	Leu	Val	Trp	Leu	Ala	Cys	Leu	Leu	Phe	Ala	Ala
				455					460					465
Ala	Leu	Ser	Leu	Ile	Leu	Leu	Leu	Lys	Lys	Asp	His	Ala	Lys	Gly
				470					475					480
Trp	Leu	Arg	Leu	Leu	Lys	Gln	Asp	Val	Arg	Ser	Gly	Ala	Ala	Ala
				485					490					495
Arg	Gly	Arg	Ala	Ala	Leu	Leu	Leu	Tyr	Ser	Ala	Asp	Asp	Ser	Gly
				500					505					510
Phe	Glu	Arg	Leu	Val	Gly	Ala	Leu	Ala	Ser	Ala	Leu	Cys	Gln	Leu
				515					520					525
Pro	Leu	Arg	Val	Ala	Val	Asp	Leu	Trp	Ser	Arg	Arg	Glu	Leu	Ser
				530					535					540
Ala	Gln	Gly	Pro	Val	Ala	Trp	Phe	His	Ala	Gln	Arg	Arg	Gln	Thr
				545					550					555
Leu	Gln	Glu	Gly	Gly	Val	Val	Val	Leu	Leu	Phe	Ser	Pro	Gly	Ala
				560					565					570
Val	Ala	Leu	Cys	Ser	Glu	Trp	Leu	Gln	Asp	Gly	Val	Ser	Gly	Pro
				575					580					585
Gly	Ala	His	Gly	Pro	His	Asp	Ala	Phe	Arg	Ala	Ser	Leu	Ser	Cys
				590					595					600

-continued

Val	Leu	Pro	Asp	Phe	Leu	Gln	Gly	Arg	Ala	Pro	Gly	Ser	Tyr	Val
605									610					615
Gly	Ala	Cys	Phe	Asp	Arg	Leu	Leu	His	Pro	Asp	Ala	Val	Pro	Ala
620									625					630
Leu	Phe	Arg	Thr	Val	Pro	Val	Phe	Thr	Leu	Pro	Ser	Gln	Leu	Pro
635									640					645
Asp	Phe	Leu	Gly	Ala	Leu	Gln	Gln	Pro	Arg	Ala	Pro	Arg	Ser	Gly
650									655					660
Arg	Leu	Gln	Glu	Arg	Ala	Glu	Gln	Val	Ser	Arg	Ala	Leu	Gln	Pro
665									670					675
Ala	Leu	Asp	Ser	Tyr	Phe	His	Pro	Pro	Gly	Thr	Pro	Ala	Pro	Gly
680									685					690
Arg	Gly	Val	Gly	Pro	Gly	Ala	Gly	Pro	Gly	Ala	Gly	Asp	Gly	Thr
695									700					705

<210> SEQ ID NO 163
<211> LENGTH: 2478
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 163

gtcagtgccg	gaggccggtc	agccaccaag	atgactgaca	ggttcagctc	50
tctgcagcac	actaccctca	agccacctga	tgtgacctgt	atctccaaag	100
tgagatcgat	tcagatgatt	gttcataccta	ccccacgcc	aatccgtgca	150
ggcgatggcc	accggctaac	cctggaagac	atcttccatg	acctgttcta	200
ccacttagag	ctccaggta	accgcacct	ccaaatgcac	cttgaggagg	250
agcagagaga	atatgagttc	ttcggcctga	cccctgacac	agagttcctt	300
ggcaccatca	tgatttgctg	tcccacctgg	gccaaggaga	gtgccccta	350
catgtgccga	gtgaagacac	tgccagaccg	gacatggacc	tactccttct	400
ccggagcctt	cctgtttctc	atgggcttcc	tcgtcgagct	actctgctac	450
ctgagctaca	gatattgtac	caagccgcct	gcacctccca	actccctgaa	500
cgtccagcga	gtcctgactt	tccagccgct	gcgcttcata	caggagcacg	550
tcctgatccc	tgtctttgac	ctcagcggcc	ccagcagctc	ggcccagcct	600
gtccagtact	cccagatcag	ggtgtcttga	cccagggagc	ccgcaggagc	650
tccacagcgg	catagcctgt	ccgagatcac	ctacttaggg	cagccagaca	700
tctccatcct	ccagccctcc	aacgtgccac	ctccccagat	cctctcccca	750
ctgtcctatg	ccccaaacgc	tgcccctgag	gtcgggcccc	catcctatgc	800
acctcaggtg	acccccgaag	ctcaattccc	attctacgcc	ccacaggcca	850
tctctaaggt	ccagccttcc	tcctatgccc	ctcaagccac	tccggacagc	900
tggcctccct	cctatggggg	atgcatggaa	ggttctggca	aagactcccc	950
cactggggaca	ctttctagtc	ctaaacacct	taggcctaaa	ggtcagcttc	1000
agaaagagcc	accagctgga	agctgcatgt	taggtggcct	ttctctgcag	1050
gaggtgacct	ccttggttat	ggaggaatcc	caagaagcaa	aatcattgca	1100
ccagcccctg	gggatttgca	cagacagaac	atctgaccca	aatgtgctac	1150

-continued

acagtgggga ggaagggaca ccacagtacc taaagggcca gctccccctc	1200
ctctcctcag tccagatcga gggccacccc atgtccctcc ctttgcaacc	1250
tccttcgcgt ccatgttccc cctcggaacca aggtccaagt ccctggggcc	1300
tgctggagtc ccttgtgtgt cccaaggatg aagccaagag cccagcccct	1350
gagacctcag acctggagca gcccacagaa ctggattctc ttttcagagg	1400
cctggccctg actgtgcagt gggagtcctg aggggaatgg gaaaggcttg	1450
gtgcttcctc cctgtcccta cccagtgtca catccttggc tgtcaatccc	1500
atgcctgccc atgccacaca ctctgcgato tggcctcaga cgggtgccct	1550
tgagagaagc agagggagtg gcatgcaggg cccctgccat gggtgcgctc	1600
ctcaccgga caaagcagca tgataaggac tgcagcgggg gagctctggg	1650
gagcagcttg tgtagacaag cgctgtctcg ctgagccctg caaggcagaa	1700
atgacagtgc aaggaggaaa tgcagggaaa ctcccagagt ccagagcccc	1750
acctcctaac accatggatt caaagtgtct aggggaatttg cctctccttg	1800
ccccattcct ggccagtttc acaatctagc tcgacagagc atgaggcccc	1850
tgccctctct gtcatgttcc aaaggtggga agagagcctg gaaaagaacc	1900
aggcctggaa aagaaccaga aggaggtcgg gcagaaccag aacaacctgc	1950
acttctgcca aggccagggc cagcaggacg gcaggactct agggaggggt	2000
gtggcctgca gctcattccc agccagggca actgcctgac gttgcacgat	2050
ttcagcttca ttctctgat agaacaaagc gaaatgcagg tccaccaggg	2100
agggagacac acaagccttt tctgcaggca ggagtttcag accctatcct	2150
gagaatgggg tttgaaagga aggtgagggc tgtggcccct ggacgggtac	2200
aataacacac tgtactgatg tcacaacttt gcaagctctg ccttgggttc	2250
agcccatctg ggctcaaatt ccagcctcac cactcacaag ctgtgtgact	2300
tcaaacaaat gaaatcagtg cccagaacct cggtttcctc atctgtaatg	2350
tggggatcat aacacctacc tcatggagtt gtggtgaaga tgaaatgaag	2400
tcatgtcttt aaagtgttta atagtgcctg gtacatgggc agtgcccaat	2450
aaacggtagc tatttaaaaa aaaaaaaa	2478

<210> SEQ ID NO 164

<211> LENGTH: 574

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 164

Met	Arg	Thr	Leu	Leu	Thr	Ile	Leu	Thr	Val	Gly	Ser	Leu	Ala	Ala
1				5					10					15
His	Ala	Pro	Glu	Asp	Pro	Ser	Asp	Leu	Leu	Gln	His	Val	Lys	Phe
				20					25					30
Gln	Ser	Ser	Asn	Phe	Glu	Asn	Ile	Leu	Thr	Trp	Asp	Ser	Gly	Pro
				35					40					45
Glu	Gly	Thr	Pro	Asp	Thr	Val	Tyr	Ser	Ile	Glu	Tyr	Lys	Thr	Tyr
				50					55					60
Gly	Glu	Arg	Asp	Trp	Val	Ala	Lys	Lys	Gly	Cys	Gln	Arg	Ile	Thr
				65					70					75

-continued

Arg Lys Ser Cys Asn Leu Thr Val Glu Thr Gly Asn Leu Thr Glu	80	85	90
Leu Tyr Tyr Ala Arg Val Thr Ala Val Ser Ala Gly Gly Arg Ser	95	100	105
Ala Thr Lys Met Thr Asp Arg Phe Ser Ser Leu Gln His Thr Thr	110	115	120
Leu Lys Pro Pro Asp Val Thr Cys Ile Ser Lys Val Arg Ser Ile	125	130	135
Gln Met Ile Val His Pro Thr Pro Thr Pro Ile Arg Ala Gly Asp	140	145	150
Gly His Arg Leu Thr Leu Glu Asp Ile Phe His Asp Leu Phe Tyr	155	160	165
His Leu Glu Leu Gln Val Asn Arg Thr Tyr Gln Met His Leu Gly	170	175	180
Gly Lys Gln Arg Glu Tyr Glu Phe Phe Gly Leu Thr Pro Asp Thr	185	190	195
Glu Phe Leu Gly Thr Ile Met Ile Cys Val Pro Thr Trp Ala Lys	200	205	210
Glu Ser Ala Pro Tyr Met Cys Arg Val Lys Thr Leu Pro Asp Arg	215	220	225
Thr Trp Thr Tyr Ser Phe Ser Gly Ala Phe Leu Phe Ser Met Gly	230	235	240
Phe Leu Val Ala Val Leu Cys Tyr Leu Ser Tyr Arg Tyr Val Thr	245	250	255
Lys Pro Pro Ala Pro Pro Asn Ser Leu Asn Val Gln Arg Val Leu	260	265	270
Thr Phe Gln Pro Leu Arg Phe Ile Gln Glu His Val Leu Ile Pro	275	280	285
Val Phe Asp Leu Ser Gly Pro Ser Ser Leu Ala Gln Pro Val Gln	290	295	300
Tyr Ser Gln Ile Arg Val Ser Gly Pro Arg Glu Pro Ala Gly Ala	305	310	315
Pro Gln Arg His Ser Leu Ser Glu Ile Thr Tyr Leu Gly Gln Pro	320	325	330
Asp Ile Ser Ile Leu Gln Pro Ser Asn Val Pro Pro Pro Gln Ile	335	340	345
Leu Ser Pro Leu Ser Tyr Ala Pro Asn Ala Ala Pro Glu Val Gly	350	355	360
Pro Pro Ser Tyr Ala Pro Gln Val Thr Pro Glu Ala Gln Phe Pro	365	370	375
Phe Tyr Ala Pro Gln Ala Ile Ser Lys Val Gln Pro Ser Ser Tyr	380	385	390
Ala Pro Gln Ala Thr Pro Asp Ser Trp Pro Pro Ser Tyr Gly Val	395	400	405
Cys Met Glu Gly Ser Gly Lys Asp Ser Pro Thr Gly Thr Leu Ser	410	415	420
Ser Pro Lys His Leu Arg Pro Lys Gly Gln Leu Gln Lys Glu Pro	425	430	435
Pro Ala Gly Ser Cys Met Leu Gly Gly Leu Ser Leu Gln Glu Val	440	445	450

-continued

Thr	Ser	Leu	Ala	Met	Glu	Glu	Ser	Gln	Glu	Ala	Lys	Ser	Leu	His
				455					460					465
Gln	Pro	Leu	Gly	Ile	Cys	Thr	Asp	Arg	Thr	Ser	Asp	Pro	Asn	Val
				470					475					480
Leu	His	Ser	Gly	Glu	Glu	Gly	Thr	Pro	Gln	Tyr	Leu	Lys	Gly	Gln
				485					490					495
Leu	Pro	Leu	Leu	Ser	Ser	Val	Gln	Ile	Glu	Gly	His	Pro	Met	Ser
				500					505					510
Leu	Pro	Leu	Gln	Pro	Pro	Ser	Gly	Pro	Cys	Ser	Pro	Ser	Asp	Gln
				515					520					525
Gly	Pro	Ser	Pro	Trp	Gly	Leu	Leu	Glu	Ser	Leu	Val	Cys	Pro	Lys
				530					535					540
Asp	Glu	Ala	Lys	Ser	Pro	Ala	Pro	Glu	Thr	Ser	Asp	Leu	Glu	Gln
				545					550					555
Pro	Thr	Glu	Leu	Asp	Ser	Leu	Phe	Arg	Gly	Leu	Ala	Leu	Thr	Val
				560					565					570
Gln	Trp	Glu	Ser											

<210> SEQ ID NO 165
<211> LENGTH: 1060
<212> TYPE: DNA
<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 165

tggcctactg gaaaaaaaaa aaaaaaaaaa aaaagtcacc cgggcccgcg	50
gtggccacaa catggctgcg gcgccggggc tgctcttctg gctgttcgtg	100
ctggggggcg tctggtgggt cccgggccag tcggatctca gccacggacg	150
gcgttttctg gacctcaaag tgtgcgggga cgaagagtgc agcatgttaa	200
tgtaccgtgg gaaagctctt gaagacttca cgggccctga ttgtcgtttt	250
gtgaatttta aaaaaggtga cgatgtatat gtctactaca aactggcagg	300
gggatccctt gaactttggg ctggaagtgt tgaacacagt tttggatatt	350
ttccaaaaga tttgatcaag gtacttcata aatacacgga agaagagcta	400
catattccag cagatgagac agactttgtc tgctttgaag gaggaagaga	450
tgattttaat agttataatg tagaagagct tttaggatct ttggaactgg	500
aggactctgt acctgaagag tcgaagaaaag ctgaagaagt ttctcagcac	550
agagagaaat ctctgagga gtctcggggg cgtgaacttg accctgtgcc	600
tgagcccag gcatccagag ctgattcaga ggatggagaa ggtgctttct	650
cagagagcac cgaggggctg cagggacagc cctcagctca ggagagccac	700
cctcacacca gcggtcctgc ggctaacgct cagggagtgc agtcttcgtt	750
ggacactttt gaagaaattc tgcacgataa attgaaagtg ccgggaagcg	800
aaagcagaac tggcaatagt tctctgcct cggtgagcg ggagaagaca	850
gatgcttaca aagtcctgaa aacagaaatg agtcagagag gaagtggaca	900
gtgcgttatt cattacagca aaggatttcg ttggcatcaa aatctaagtt	950
tgttttacaa agattgtttt tagtactaag ctgccttggc agtttgcatt	1000
tttgagccaa acaaaaatat attattttcc cttctaagta aaaaaaaaaa	1050

-continued

aaaaaaaaaaaa

1060

<210> SEO ID NO 166

```
<211> LENGTH: 303
```

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 166

Met Ala Ala Ala Pro Gly Leu Leu Phe Trp Leu Phe Val Leu Gly
1 5 10 15

Ala Leu Trp Trp Val Pro Gly Gln Ser Asp Leu Ser His Gly Arg
20 25 30

Arg Phe Ser Asp Leu Lys Val Cys Gly Asp Glu Glu Cys Ser Met
35 40 45

Leu Met Tyr Arg Gly Lys Ala Leu Glu Asp Phe Thr Gly Pro Asp
50 55 60

Cys Arg Phe Val Asn Phe Lys Lys Gly Asp Asp Val Tyr Val Tyr
65 70 75

Tyr Lys Leu Ala Gly Gly Ser Leu Glu Leu Trp Ala Gly Ser Val
80 85 90

Glu His Ser Phe Gly Tyr Phe Pro Lys Asp Leu Ile Lys Val Leu
95 100 105

His Lys Tyr Thr Glu Glu Glu Leu His Ile Pro Ala Asp Glu Thr
110 115 120

Asp Phe Val Cys Phe Glu Gly Gly Arg Asp Asp Phe Asn Ser Tyr
125 130 135

Asn Val Glu Glu Leu Leu Gly Ser Leu Glu Leu Glu Asp Ser Val
140 145 150

Pro Glu Glu Ser Lys Lys Ala Glu Glu Val Ser Gln His Arg Glu
155 160 165

Lys Ser Pro Glu Glu Ser Arg Gly Arg Glu Leu Asp Pro Val Pro
170 175 180

Glu Pro Glu Ala Phe Arg Ala Asp Ser Glu Asp Gly Glu Gly Ala
185 190 195

Phe Ser Glu Ser Thr Glu Gly Leu Gln Gly Gln Pro Ser Ala Gln
200 205 210

Glu Ser His Pro His Thr Ser Gly Pro Ala Ala Asn Ala Gln Gly
215 220 225

Val Gln Ser Ser Leu Asp Thr Phe Glu Glu Ile Leu His Asp Lys
230 235 240

Leu Lys Val Pro Gly Ser Glu Ser Arg Thr Gly Asn Ser Ser Pro
245 250 255

Ala Ser Val Glu Arg Glu Lys Thr Asp Ala Tyr Lys Val Leu Lys
260 265 270

Thr Glu Met Ser Gln Arg Gly Ser Gly Gln Cys Val Ile His Tyr
275 280 285

Ser Lys Gly Phe Arg Trp His Gln Asn Leu Ser Leu Phe Tyr Lys
290 295 300

Asp Cys Phe

<210> SEO ID NO 167

```
<210> SEQ ID NO 18
<211> LENGTH: 2570
```

<212> TYPE: DNA

-continued

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 167

ccaggaccag ggcgcaccgg ctccagcctct cacttgctcag aggcggggga	50
agagaagcaa agcgcaacgg tgtgtgtccaa gccgggggctt ctgcttcgcc	100
tctaggacat acacgggacc ccctaacttc agtcccccaa acgcgcaccc	150
tcgaagtctt gaactccagc cccgcacatc cagcgcgggc acaggcgcg	200
caggcggcag gtcccggccg aaggcgatgc gcgcaggggg tcgggcagct	250
gggctcgggc ggcgggagta gggcccgga gggaggcagg gaggctgcat	300
attcagatgc gcgggctgcg ccctggggcag aggcgcgcct cgctccacgc	350
aacacctgct gctgccaccg cgcgcgatg agcgcgctgg tctcgtgct	400
gctggcgccc gcgctgctct gcggccacgg agccttctgc cgcgcgctgg	450
tcagcgccca aaaggtgtgt tttgtgact tcaagcatcc ctgctacaaa	500
atggcctact tccatgaact gtccagccga gtgagctttc aggaggcacg	550
cctggcttgt gagagtgagg gaggagtctt cctcagcctt gagaatgaag	600
cagaacagaa gttaatagag agcatgttgc aaaacctgac aaaacccggg	650
acagggattt ctgatgtgta tttctggata gggctttgga ggaatggaga	700
tgggcaaaac tctggtgcct gccagatct ctaccagtgg tctgatggaa	750
gcaattccca gtaccgaaac tggtagacag atgaaccttc ctgcggaagt	800
gaaaagtgtg ttgtgatgta tcaccaacca actgccaatc ctggccttgg	850
gggtccctac ctttaccagt ggaatgatga cagggtgaac atgaagcaca	900
attatatattg caagtatgaa ccagagatta atccaacagc ccctgtagaa	950
aagccttacc ttacaaatca accaggagac acccatcaga atgtggttgt	1000
tactgaagca ggtataattc ccaatctaata ttatgttgtt ataccaacaa	1050
taccctctgt cttactgata ctggttgctt ttggaacctg ttgtttccag	1100
atgtctcata aaagtaaagg aagaacaaaa actagtccaa accagtctac	1150
actgtggatt tcaaagagta ccagaaaaga aagtgcatg gaagtataat	1200
aactcattga cttggttcca gaattttgta attctggatc tgtataagga	1250
atggcatcag aacaatagct tggaaatggc tgaaatcaca aaggatctgc	1300
aagatgaact gtaagctccc ccttgaggca aatattaaag taatttttat	1350
atgtctatta tttcatttaa agaatatgct gtgctaataa tggagtgaga	1400
catgcttatt ttgctaaagg atgcacccaa acttcaaact tcaagcaaat	1450
gaaatggaca atgcagataa agttgttatc aacacgtcgg gagtatgtgt	1500
gttagaagca attcctttta tttctttcac ctttcataag ttgttatcta	1550
gtcaatgtaa tgtatatgtt attgaaattt acagtgtgca aaagtatatt	1600
acctttgcat aagtgtttga taaaaatgaa ctgttctaata atttattttt	1650
atggcatctc atttttcaat acatgctctt ttgattaaag aaacttatta	1700
ctgttgtcaa ctgaattcac acacacacaa atatagtacc atagaaaaag	1750
tttgttttct cgaaataatt catctttcag cttctctgct tttggtaaat	1800

-continued

gtctaggaaa tctcttcaga aataagaagc tatttcatta agtgtgatat	1850
aaacctcctc aaacatttta ctagaggca aggattgtct aatttcaatt	1900
gtgcaagaca tgtgccttat aattattttt agcttaaaat taaacagatt	1950
ttgtaataat gtaactttgt taataggtgc ataaacacta atgcagtcaa	2000
tttgaacaaa agaagtgaca tacacaatat aaatcatatg tcttcacacg	2050
ttgcctatat aatgagaagc agctctctga gggttctgaa atcaatgtgg	2100
tccctctctt gcccaactaaa caaagatggg tggtcggggg ttgggattga	2150
cactggaggc agatagtgtc aaagttagtc taaggtttcc ctagctgtat	2200
ttagcctctg actatattag tatacaaaga ggtcatgtgg ttgagaccag	2250
gtgaatagtc actatcagtg tggagacaag cacagcacac agacatttta	2300
ggaaggaaaag gaactacgaa atcgtgtgaa aatgggttgg aacccatcag	2350
tgatcgcata ttcatgtatg agggtttgct tgagatagaa aatggtggct	2400
cctttctgtc ttatctccta gtttcttcaa tgcttacgcc ttgttcttct	2450
caagagaaaag ttgtaactct ctggtcttca tatgtccctg tgctcctttt	2500
aaccaataaa agagttcttg tttctggggg aaaaaaaaaa aaaaaaaaaa	2550
aaaaaaaaaa aaaaaaaaaa	2570

<210> SEQ ID NO 168

<211> LENGTH: 273

<212> TYPE: PRT

<213> ORGANISM: Homo Sapien

<400> SEQUENCE: 168

Met Ser Arg Val Val Ser Leu Leu Leu Gly Ala Ala Leu Leu Cys	
1 5 10 15	
Gly His Gly Ala Phe Cys Arg Arg Val Val Ser Gly Gln Lys Val	
20 25 30	
Cys Phe Ala Asp Phe Lys His Pro Cys Tyr Lys Met Ala Tyr Phe	
35 40 45	
His Glu Leu Ser Ser Arg Val Ser Phe Gln Glu Ala Arg Leu Ala	
50 55 60	
Cys Glu Ser Glu Gly Gly Val Leu Leu Ser Leu Glu Asn Glu Ala	
65 70 75	
Glu Gln Lys Leu Ile Glu Ser Met Leu Gln Asn Leu Thr Lys Pro	
80 85 90	
Gly Thr Gly Ile Ser Asp Gly Asp Phe Trp Ile Gly Leu Trp Arg	
95 100 105	
Asn Gly Asp Gly Gln Thr Ser Gly Ala Cys Pro Asp Leu Tyr Gln	
110 115 120	
Trp Ser Asp Gly Ser Asn Ser Gln Tyr Arg Asn Trp Tyr Thr Asp	
125 130 135	
Glu Pro Ser Cys Gly Ser Glu Lys Cys Val Val Met Tyr His Gln	
140 145 150	
Pro Thr Ala Asn Pro Gly Leu Gly Gly Pro Tyr Leu Tyr Gln Trp	
155 160 165	
Asn Asp Asp Arg Cys Asn Met Lys His Asn Tyr Ile Cys Lys Tyr	
170 175 180	

-continued

Glu	Pro	Glu	Ile	Asn	Pro	Thr	Ala	Pro	Val	Glu	Lys	Pro	Tyr	Leu
				185					190					195
Thr	Asn	Gln	Pro	Gly	Asp	Thr	His	Gln	Asn	Val	Val	Val	Thr	Glu
				200					205					210
Ala	Gly	Ile	Ile	Pro	Asn	Leu	Ile	Tyr	Val	Val	Ile	Pro	Thr	Ile
				215					220					225
Pro	Leu	Leu	Leu	Leu	Ile	Leu	Val	Ala	Phe	Gly	Thr	Cys	Cys	Phe
				230					235					240
Gln	Met	Leu	His	Lys	Ser	Lys	Gly	Arg	Thr	Lys	Thr	Ser	Pro	Asn
				245					250					255
Gln	Ser	Thr	Leu	Trp	Ile	Ser	Lys	Ser	Thr	Arg	Lys	Glu	Ser	Gly
				260					265					270
Met	Glu	Val												

<210> SEQ ID NO 169
<211> LENGTH: 43
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide probe

<400> SEQUENCE: 169

tgtaaaacga cggccagtta aatagacctg caattattaa tct 43

<210> SEQ ID NO 170
<211> LENGTH: 41
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic oligonucleotide probe

<400> SEQUENCE: 170

caggaaacag ctatgaccac ctgcacacct gcaaatccat t 41

1. An isolated nucleic acid having at least 80% nucleic acid sequence identity to:

(a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);

(b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;

(c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);

(d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;

(e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);

(f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or

(g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.

2. The isolated nucleic acid of claim 1 having at least 85% nucleic acid sequence identity to:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);

(b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;

(c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);

(d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;

(e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);

(f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or

(g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
3. The isolated nucleic acid of claim 1 having at least 90% nucleic acid sequence identity to:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);

- (b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);
 - (f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or
 - (g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
4. The isolated nucleic acid of claim 1 having at least 95% nucleic acid sequence identity to:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);
 - (f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or
 - (g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
5. The isolated nucleic acid of claim 1 having at least 99% nucleic acid sequence identity to:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);
 - (f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or
 - (g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
6. An isolated nucleic acid comprising:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);
 - (f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or
 - (g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
7. The isolated nucleic acid of claim 6 comprising a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118).
8. The isolated nucleic acid of claim 6 comprising a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide.
9. The isolated nucleic acid of claim 6 comprising a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118).
10. The isolated nucleic acid of claim 6 comprising a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide.
11. The isolated nucleic acid of claim 6 comprising the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117).
12. The isolated nucleic acid of claim 6 comprising the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117).
13. The isolated nucleic acid of claim 6 comprising the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.
14. An isolated nucleic acid that hybridizes to:
- (a) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (b) a nucleic acid sequence encoding the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (c) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118);
 - (d) a nucleic acid sequence encoding the extracellular domain of the polypeptide shown in **FIG. 118** (SEQ ID NO:118), lacking its associated signal peptide;
 - (e) the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117);
 - (f) the full-length coding sequence of the nucleic acid sequence shown in **FIG. 117** (SEQ ID NO:117); or

(g) the full-length coding sequence of the cDNA deposited under ATCC accession number 203363.

15. The isolated nucleic acid of claim 14, wherein said hybridization occurs under stringent conditions.

16. The isolated nucleic acid of claim 14 which is at least 10 nucleotides in length.

17. A vector comprising the nucleic acid of claim 1.

18. The vector of claim 17, wherein said nucleic acid is operably linked to control sequences recognized by a host cell transformed with the vector.

19. A host cell comprising the vector of claim 17.

20. The host cell of claim 19, wherein said cell is a CHO cell, an *E. coli* or a yeast cell.

* * * * *