



US 20110091437A1

(19) **United States**(12) **Patent Application Publication**
Foster et al.(10) **Pub. No.: US 2011/0091437 A1**(43) **Pub. Date: Apr. 21, 2011**(54) **FUSION PROTEINS****Publication Classification**(75) Inventors: **Keith Foster**, Salisbury (GB); **John Chaddock**, Salisbury (GB); **Philip Marks**, Salisbury (GB); **Patrick Stancombe**, Salisbury (GB); **Kei Roger Aoki**, Irvine, CA (US); **Joseph Francis**, Irvine, CA (US); **Lance Steward**, Irvine, CA (US)(51) **Int. Cl.**
A61K 38/48 (2006.01)
C12N 9/96 (2006.01)
C12N 15/57 (2006.01)
C12N 15/63 (2006.01)
C12P 21/02 (2006.01)
A61P 29/00 (2006.01)
(52) **U.S. Cl. 424/94.3**; 435/188; 536/23.2; 435/320.1; 435/69.7(73) Assignees: **SYNTAXIN LIMITED**, Abingdon (GB); **ALLERGAN, INC.**, Irvine, CA (US)(21) Appl. No.: **12/868,510**(22) Filed: **Aug. 25, 2010****Related U.S. Application Data**

(63) Continuation-in-part of application No. 12/303,078, filed on Sep. 21, 2009, filed as application No. PCT/GB2007/002049 on Jun. 1, 2007.

(30) **Foreign Application Priority Data**

Jun. 1, 2006 (GB) 0610867.4

ABSTRACT

A single chain, polypeptide fusion protein, comprising: a non-cytotoxic protease, or a fragment thereof, which protease or protease fragment is capable of cleaving a protein of the exocytic fusion apparatus of a nociceptive sensory afferent; a dynorphin Targeting Moiety that is capable of binding to a Binding Site on the nociceptive sensory afferent, which Binding Site is capable of undergoing endocytosis to be incorporated into an endosome within the nociceptive sensory afferent; a protease cleavage site at which site the fusion protein is cleavable by a protease, wherein the protease cleavage site is located between the non-cytotoxic protease or fragment thereof and the dynorphin Targeting Moiety; and a translocation domain that is capable of translocating the protease or protease fragment from within an endosome, across the endosomal membrane and into the cytosol of the nociceptive sensory afferent. Nucleic acid sequences encoding the polypeptide fusion proteins, methods of preparing same and uses thereof are also described.

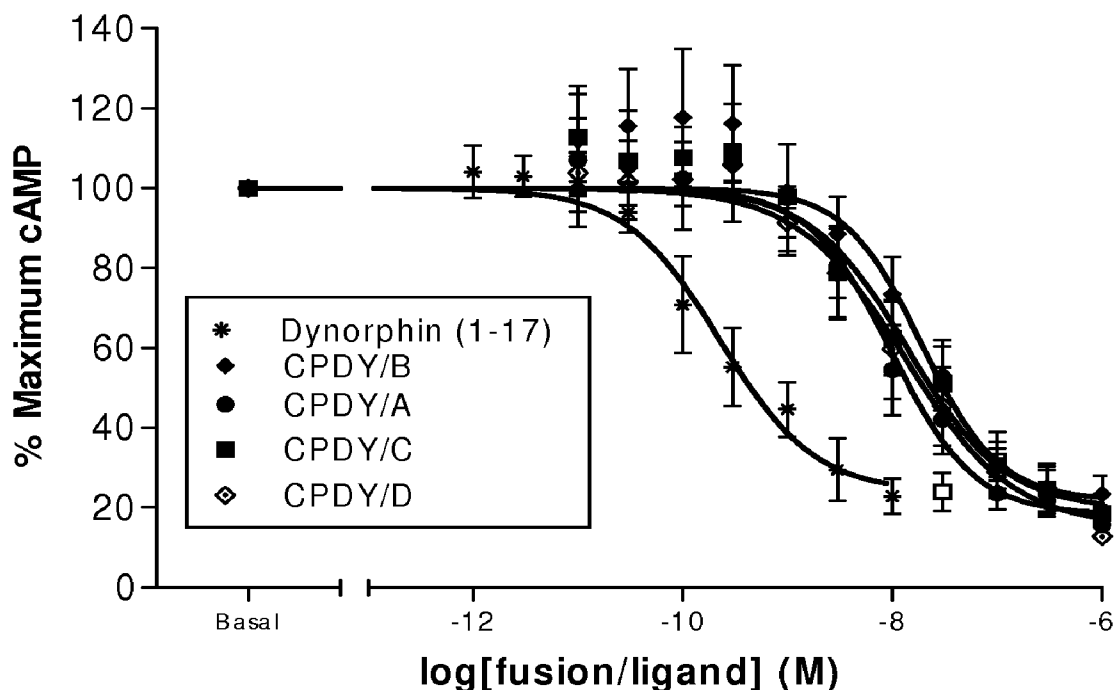


Figure 1

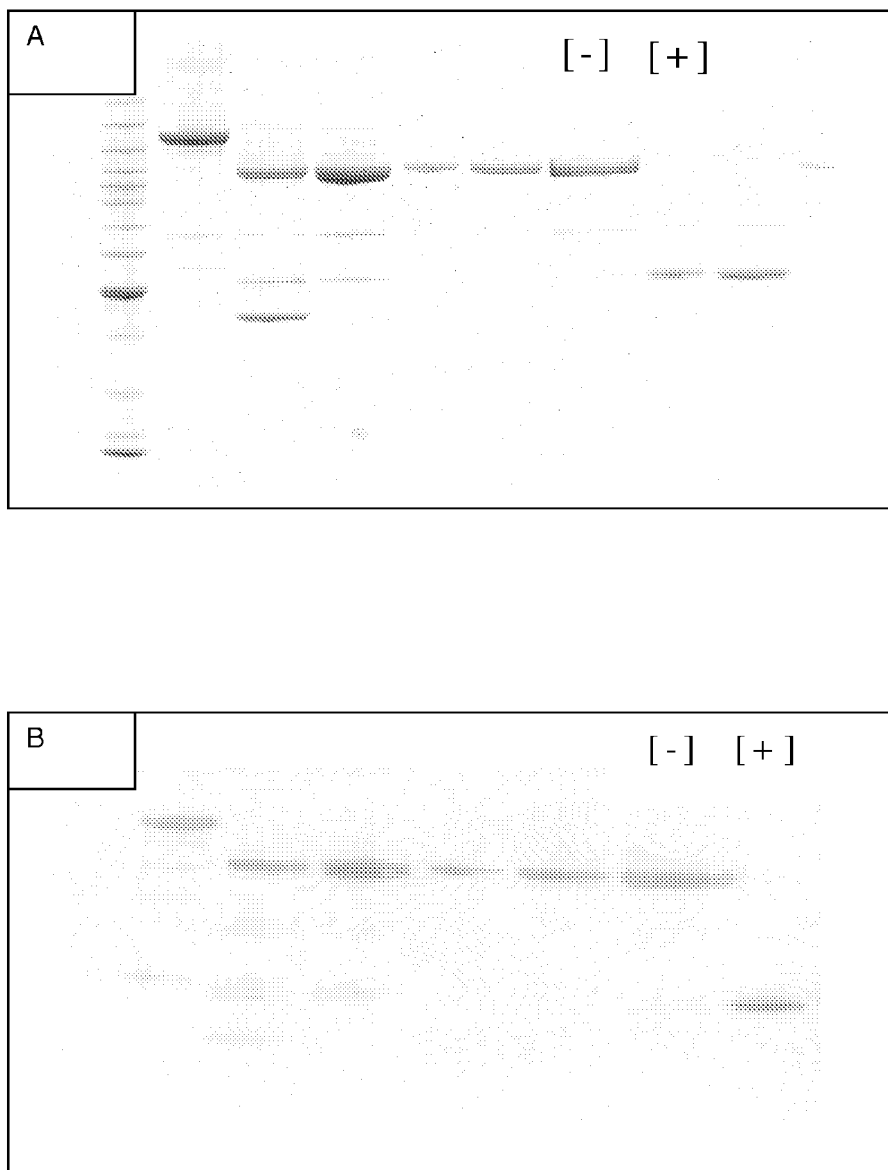


Figure 2

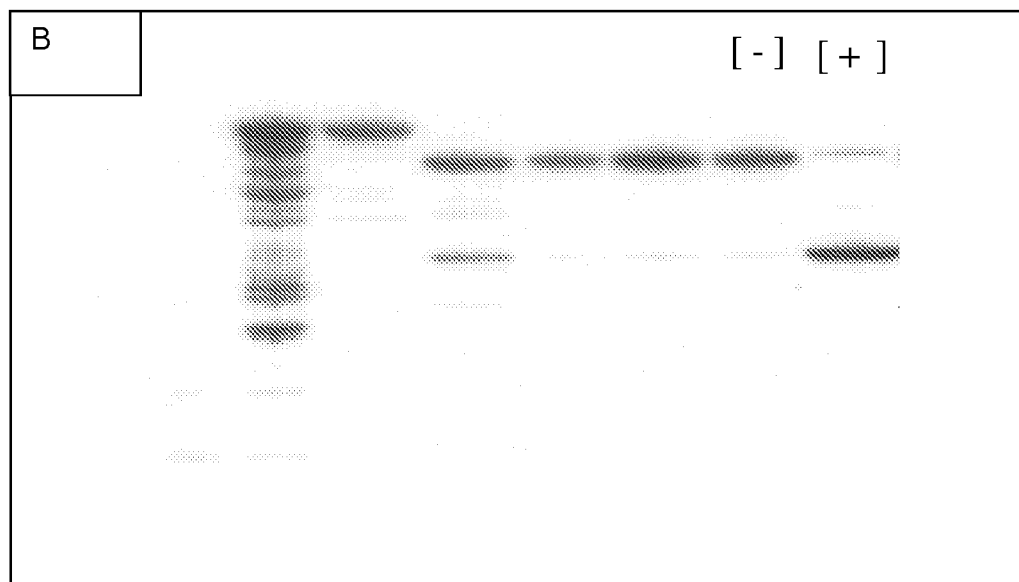
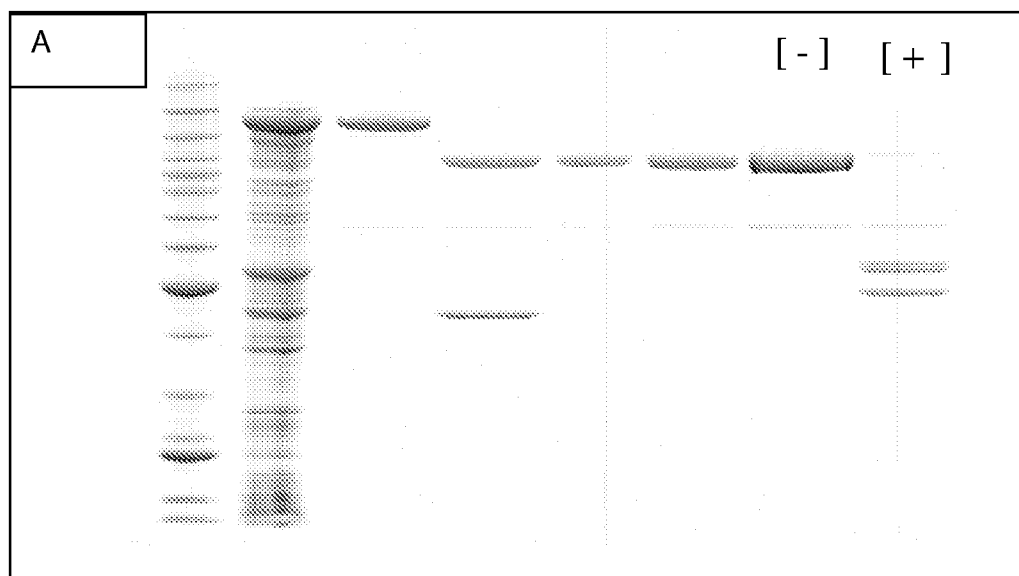


Figure 3

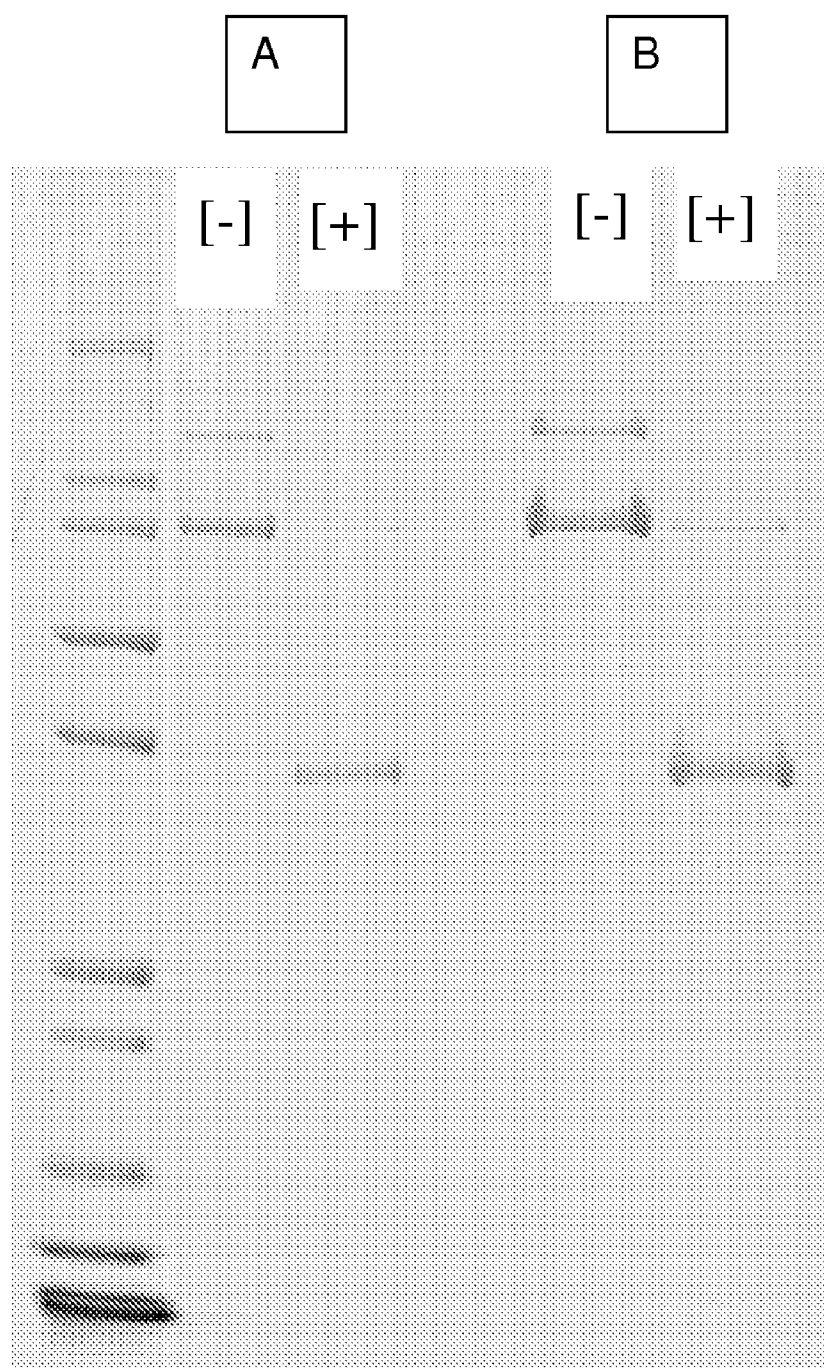


Figure 4

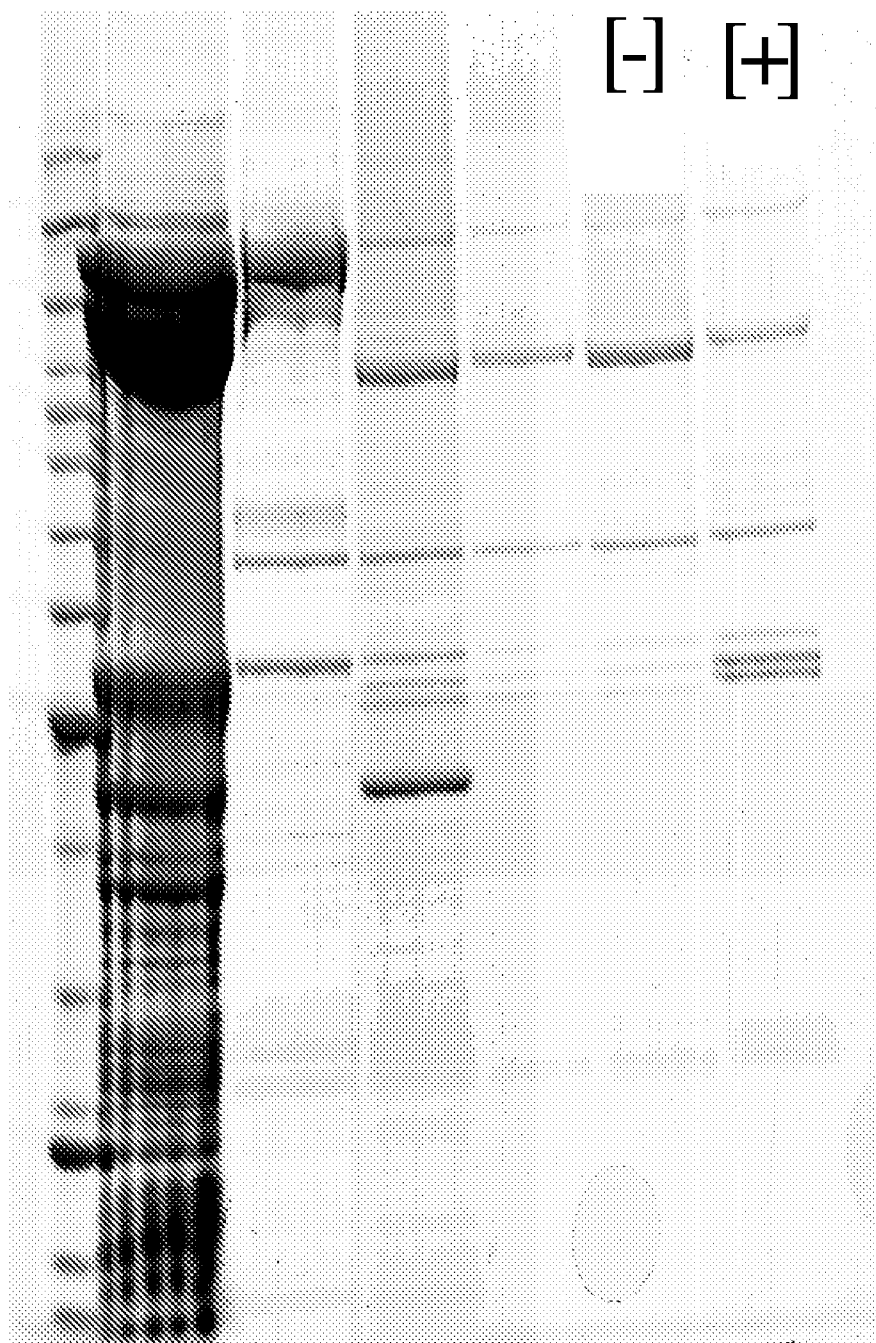


Figure 5

Competition Assay : Nociceptin-LH_N/A Fusions
vs 1 nM [³H]-Nociceptin on eDRGs (4 °C)

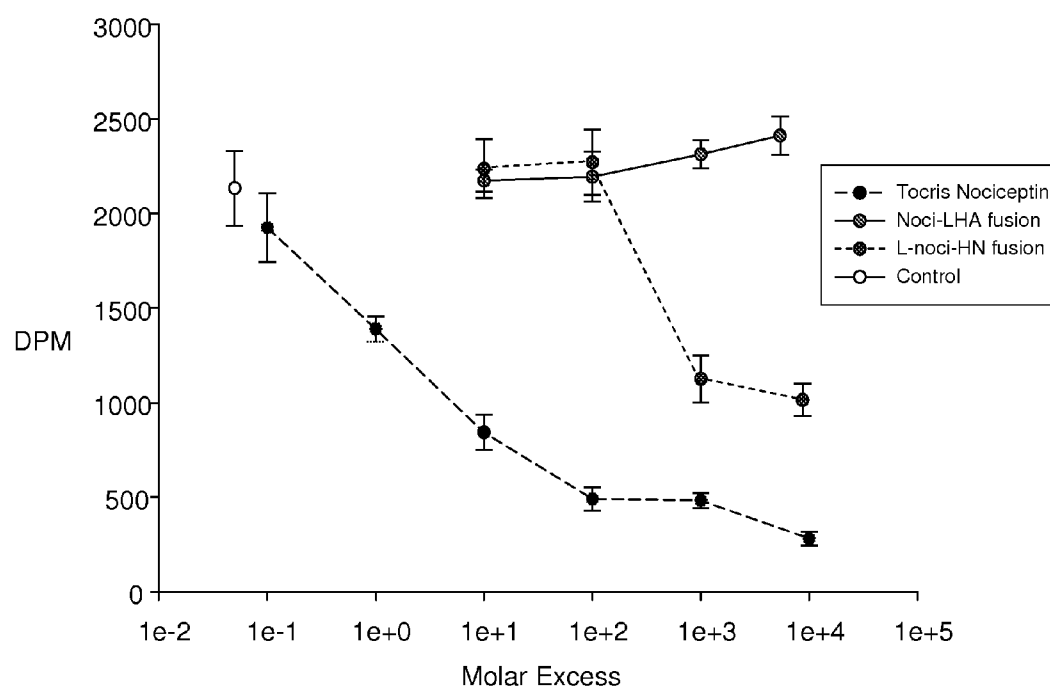


Figure 6

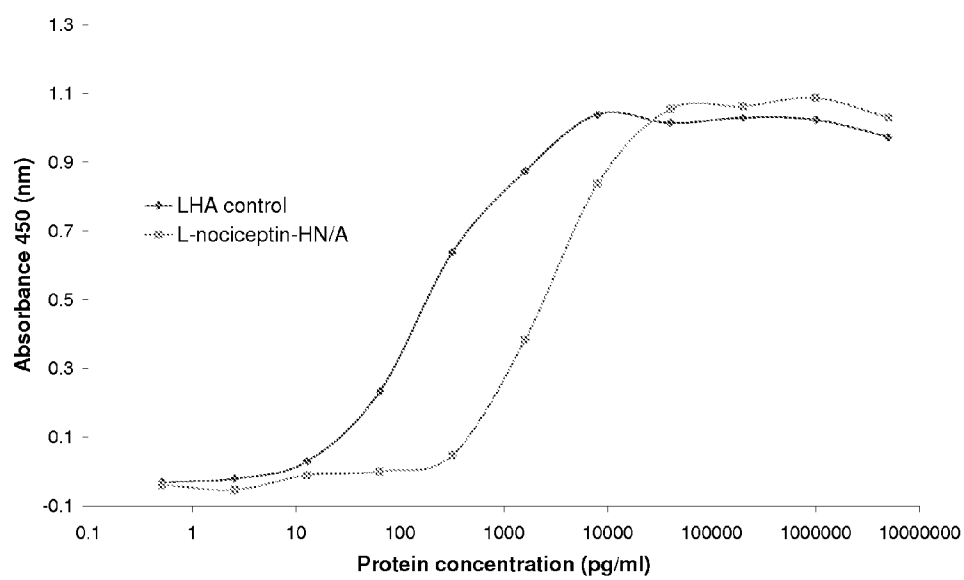


Figure 7

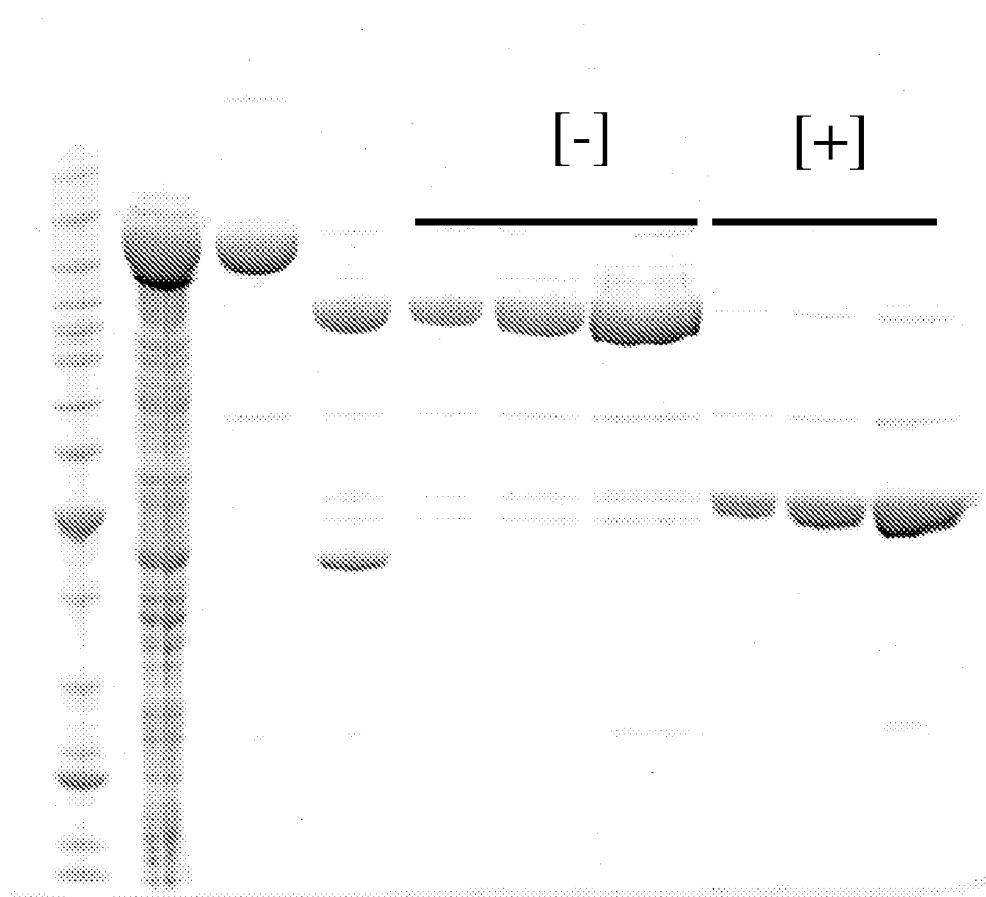


Figure 8

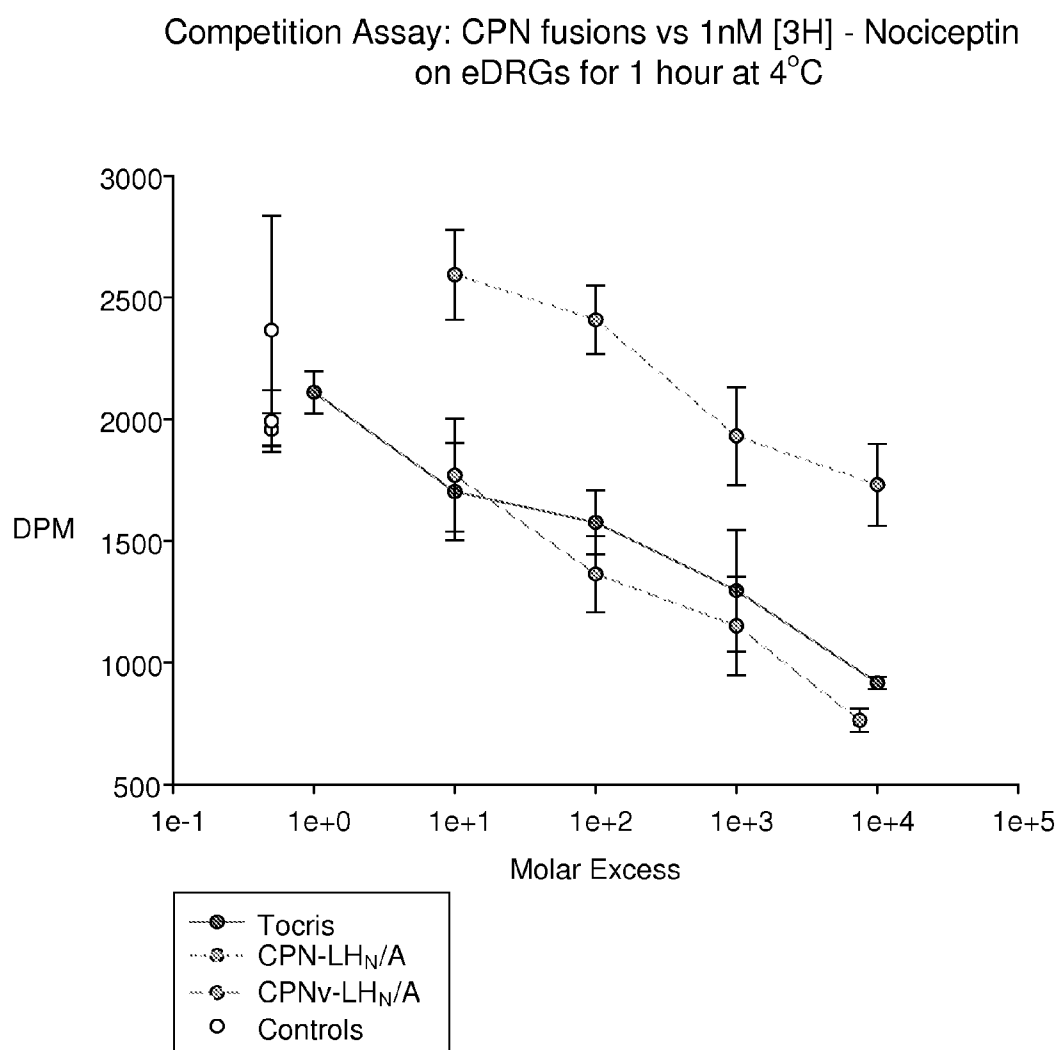


Figure 9

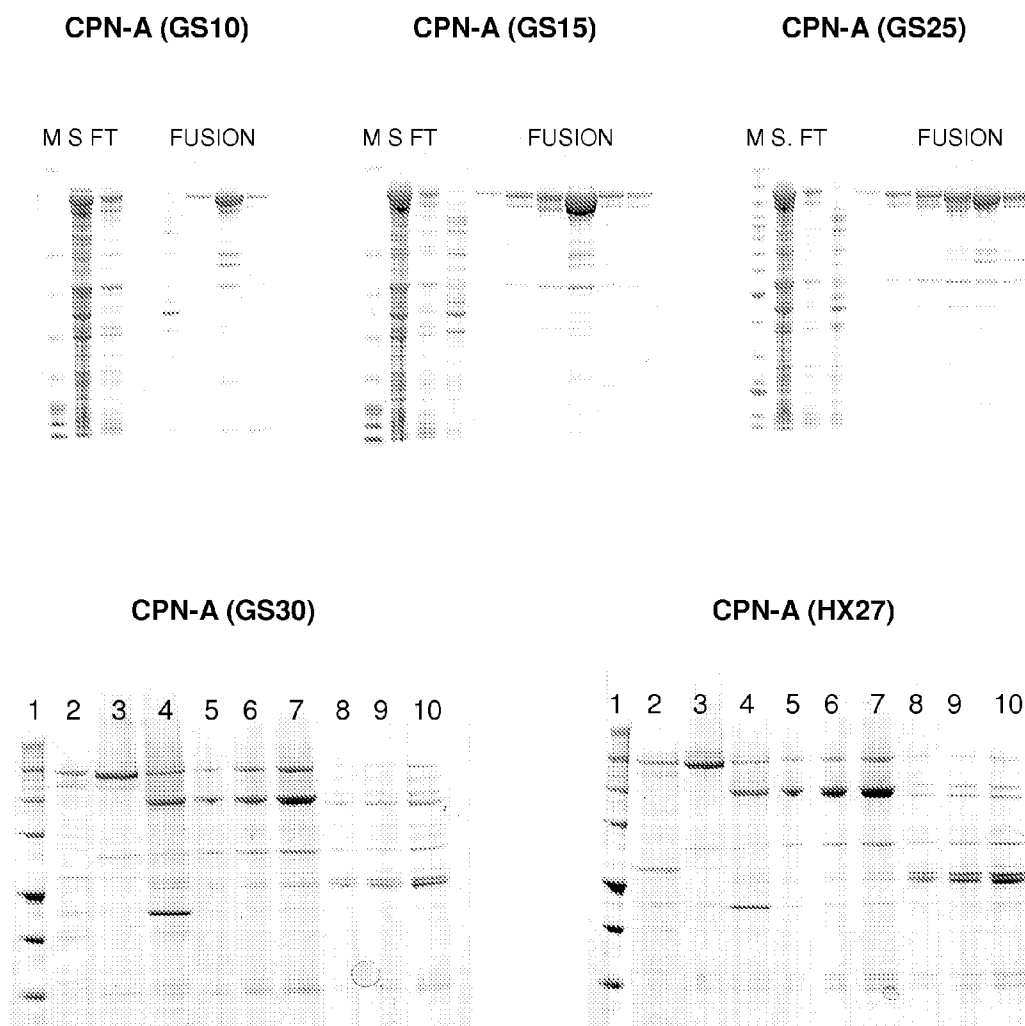


Figure 10

CPN-A on eDRG for 1 Day

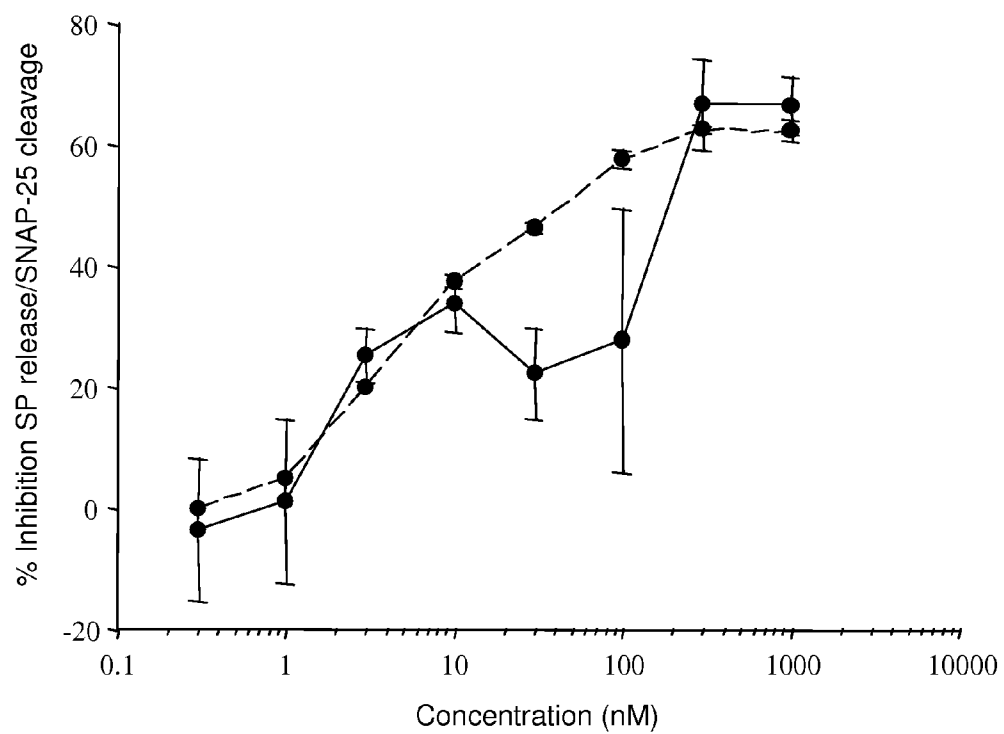


Figure 11

Duration of action following eDRG exposure for 1 Day

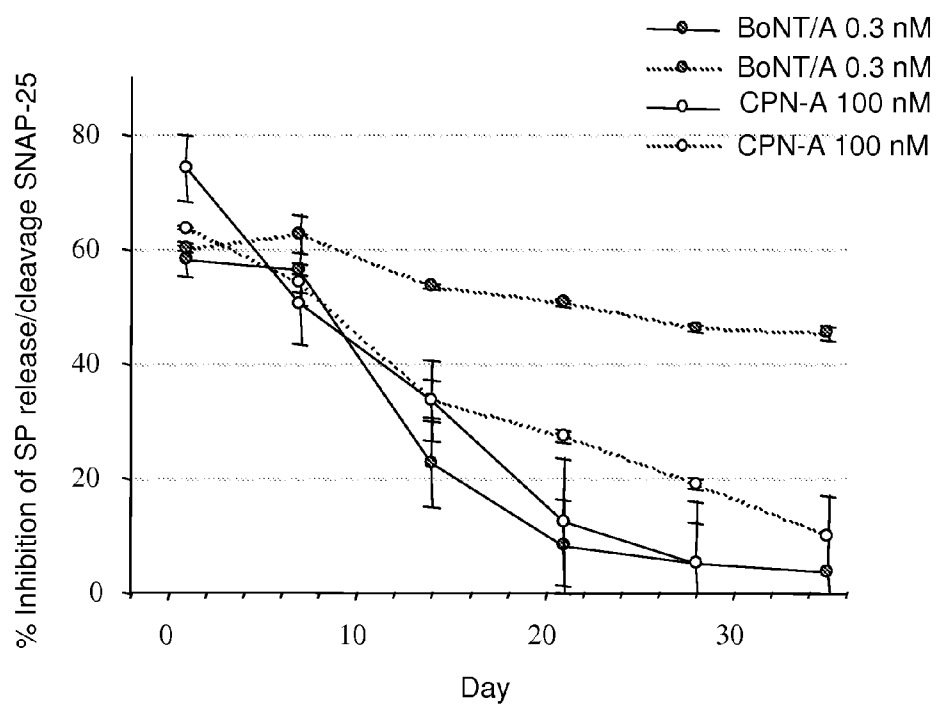


Figure 12

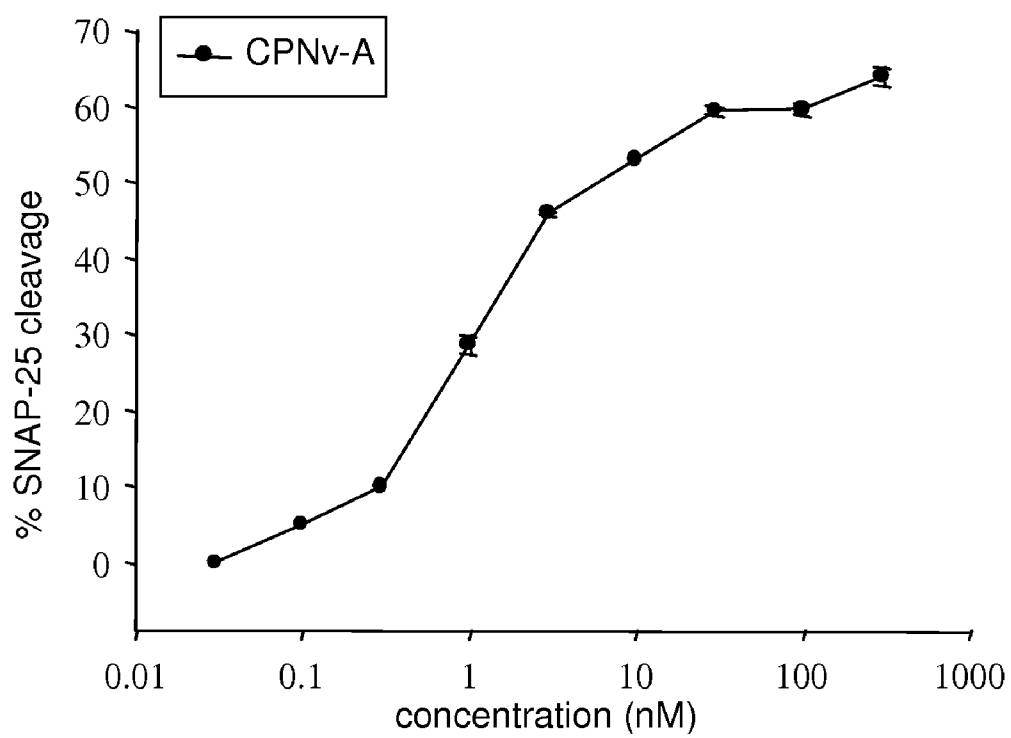


Figure 13

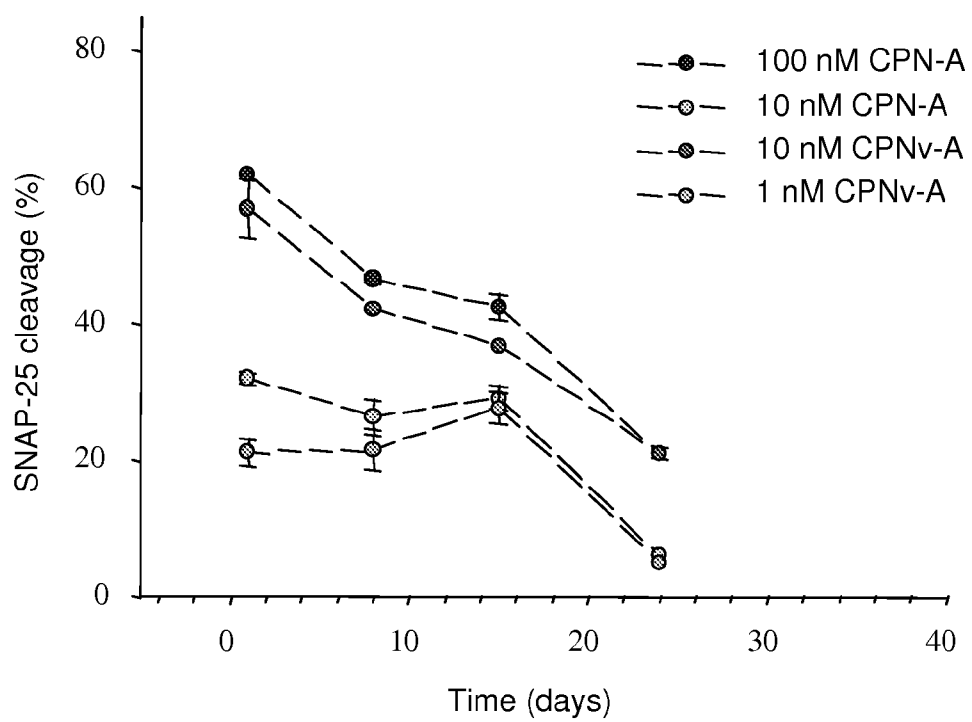


Figure 14

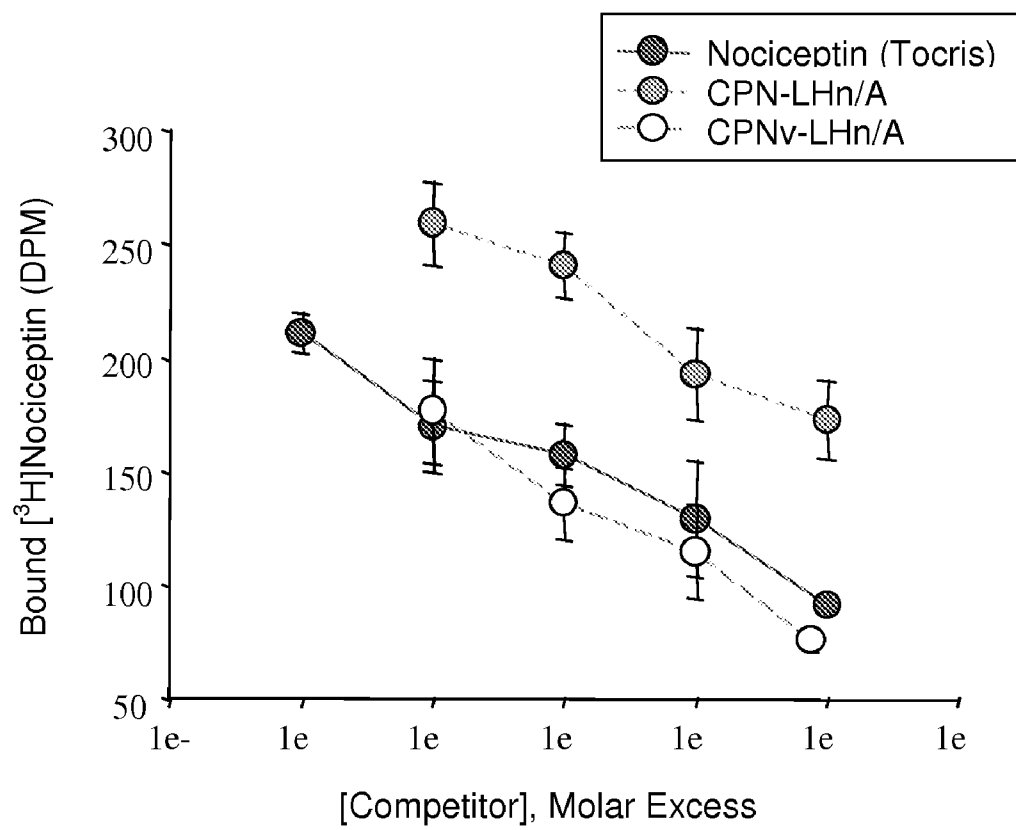


Figure 15

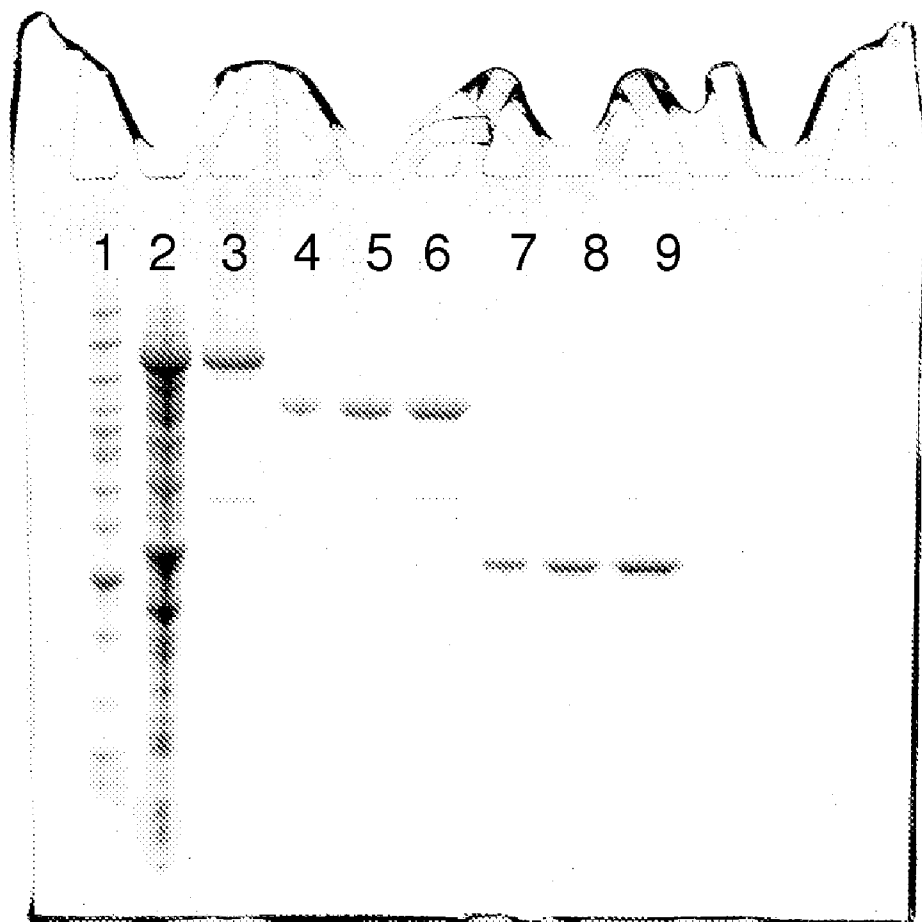


Figure 16

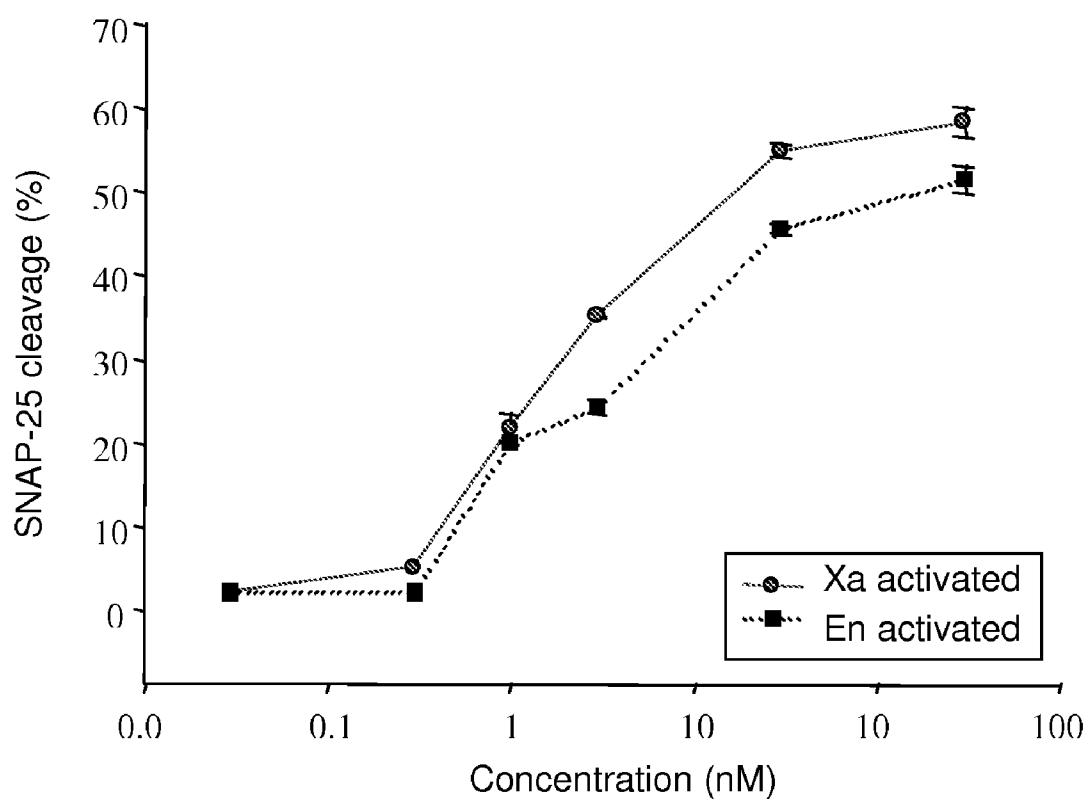


Figure 17

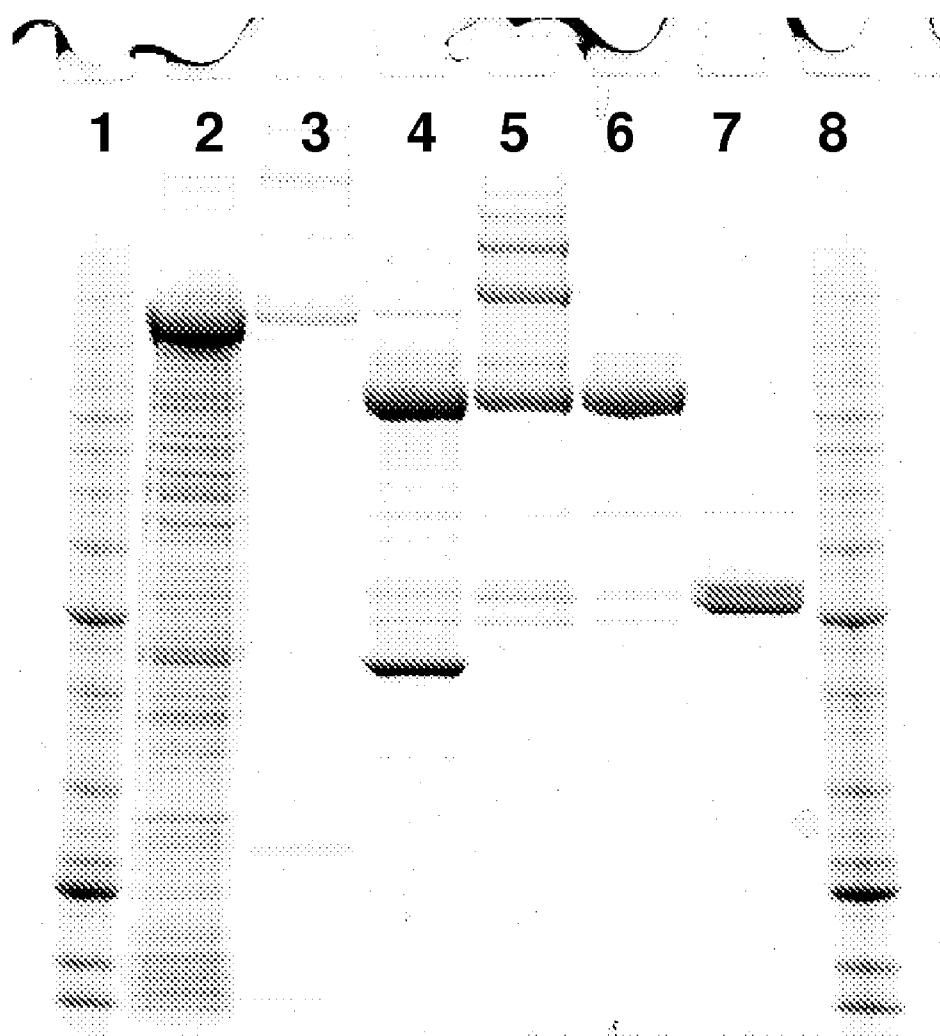


Figure 18

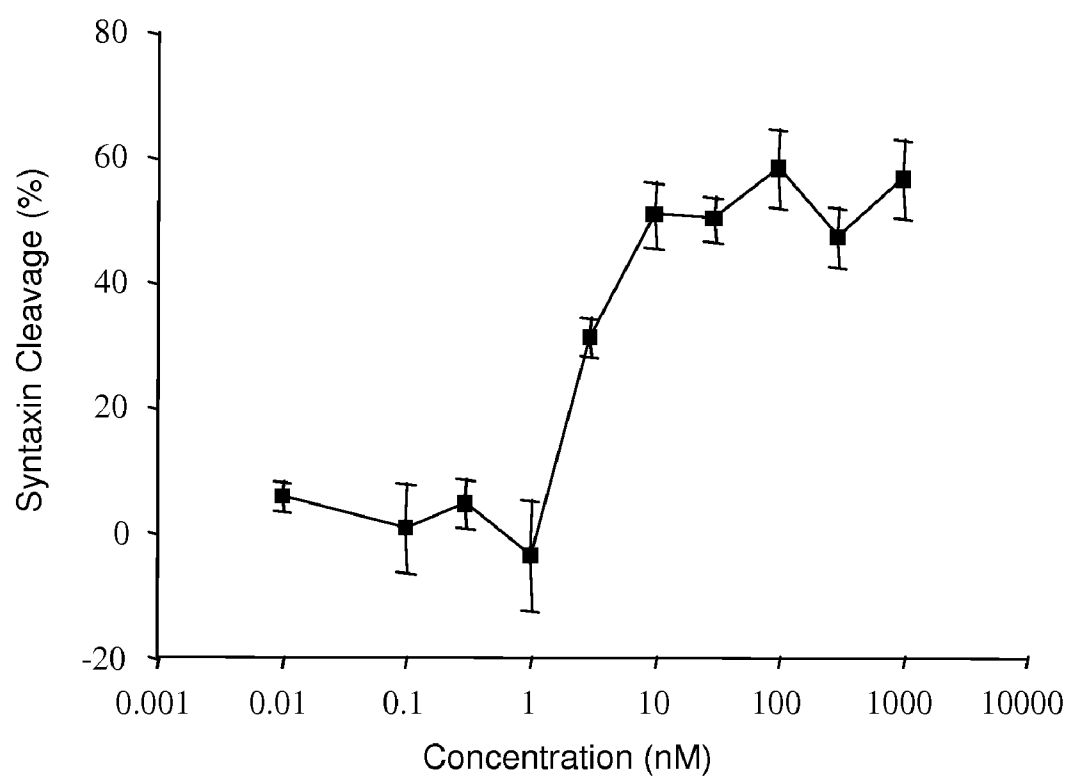


Figure 19

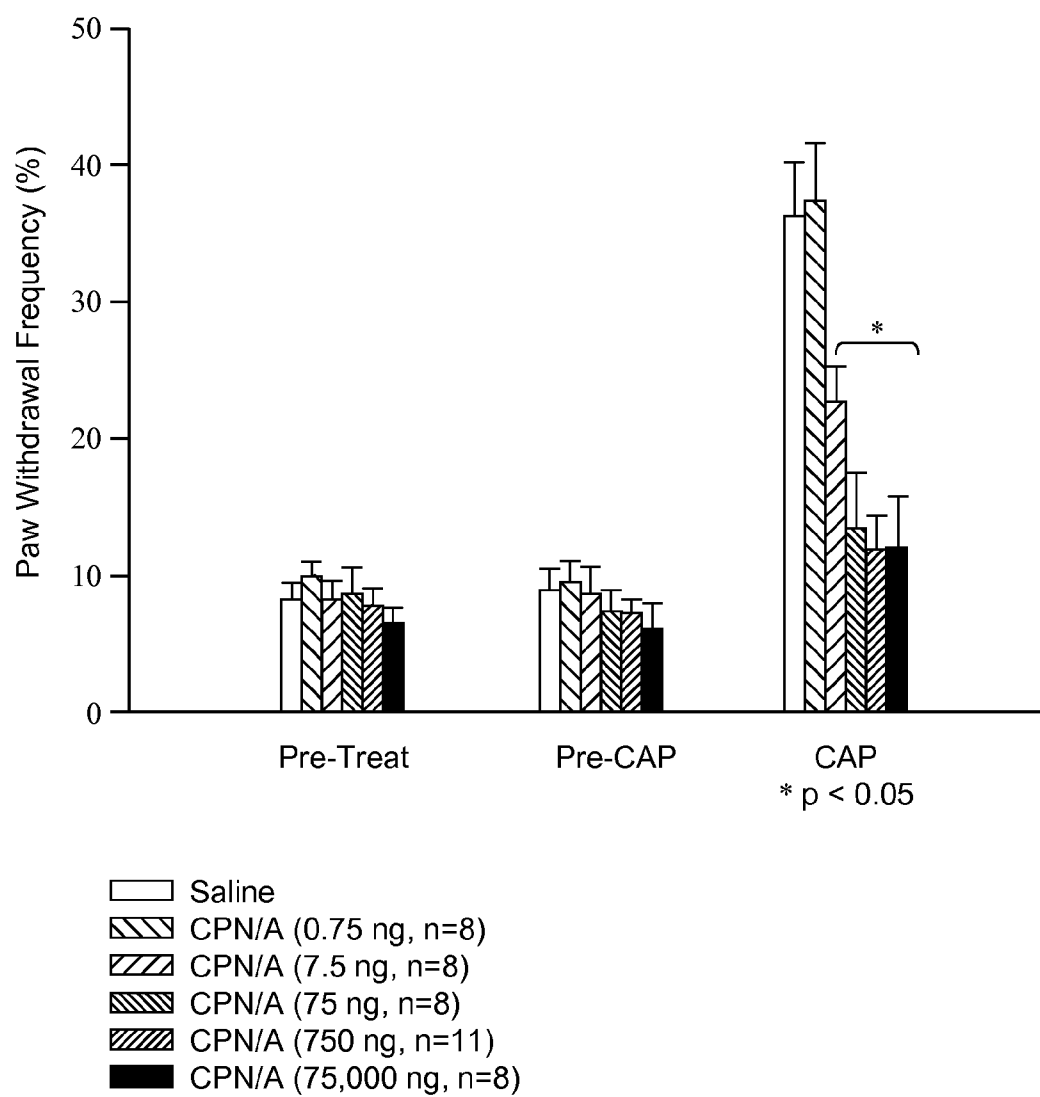


Figure 20

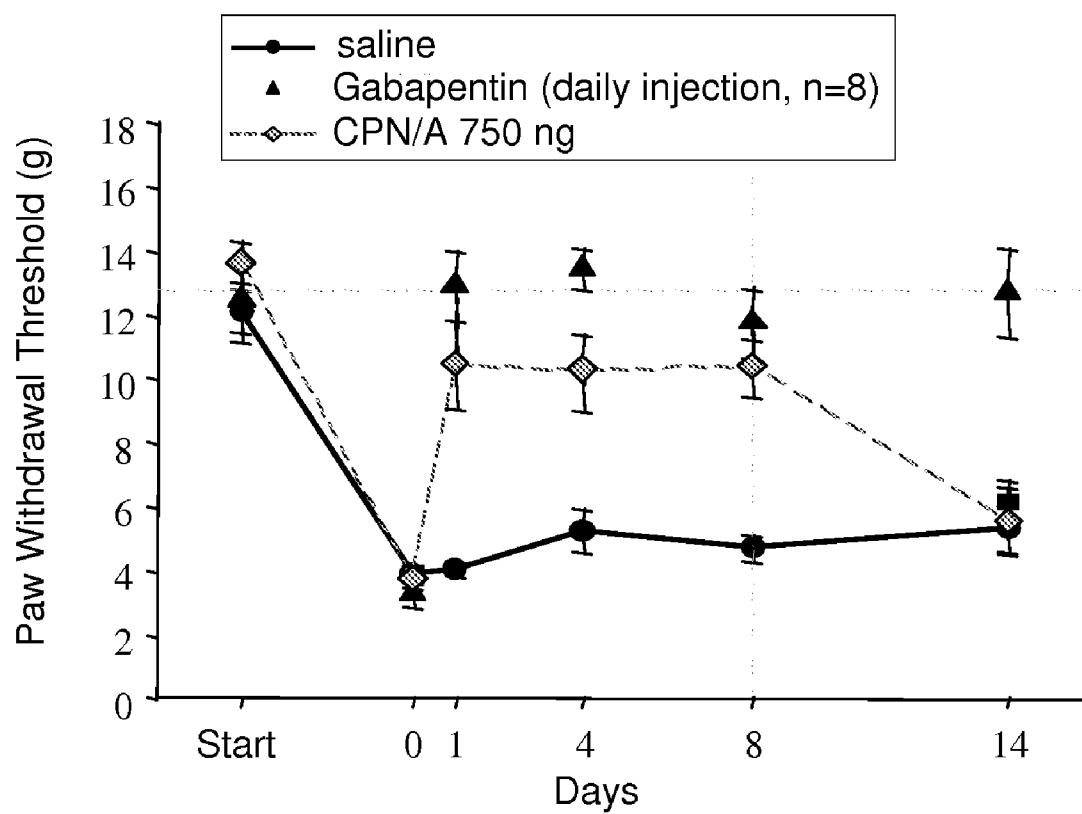


Figure 21

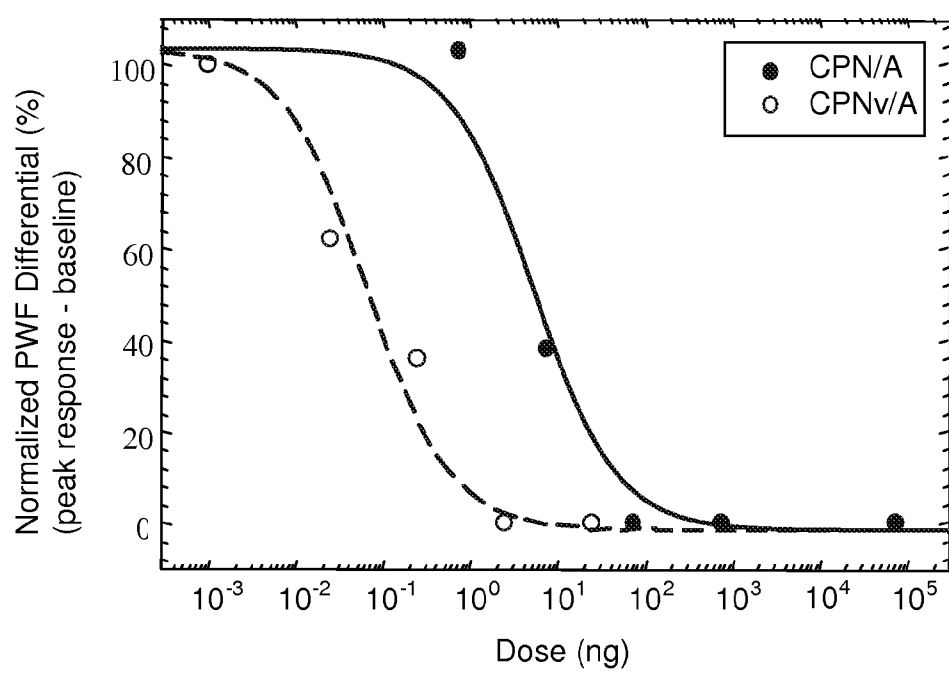


Figure 22

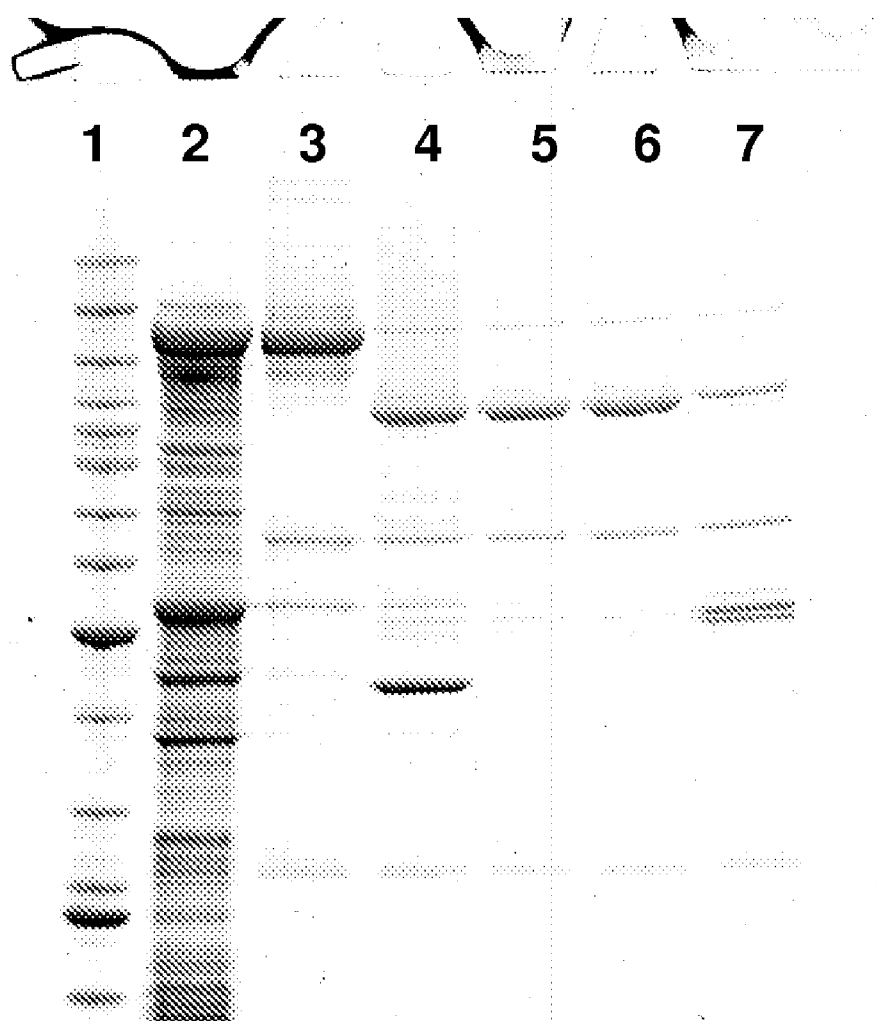


Figure 23

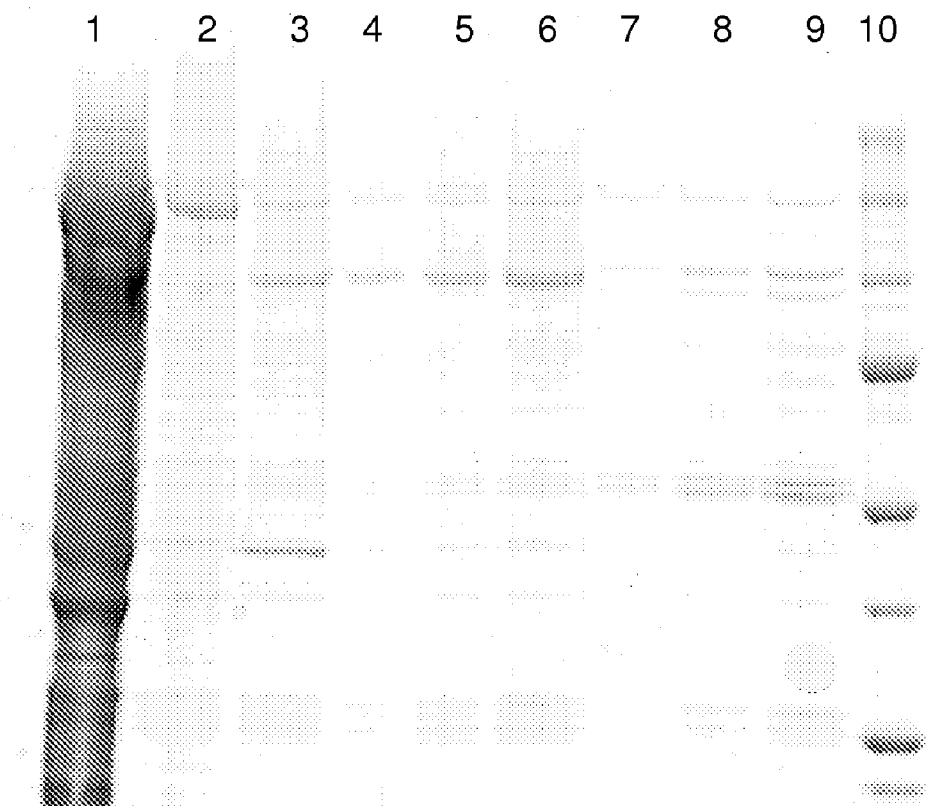


Figure 24

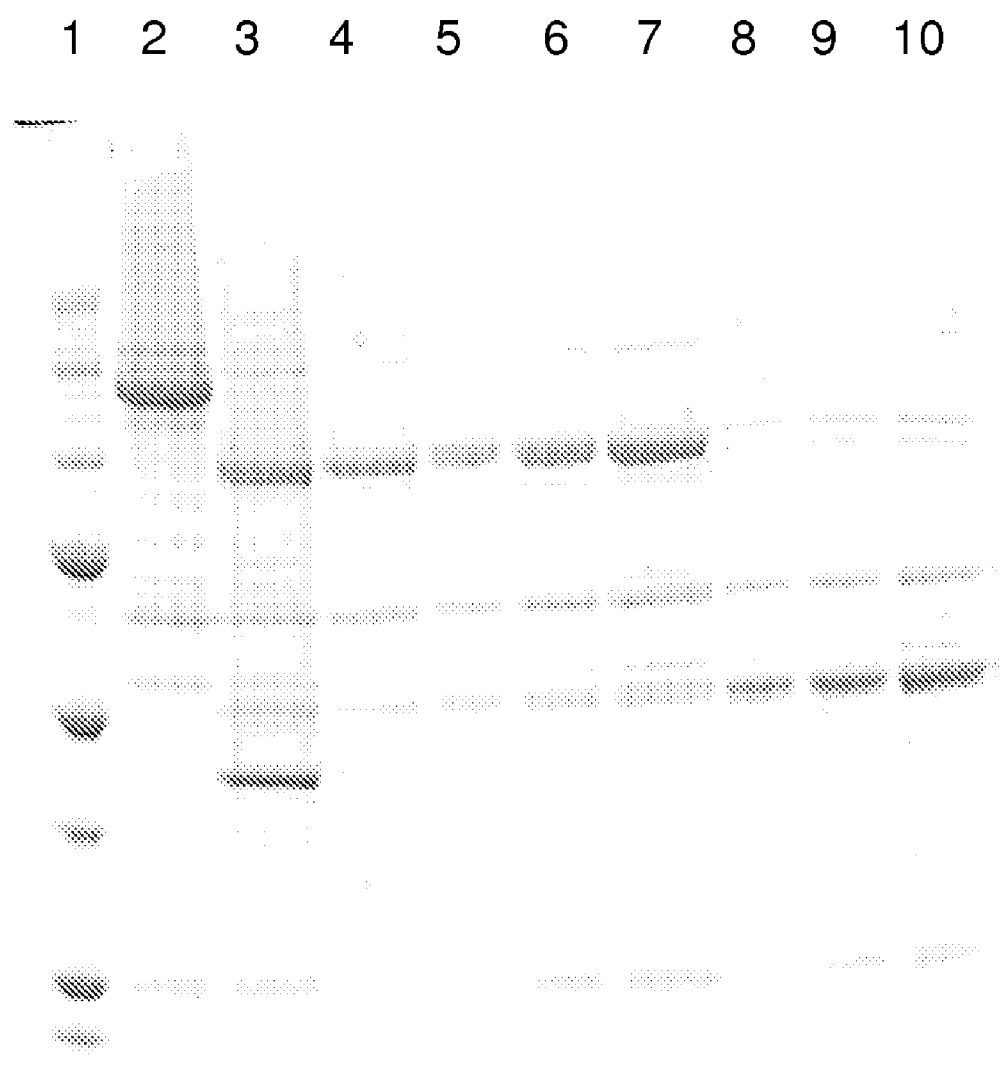


Figure 25

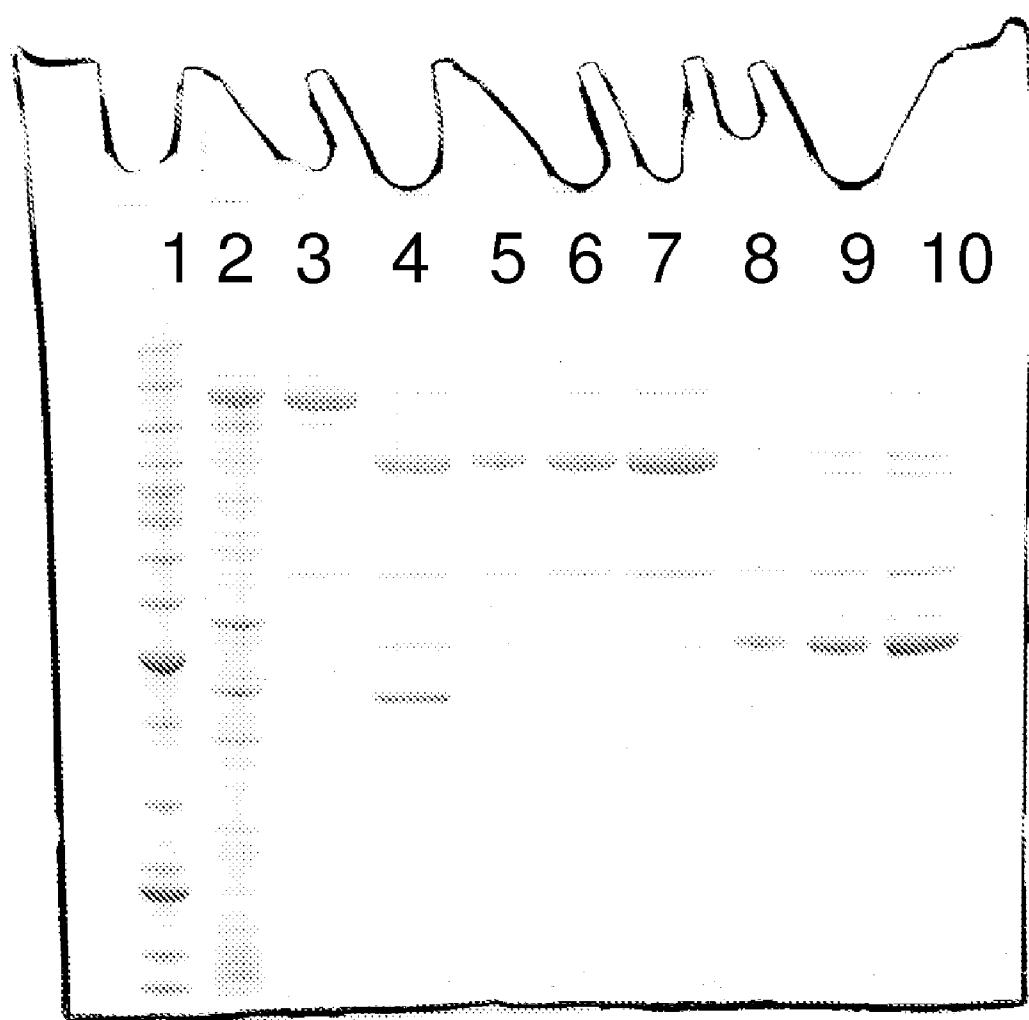


Figure 26

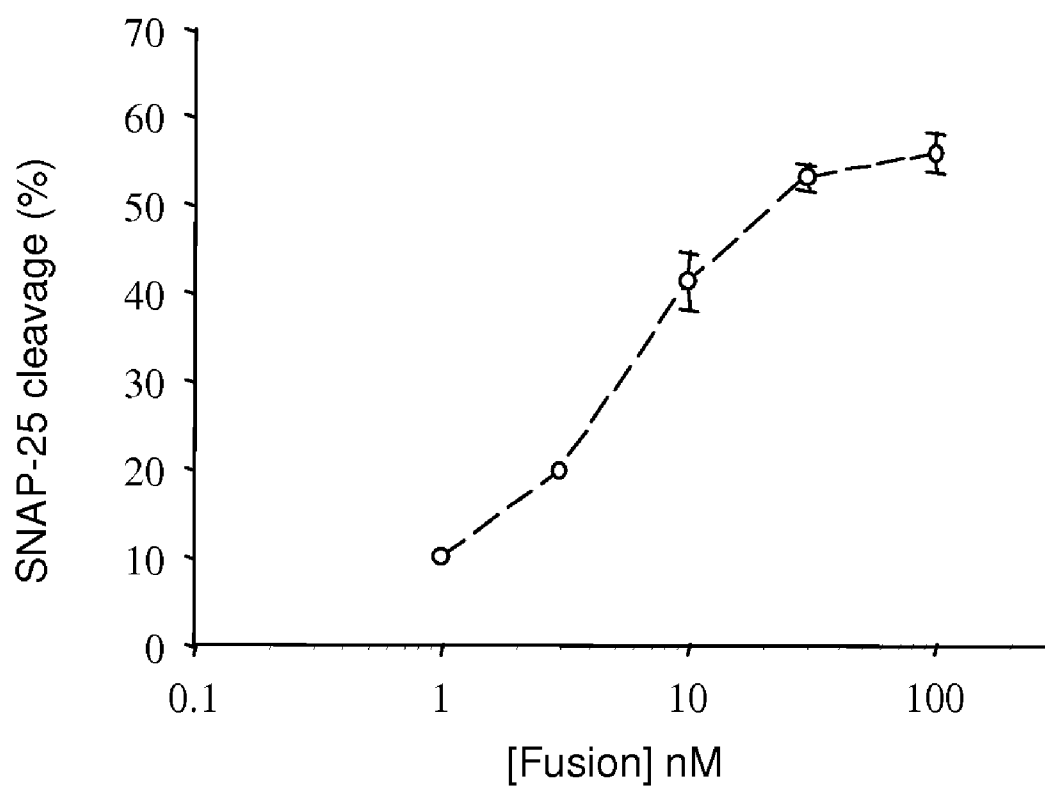


Figure 27

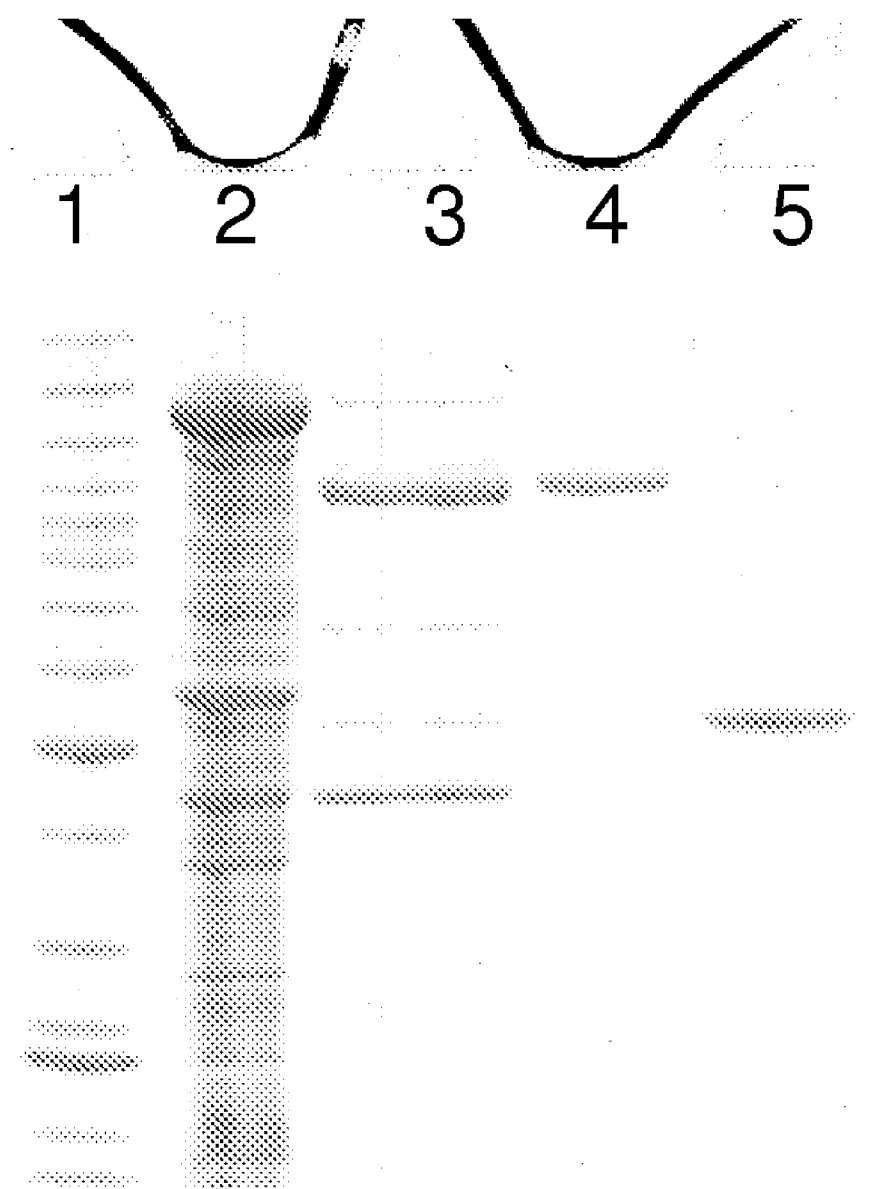


Figure 28

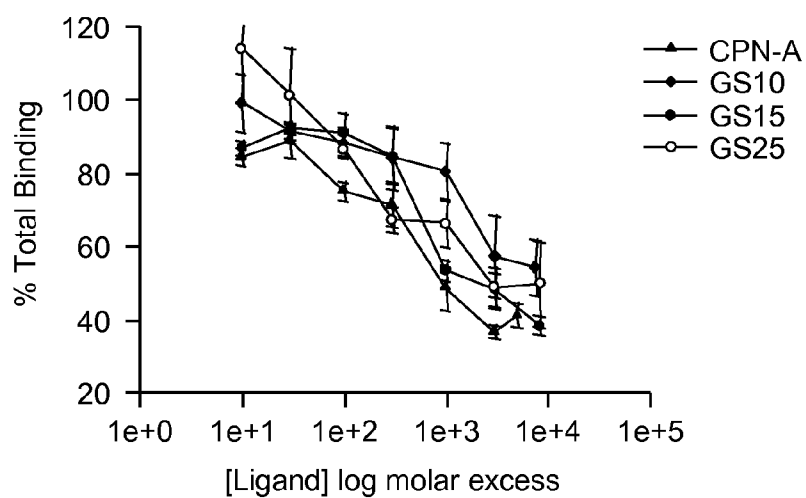
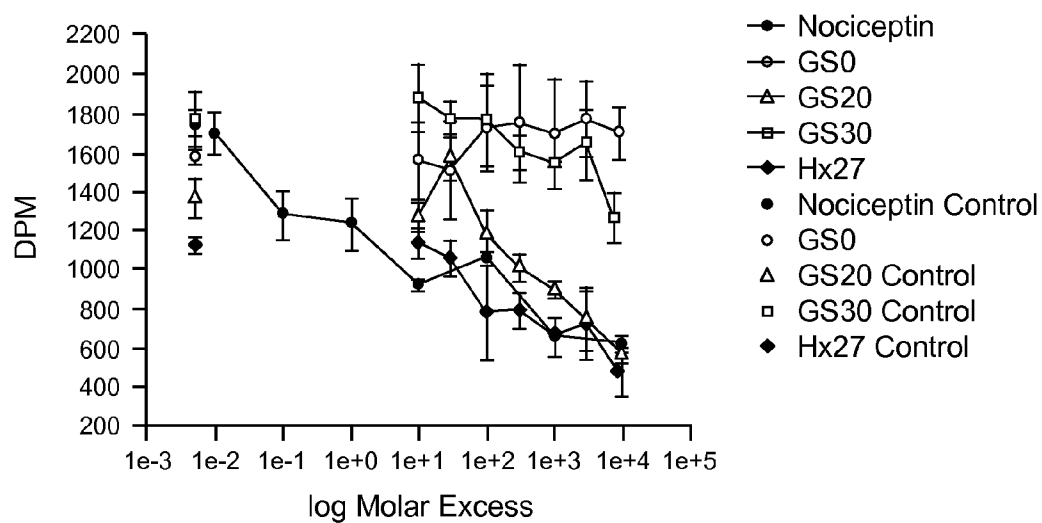


Figure 29

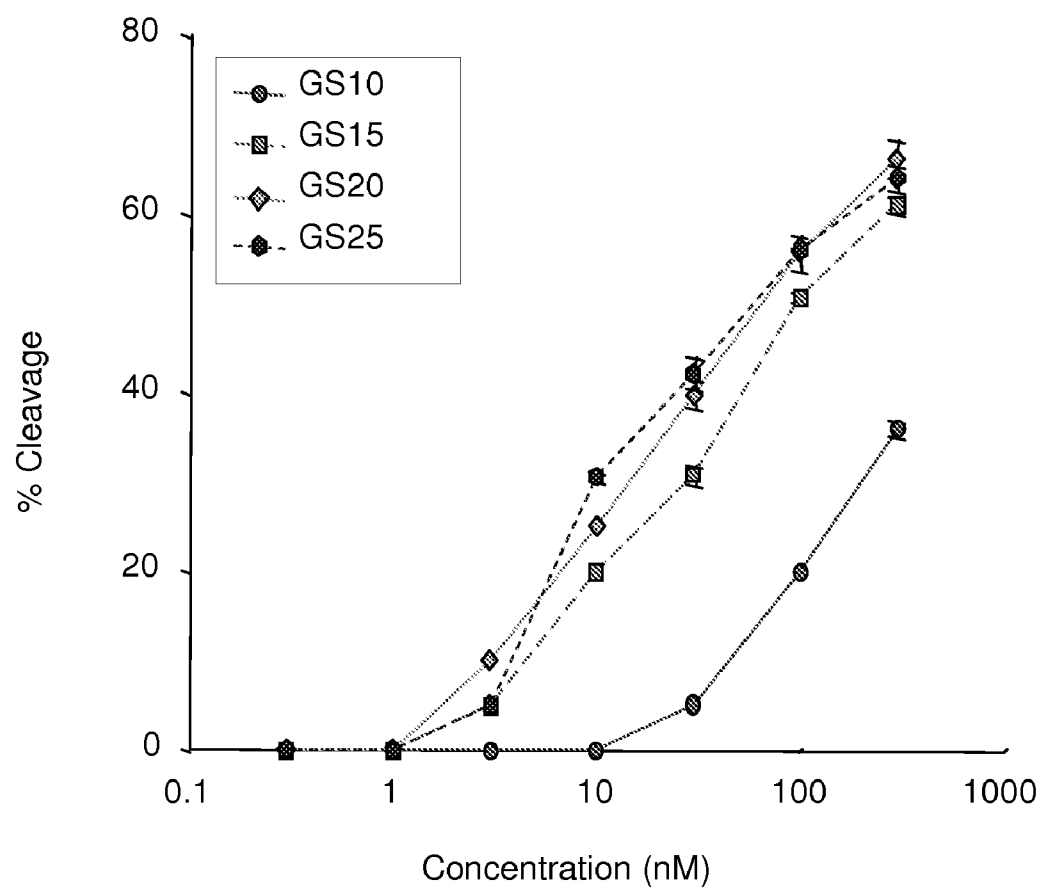


Figure 30

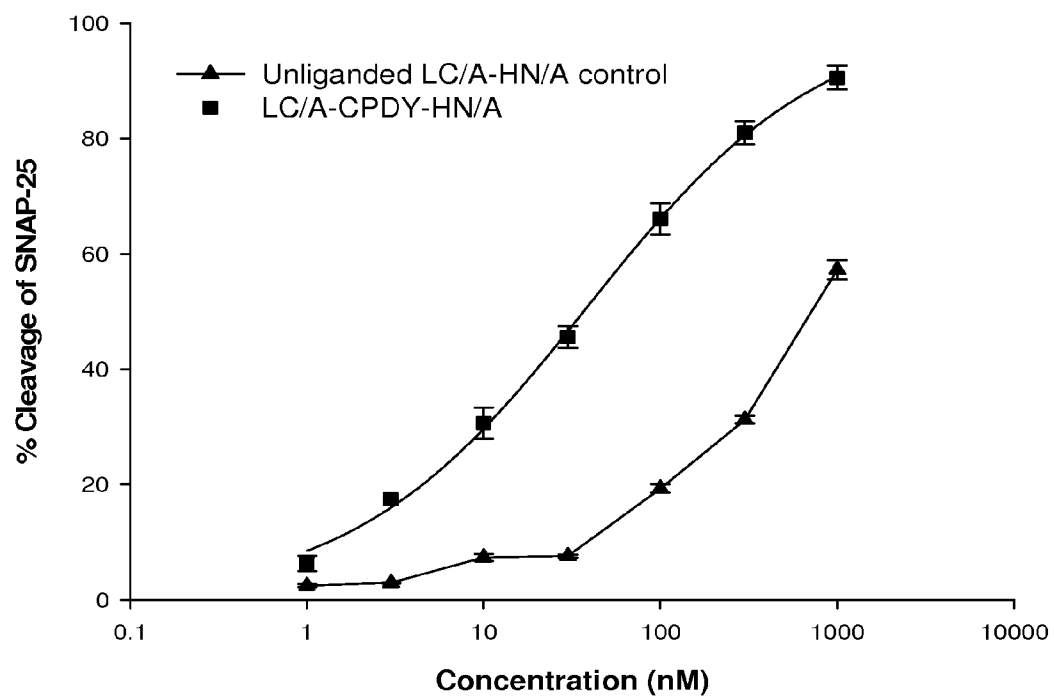


Figure 31

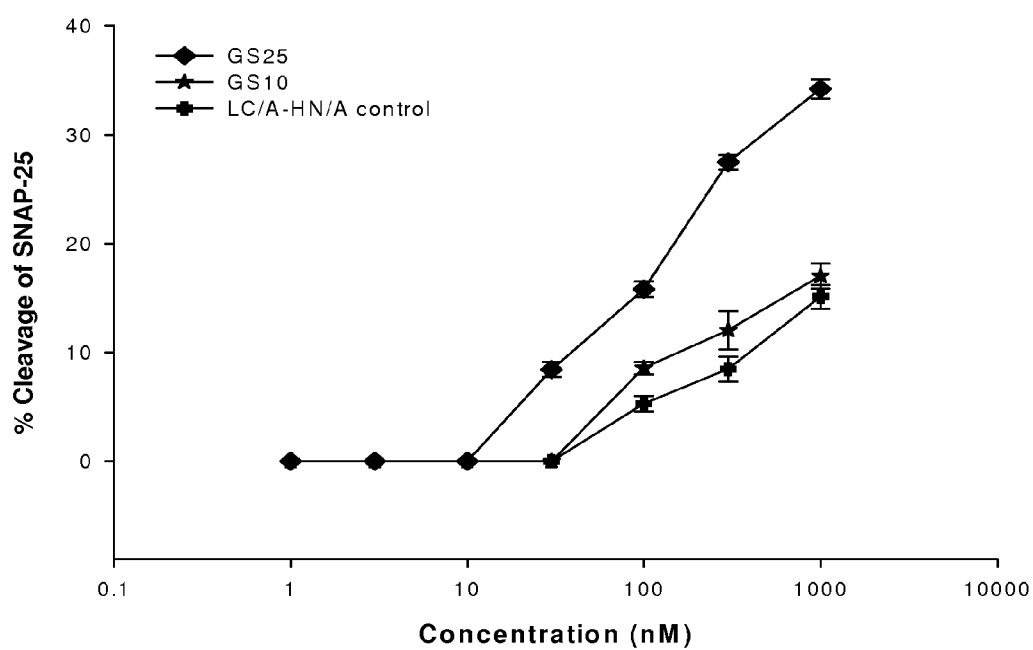


Figure 32

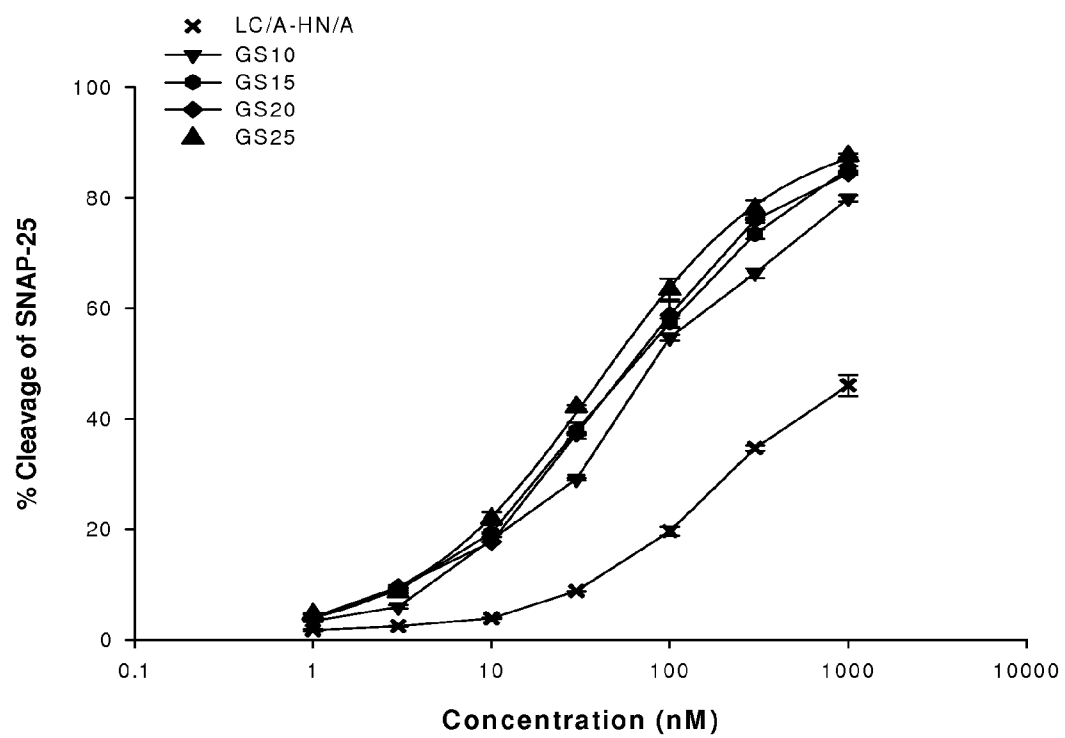


Figure 33

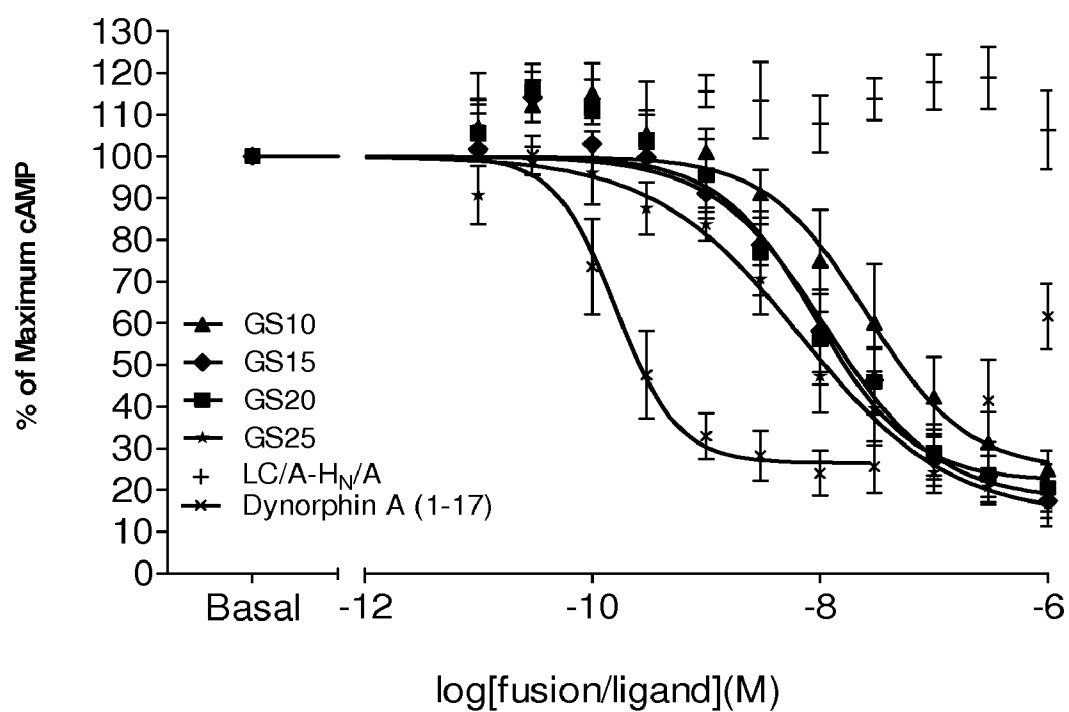
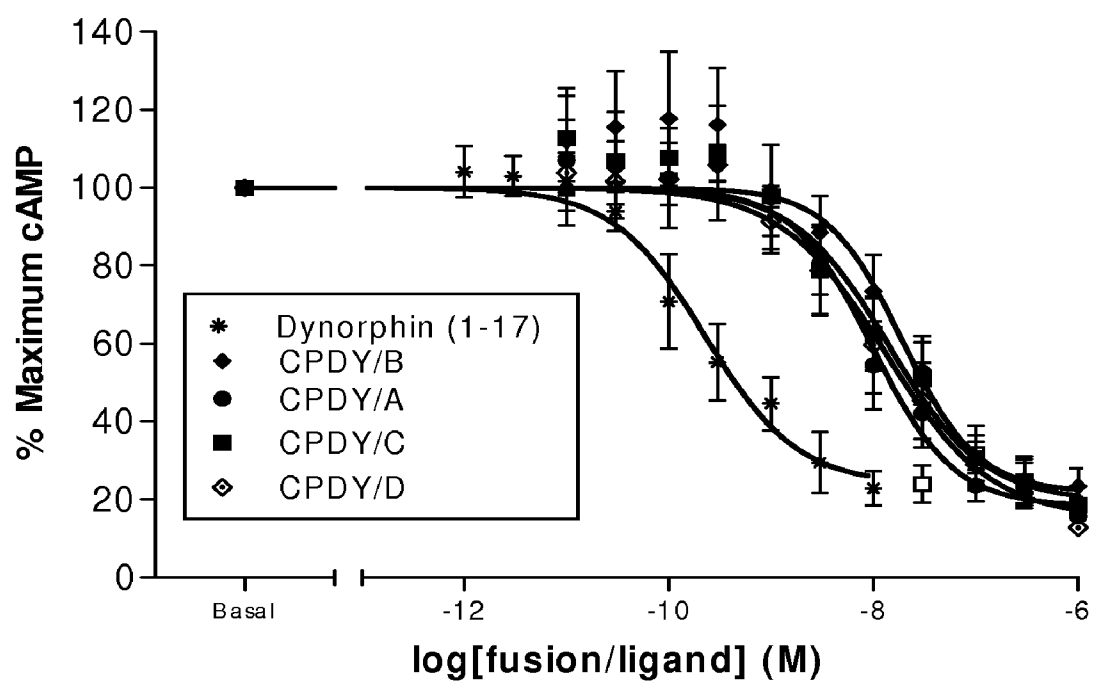


Figure 34



FUSION PROTEINS

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] This application is a continuation-in-part of U.S. application Ser. No. 12/303,078, filed Dec. 1, 2008, which is the National Stage of International Application No. PCT/GB/2007/02049, filed Jun. 1, 2007, which claims priority from Great Britain Application No. 0610867.4, filed Jun. 1, 2006. Each application is expressly incorporated herein by reference in its entirety.

STATEMENT REGARDING SEQUENCE LISTING

[0002] The sequence listing associated with this application is provided in text format in lieu of a paper copy and is hereby incorporated by reference into the specification. The name of the text file containing the sequence listing is 35492_SEQ_FINAL.txt. The text file is 372 KB; was created on Aug. 23, 2010; and is being submitted via EFS-Web with the filing of the specification.

FIELD OF THE INVENTION

[0003] This invention relates to non-cytotoxic fusion proteins, and to the therapeutic application thereof as analgesic molecules.

BACKGROUND OF THE INVENTION

[0004] Toxins may be generally divided into two groups according to the type of effect that they have on a target cell. In more detail, the first group of toxins kill their natural target cells, and are therefore known as cytotoxic toxin molecules. This group of toxins is exemplified inter alia by plant toxins such as ricin, and abrin, and by bacterial toxins such as diphtheria toxin, and *Pseudomonas* exotoxin A. Cytotoxic toxins have attracted much interest in the design of “magic bullets” (e.g. immunoconjugates, which comprise a cytotoxic toxin component and an antibody that binds to a specific marker on a target cell) for the treatment of cellular disorders and conditions such as cancer. Cytotoxic toxins typically kill their target cells by inhibiting the cellular process of protein synthesis.

[0005] The second group of toxins, which are known as non-cytotoxic toxins, do not (as their name confirms) kill their natural target cells. Non-cytotoxic toxins have attracted much less commercial interest than have their cytotoxic counterparts, and exert their effects on a target cell by inhibiting cellular processes other than protein synthesis. Non-cytotoxic toxins are produced by a variety of plants, and by a variety of microorganisms such as *Clostridium* sp. and *Neisseria* sp.

[0006] Clostridial neurotoxins are proteins that typically have a molecular mass of the order of 150 kDa. They are produced by various species of bacteria, especially of the genus *Clostridium*, most importantly *C. tetani* and several strains of *C. botulinum*, *C. butyricum* and *C. argentinense*. There are at present eight different classes of the clostridial neurotoxin, namely: tetanus toxin, and botulinum neurotoxin in its serotypes A, B, C1, D, E, F and G, and they all share similar structures and modes of action.

[0007] Clostridial neurotoxins represent a major group of non-cytotoxic toxin molecules, and are synthesised by the host bacterium as single polypeptides that are modified post-

translationally by a proteolytic cleavage event to form two polypeptide chains joined together by a disulphide bond. The two chains are termed the heavy chain (H-chain), which has a molecular mass of approximately 100 kDa, and the light chain (L-chain), which has a molecular mass of approximately 50 kDa.

[0008] L-chains possess a protease function (zinc-dependent endopeptidase activity) and exhibit a high substrate specificity for vesicle and/or plasma membrane associated proteins involved in the exocytic process. L-chains from different clostridial species or serotypes may hydrolyse different but specific peptide bonds in one of three substrate proteins, namely synaptobrevin, syntaxin or SNAP-25. These substrates are important components of the neurosecretory machinery. *Neisseria* sp., most importantly from the species *N. gonorrhoeae*, produce functionally similar non-cytotoxic proteases. An example of such a protease is IgA protease (see WO99/58571).

[0009] It has been well documented in the art that toxin molecules may be re-targeted to a cell that is not the toxin's natural target cell. When so re-targeted, the modified toxin is capable of binding to a desired target cell and, following subsequent translocation into the cytosol, is capable of exerting its effect on the target cell. Said re-targeting is achieved by replacing the natural Targeting Moiety (TM) of the toxin with a different TM. In this regard, the TM is selected so that it will bind to a desired target cell, and allow subsequent passage of the modified toxin into an endosome within the target cell. The modified toxin also comprises a translocation domain to enable entry of the non-cytotoxic protease into the cell cytosol. The translocation domain can be the natural translocation domain of the toxin or it can be a different translocation domain obtained from a microbial protein with translocation activity.

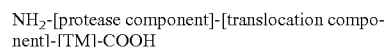
[0010] The above-mentioned TM replacement may be effected by conventional chemical conjugation techniques, which are well known to a skilled person. In this regard, reference is made to Hermanson, G. T. (1996), Bioconjugate techniques, Academic Press, and to Wong, S. S. (1991), Chemistry of protein conjugation and cross-linking, CRC Press.

[0011] Chemical conjugation is, however, often imprecise. For example, following conjugation, a TM may become joined to the remainder of the conjugate at more than one attachment site.

[0012] Chemical conjugation is also difficult to control. For example, a TM may become joined to the remainder of the modified toxin at an attachment site on the protease component and/or on the translocation component. This is problematic when attachment to only one of said components (preferably at a single site) is desired for therapeutic efficacy.

[0013] Thus, chemical conjugation results in a mixed population of modified toxin molecules, which is undesirable.

[0014] As an alternative to chemical conjugation, TM replacement may be effected by recombinant preparation of a single polypeptide fusion protein (see WO98/07864). This technique is based on the in vivo bacterial mechanism by which native clostridial neurotoxin (i.e. holotoxin) is prepared, and results in a fusion protein having the following structural arrangement:



[0015] According to WO98/07864, the TM is placed towards the C-terminal end of the fusion protein. The fusion

protein is then activated by treatment with a protease, which cleaves at a site between the protease component and the translocation component. A di-chain protein is thus produced, comprising the protease component as a single polypeptide chain covalently attached (via a disulphide bridge) to another single polypeptide chain containing the translocation component plus TM. Whilst the WO98/07864 methodology follows (in terms of structural arrangement of the fusion protein) the natural expression system of clostridial holotoxin, the present inventors have found that this system may result in the production of certain fusion proteins that have a substantially-reduced binding ability for the intended target cell.

[0016] There is therefore a need for an alternative or improved system for constructing a non-cytotoxic fusion protein.

SUMMARY OF THE INVENTION

[0017] The present invention addresses one or more of the above-mentioned problems by providing a single chain, polypeptide fusion protein, comprising:

- [0018] a. a non-cytotoxic protease, or a fragment thereof, which protease or protease fragment is capable of cleaving a protein of the exocytic fusion apparatus in a nociceptive sensory afferent;
- [0019] b. a Targeting Moiety that is capable of binding to a Binding Site on the nociceptive sensory afferent, which Binding Site is capable of undergoing endocytosis to be incorporated into an endosome within the nociceptive sensory afferent;
- [0020] c. a protease cleavage site at which site the fusion protein is cleavable by a protease, wherein the protease cleavage site is located between the non-cytotoxic protease or fragment thereof and the Targeting Moiety; and
- [0021] d. a translocation domain that is capable of translocating the protease or protease fragment from within an endosome, across the endosomal membrane and into the cytosol of the nociceptive sensory afferent.

BRIEF DESCRIPTION OF THE DRAWINGS

[0022] FIG. 1—Purification of a LC/A-nociceptin-H_N/A fusion protein

[0023] Using the methodology outlined in Example 9, a LC/A-nociceptin-H_N/A fusion protein was purified from *E. coli* BL21 cells. Briefly, the soluble products obtained following cell disruption were applied to a nickel-charged affinity capture column. Bound proteins were eluted with 100 mM imidazole, treated with Factor Xa to activate the fusion protein and remove the maltose-binding protein (MBP) tag, then re-applied to a second nickel-charged affinity capture column. Samples from the purification procedure were assessed by SDS-PAGE (Panel A) and Western blotting (Panel B). Anti-nociceptin antisera (obtained from Abcam) were used as the primary antibody for Western blotting. The final purified material in the absence and presence of reducing agent is identified in the lanes marked [−] and [+] respectively.

[0024] FIG. 2—Purification of a nociceptin-LC/A-H_N/A fusion protein

[0025] Using the methodology outlined in Example 9, a nociceptin-LC/A-H_N/A fusion protein was purified from *E. coli* BL21 cells. Briefly, the soluble products obtained following cell disruption were applied to a nickel-charged affinity capture column. Bound proteins were eluted with 100 mM imidazole, treated with Factor Xa to activate the fusion protein

and remove the maltose-binding protein (MBP) tag, then re-applied to a second nickel-charged affinity capture column. Samples from the purification procedure were assessed by SDS-PAGE (Panel A) and Western blotting (Panel B). Anti-nociceptin antisera (obtained from Abcam) were used as the primary antibody for Western blotting. The final purified material in the absence and presence of reducing agent is identified in the lanes marked [−] and [+] respectively.

[0026] FIG. 3—Purification of a LC/C-nociceptin-H_N/C fusion protein

[0027] Using the methodology outlined in Example 9, an LC/C-nociceptin-H_N/C fusion protein was purified from *E. coli* BL21 cells. Briefly, the soluble products obtained following cell disruption were applied to a nickel-charged affinity capture column. Bound proteins were eluted with 100 mM imidazole, treated with Factor Xa to activate the fusion protein and remove the maltose-binding protein (MBP) tag, then re-applied to a second nickel-charged affinity capture column. Samples from the purification procedure were assessed by SDS-PAGE (Panel A) and Western blotting (Panel B). Anti-nociceptin antisera (obtained from Abcam) were used as the primary antibody for Western blotting. The final purified material in the absence and presence of reducing agent is identified in the lanes marked [−] and [+] respectively.

[0028] FIG. 4—Purification of a LC/A-met enkephalin-H_N/A fusion protein

[0029] Using the methodology outlined in Example 9, an LC/A-met enkephalin-H_N/A fusion protein was purified from *E. coli* BL21 cells. Briefly, the soluble products obtained following cell disruption were applied to a nickel-charged affinity capture column. Bound proteins were eluted with 100 mM imidazole, treated with Factor Xa to activate the fusion protein and remove the maltose-binding protein (MBP) tag, then re-applied to a second nickel-charged affinity capture column. Samples from the purification procedure were assessed by SDS-PAGE. The final purified material in the absence and presence of reducing agent is identified in the lanes marked [−] and [+] respectively.

[0030] FIG. 5—Comparison of binding efficacy of a LC/A-nociceptin-H_N/A fusion protein and a nociceptin-LC/A-H_N/A fusion protein

[0031] The ability of nociceptin fusions to bind to the ORL₁ receptor was assessed using a simple competition-based assay. Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of test material in the presence of 1 nM [3H]-nociceptin. The reduction in specific binding of the radiolabelled ligand was assessed by scintillation counting, and plotted in comparison to the efficacy of unlabelled ligand (Tocris nociceptin). It is clear that the LC/A-nociceptin-H_N/A fusion is far superior to the nociceptin-LC/A-H_N/A fusion at interacting with the ORL₁ receptor.

[0032] FIG. 6—In vitro catalytic activity of a LC/A-nociceptin-H_N/A fusion protein

[0033] The in vitro endopeptidase activity of the purified LC/A-nociceptin-H_N/A fusion protein was determined essentially as described in Chaddock et al 2002, Prot. Express Purif. 25, 219-228. Briefly, SNAP-25 peptide immobilised to an ELISA plate was exposed to varying concentrations of fusion protein for 1 hour at 37° C. Following a series of washes, the amount of cleaved SNAP-25 peptide was quantified by reactivity with a specific antisera.

[0034] FIG. 7—Purification of a LC/A-nociceptin variant- H_N/A fusion protein

[0035] Using the methodology outlined in Example 9, an LC/A-nociceptin variant- H_N/A fusion protein was purified from *E. coli* BL21 cells. Briefly, the soluble products obtained following cell disruption were applied to a nickel-charged affinity capture column. Bound proteins were eluted with 100 mM imidazole, treated with Factor Xa to activate the fusion protein and remove the maltose-binding protein (MBP) tag, then re-applied to a second nickel-charged affinity capture column. Samples from the purification procedure were assessed by SDS-PAGE. The final purified material in the absence and presence of reducing agent is identified in the lanes marked [–] and [+] respectively.

[0036] FIG. 8—Comparison of binding efficacy of a LC/A-nociceptin- H_N/A fusion protein and a LC/A-nociceptin variant- H_N/A fusion protein

[0037] The ability of nociceptin fusions to bind to the ORL₁ receptor was assessed using a simple competition-based assay. Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of test material in the presence of 1 nM [3H]-nociceptin. The reduction in specific binding of the radiolabelled ligand was assessed by scintillation counting, and plotted in comparison to the efficacy of unlabelled ligand (Tocris nociceptin). It is clear that the LC/A-nociceptin variant- H_N/A fusion (CPNv-LHA) is superior to the LC/A-nociceptin variant- H_N/A fusion (CPN-LHA) at interacting with the ORL₁ receptor.

[0038] FIG. 9—Expressed/purified LC/A-nociceptin- H_N/A fusion protein family with variable spacer length product(s)

[0039] Using the methodology outlined in Example 9, variants of the LC/A-CPN- H_N/A fusion consisting of GS10, GS30 and HX27 are purified from *E. coli* cell paste. Samples from the purification of LC/A-CPN(GS10)- H_N/A , LC/A-CPN(GS15)- H_N/A , LC/A-CPN(GS25)- H_N/A , LC/A-CPN(GS30)- H_N/A and LC/A-CPN(HX27)- H_N/A were assessed by SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPBE-A. Top panel: M=benchmark molecular mass markers; S=total *E. coli* protein soluble fraction; FT=proteins that did not bind to the Ni²⁺-charged Sepharose column; FUSION=fusion protein eluted by the addition of imidazole. Bottom panel: Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=purified material following initial capture on Ni²⁺-charged Sepharose; Lane 4=Factor Xa treated material prior to final capture on Ni²⁺-charged Sepharose; Lane 5=purified final material post activation with Factor Xa (5 μ l); Lane 6=purified final material post activation with Factor Xa (10 μ l); Lane 7=purified final material post activation with Factor Xa (20 μ l); Lane 8=purified final material post activation with Factor Xa+DTT (5 μ l); Lane 9=purified final material post activation with Factor Xa+DTT (10 μ l); Lane 10=purified final material post activation with Factor Xa+DTT (20 μ l).

[0040] FIG. 10—Inhibition of SP release and cleavage of SNAP-25 by CPN-A

[0041] Briefly, primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPN-A for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis

and plotted against fusion concentration (dashed line). Material was also recovered for an analysis of substance P content using a specific EIA kit. Inhibition of substance P release is illustrated by the solid line. The fusion concentration required to achieve 50% maximal SNAP-25 cleavage is estimated to be 6.30 ± 2.48 nM.

[0042] FIG. 11—Inhibition of SP release and cleavage of SNAP-25 over extended time periods after exposure of DRG to CPN-A

[0043] Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPN-A for 24 hours. Botulinum neurotoxin (BoNT/A) was used as a control. After this initial exposure, extracellular material was removed by washing, and the cells incubated at 37° C. for varying periods of time. At specific time points, cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis and illustrated by the dotted lines. Material was also recovered for an analysis of substance P content using a specific EIA kit. Inhibition of substance P release is illustrated by the solid lines.

[0044] FIG. 12—Cleavage of SNAP-25 by CPNv-A Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPNv-A for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. The fusion concentration required to achieve 50% maximal SNAP-25 cleavage is estimated to be 1.38 ± 0.36 nM.

[0045] FIG. 13—Cleavage of SNAP-25 over extended time periods after exposure of DRG to CPNv-A Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPNv-A for 24 hours. CPN-A was used as a control. After this initial exposure, extracellular material was removed by washing, and the cells incubated at 37° C. for varying periods of time. At specific time points, cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis.

[0046] FIG. 14—CPNv-A fusion-mediated displacement of [3H]-nociceptin binding

[0047] The ability of nociceptin fusions to bind to the ORL₁ receptor was assessed using a simple competition-based assay. Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of test material in the presence of 1 nM [3H]-nociceptin. The reduction in specific binding of the radiolabelled ligand was assessed by scintillation counting, and plotted in comparison to the efficacy of unlabelled ligand (Tocris nociceptin). It is clear that the LC/A-nociceptin variant- H_N/A fusion (labelled as CPNv-LHnA) is superior to the LC/A-nociceptin- H_N/A fusion (labelled as CPN-LHnA) at interacting with the ORL₁ receptor.

[0048] FIG. 15—Expressed/purified CPNv(Ek)-A product

[0049] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPNv(Ek)-A. Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=purified material following initial capture on Ni²⁺-charged Sepharose; Lane 4=purified final material post activation with enterokinase (5 μ l); Lane

5=purified final material post activation with enterokinase (10 μ l); Lane 6=purified final material post activation with enterokinase (20 μ l); Lane 7=purified final material post activation with enterokinase+DTT (5 μ l); Lane 8=purified final material post activation with enterokinase+DTT (10 μ l); Lane 9=purified final material post activation with enterokinase+DTT (20 μ l).

[0050] FIG. 16—Cleavage of SNAP-25 by CPNv(Ek)-A

[0051] Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPNv(Ek)-A for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. CPNv-A as prepared in Example 9 was used for comparison purposes. The percentage cleavage of SNAP-25 by CPNv(Ek)-A (labelled as En activated) and CPNv-A (labelled as Xa activated) are illustrated.

[0052] FIG. 17—Expressed/purified CPNv-C product

[0053] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPNv-C. Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=purified material following initial capture on Ni^{2+} -charged Sepharose; Lane 4=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 5=purified material following second capture on Ni^{2+} -charged Sepharose; Lane 6=final purified material; Lane 7=final purified material+DTT; Lane 8=benchmark molecular mass markers.

[0054] FIG. 18—Cleavage of syntaxin by CPNv-C

[0055] Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPNv-C for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-syntaxin to facilitate an assessment of syntaxin cleavage. The percentage of cleaved syntaxin was calculated by densitometric analysis. The fusion concentration required to achieve 50% maximal syntaxin cleavage is estimated to be 3.13 ± 1.96 nM.

[0056] FIG. 19—CPN-A efficacy in the Acute Capsaicin-Induced Mechanical Allodynia model

[0057] The ability of an LC/A-nociceptin- H_{N} /A fusion (CPN/A) to inhibit capsaicin-induced mechanical allodynia was evaluated following subcutaneous intraplantar injection in the rat hind paw. Test animals were evaluated for paw withdrawal frequency (PWF %) in response to a 10 g Von Frey filament stimulus series (10 stimuli $\times$ 3 trials) prior to recruitment into the study (Pre-Treat); after subcutaneous intraplantar treatment with CPN/A but before capsaicin (Pre-CAP); and following capsaicin challenge post-injection of CPN/A (average of responses at 15' and 30'; CAP). Capsaicin challenge was achieved by injection of 10 μ L of a 0.3% solution. Sample dilutions were prepared in 0.5% BSA/saline.

[0058] FIG. 20—CPN-A efficacy in the Streptozotocin (STZ)-Induced Peripheral Diabetic Neuropathy (Neuropathic Pain) model Male Sprague-Dawley rats (250-300 g) are treated with 65 mg/kg STZ in citrate buffer (I.V.) and blood glucose and lipid are measured weekly to define the readiness of the model. Paw Withdrawal Threshold (PWT) is measured in response to a Von Frey filament stimulus series over a period of time. Allodynia is said to be established when the PWT on two consecutive test dates (separated by 1 week)

measures below 6 g on the scale. At this point, rats are randomized to either a saline group (negative efficacy control), gabapentin group (positive efficacy control) or a test group (CPN/A). Test materials (20-25 μ l are injected subcutaneously as a single injection (except gabapentin) and the PWT is measured at 1 day post-treatment and periodically thereafter over a 2 week period. Gabapentin (30 mg/kg i.p. @ 3 ml/kg injection volume) is injected daily, 2 hours prior to the start of PWT testing.

[0059] FIG. 21—CPNv-A efficacy in the Acute Capsaicin-Induced Mechanical Allodynia model

[0060] The ability of an LC/A-nociceptin variant- H_{N} /A fusion (CPNv/A) to inhibit capsaicin-induced mechanical allodynia was evaluated following subcutaneous intraplantar injection in the rat hind paw. Test animals were evaluated for paw withdrawal frequency (PWF %) in response to a 10 g Von Frey filament stimulus series (10 stimuli $\times$ 3 trials) prior to recruitment into the study (Pre-Treat), after subcutaneous intraplantar treatment with CPNv/A but before capsaicin (Pre-CAP), and following capsaicin challenge post-injection of CPNv/A (average of responses at 15' and 30'; CAP). Capsaicin challenge was achieved by injection of 10 μ L of a 0.3% solution. Sample dilutions were prepared in 0.5% BSA/saline. These data are expressed as a normalized paw withdrawal frequency differential, in which the difference between the peak response (post-capsaicin) and the baseline response (pre-capsaicin) is expressed as a percentage. With this analysis, it can be seen that CPNv/A is more potent than CPN/A since a lower dose of CPNv/A is required to achieve similar analgesic effect to that seen with CPN/A.

[0061] FIG. 22—Expressed/purified LC/A-CPLE- H_{N} /A product

[0062] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPLE-A. Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=purified material following initial capture on Ni^{2+} -charged Sepharose; Lane 4=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 5=purified material following second capture on Ni^{2+} -charged Sepharose; Lane 6=final purified material; Lane 7=final purified material+DTT.

[0063] FIG. 23—Expressed/purified LC/A-CPBE- H_{N} /A product

[0064] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPBE-A. Lane 1=total *E. coli* protein soluble fraction; Lane 2=purified material following initial capture on Ni^{2+} -charged Sepharose; Lane 3=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 4=purified final material post activation with Factor Xa (5 μ l; Lane 5=purified final material post activation with Factor Xa (10 μ l; Lane 6=purified final material post activation with Factor Xa (20 μ l; Lane 7=purified final material post activation with Factor Xa+DTT (5 μ l; Lane 8=purified final material post activation with Factor Xa+DTT (10 μ l; Lane 9=purified final material post activation with Factor Xa+DTT (20 μ l; Lane 10=benchmark molecular mass markers.

[0065] FIG. 24—Expressed/Purified CPOP-A Product

[0066] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile

indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPOP-A. Lane 1=benchmark molecular mass markers; Lane 2=purified material following initial capture on Ni^{2+} -charged Sepharose; Lane 3=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 4=purified material following second capture on Ni^{2+} -charged Sepharose; Lane 5=purified final material post activation with Factor Xa (5 μl); Lane 6=purified final material post activation with Factor Xa (10 μl); Lane 7=purified final material post activation with Factor Xa (20 μl); Lane 8=purified final material post activation with Factor Xa+DTT (5 μl); Lane 9=purified final material post activation with Factor Xa+DTT (10 μl); Lane 10=purified final material post activation with Factor Xa+DTT (20 μl).

[0067] FIG. 25—Expressed/purified CPOPv-A product

[0068] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPOPv-A. Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=purified material following initial capture on Ni^{2+} -charged Sepharose; Lane 4=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 5=purified final material post activation with Factor Xa (5 μl); Lane 6=purified final material post activation with Factor Xa (10 μl); Lane 7=purified final material post activation with Factor Xa (20 μl); Lane 8=purified final material post activation with Factor Xa+DTT (5 μl); Lane 9=purified final material post activation with Factor Xa+DTT (10 μl); Lane 10=purified final material post activation with Factor Xa+DTT (20 μl).

[0069] FIG. 26—In vitro SNAP-25 cleavage in a DRG cell model

[0070] Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPOPv-A for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis.

[0071] FIG. 27—Expressed/purified CPNv-A-FXa-HT (removable his-tag)

[0072] Proteins were subjected to SDS-PAGE prior to staining with Coomassie Blue. The electrophoresis profile indicates purification of a disulphide-bonded di-chain species of the expected molecular mass of CPNv-A-FXa-HT. Lane 1=benchmark molecular mass markers; Lane 2=total *E. coli* protein soluble fraction; Lane 3=Factor Xa treated material prior to final capture on Ni^{2+} -charged Sepharose; Lane 4=purified final material post activation with Factor Xa; Lane 5=purified final material post activation with Factor Xa+DTT.

[0073] FIG. 28—In vitro efficacy of LC/A-nociceptin- H_N /A fusion proteins with variable spacer length, as assessed by ligand competition assay

[0074] The ability of LC/A-nociceptin- H_N /A fusions of variable spacer length to bind to the ORL_1 receptor was assessed using a simple competition-based assay. Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of test material in the presence of 1 nM [^3H]-nociceptin. The reduction in specific binding of the radiolabelled ligand was assessed by scintillation counting, and plotted in comparison to the efficacy of unlabelled ligand (Tocris nociceptin). The upper panel illustrates the displacement characteristics of the GS0, GS20, GS30 and Hx27 spacers, whilst the lower panel illustrates the displacement

achieved by the GS10, GS15 and GS25 spaced fusion proteins. It is concluded that the GS0 and GS30 spacers are ineffective, and the GS10 is poorly effective, at displacing nociceptin from the ORL_1 receptor.

[0075] FIG. 29—In vitro efficacy of LC/A-nociceptin- H_N /A fusion proteins with variable spacer length, as assessed by in vitro SNAP-25 cleavage

[0076] Primary cultures of dorsal root ganglia (DRG) were exposed to varying concentrations of CPN-A (of variable spacer length) for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. The poorly effective binding characteristics of the GS10 spaced fusion protein (see FIG. 28) are reflected in the higher concentrations of fusion required to achieve cleavage of intracellular SNAP-25. GS0 and GS30 spaced fusion proteins were completely ineffective (date not shown). GS15, 20 and 25 spaced fusion proteins were similarly effective.

[0077] FIG. 30—Cleavage of SNARE protein by dynorphin conjugates in embryonic spinal cord neurons (eSCNs)

[0078] Embryonic spinal cord neurons were exposed to varying concentrations of dynorphin conjugates of the present invention for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. It is clear that LC/A-dynorphin- H_N /A fusion is more potent than an unliganded LC/A- H_N /A control molecule. The concentration of LC/A-dynorphin- H_N /A fusion required to achieve 50% maximal SNAP-25 cleavage is estimated to be 35.3 nM and the concentration for the LC/A- H_N /A control required to achieve 50% maximal SNAP-25 cleavage could not be determined due to its low potency.

[0079] FIG. 31—Cleavage of SNARE protein by dynorphin conjugates in Chinese hamster ovary cells (CHO-K1 cells) transfected with OP2 receptor and SNAP-25

[0080] Chinese hamster ovary (CHO) cells were transfected so that they express the OP2 receptor. Said cells were further transfected to express a SNARE protein (SNAP-25). The transfected cells were exposed to varying concentrations of different dynorphin conjugates for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. It is clear that LC/A-CPDY- H_N /A conjugates are more potent than the unliganded LC/A- H_N /A control molecule (labelled as LC/A- H_N /A).

[0081] FIG. 32—Cleavage of SNARE protein by dynorphin conjugates in embryonic spinal cord neurons (eSCNs)

[0082] Embryonic spinal cord neurons were exposed to varying concentrations of dynorphin conjugates of the present invention for 24 hours. Cellular proteins were separated by SDS-PAGE, Western blotted, and probed with anti-SNAP-25 to facilitate an assessment of SNAP-25 cleavage. The percentage of cleaved SNAP-25 was calculated by densitometric analysis. It is clear that LC/A-CPDY- H_N /A conjugates are more potent than the unliganded LC/A- H_N /A control molecule (labelled as LC/A- H_N /A).

[0083] FIG. 33—Kappa receptor activation studies with a range of dynorphin conjugates

[0084] Chinese hamster ovary (CHO) cells were transfected so that they express the OP2 receptor and SNAP-25. Said cells were used to measure cAMP deletion that occurs when the receptor is activated with a dynorphin ligand, using a FRET-based cAMP kit (LANCER kit from Perkin Elmer). The transfected cells were exposed to varying concentrations of dynorphin conjugates of the present invention for 2 hours. cAMP levels were then detected by addition of a detection mix containing a fluorescently labelled cAMP tracer (Europium-streptavidin/biotin-cAMP) and fluorescently (Alexa) labelled anti-cAMP antibody and incubating at room temperature for 24 hours. Then samples are excited at 320 nm and emitted light measured at 665 nm to determine cAMP levels. It is clear that LC/A-CPDY-H_N/A conjugates are more potent than the unliganded LC/A-H_N/A control molecule (labelled as LC/A-H_N/A).

[0085] FIG. 34—Kappa receptor activation studies with a range of dynorphin conjugates

[0086] Chinese hamster ovary (CHO) cells were transfected so that they express the OP2 receptor (purchased from Perkin Elmer). Said cells were transfected so they express SNAP-25 and used to measure cAMP deletion that occurs when the receptor is activated with a dynorphin ligand, using a FRET-based cAMP kit (LANCER kit from Perkin Elmer). The transfected cells were exposed to varying concentrations of dynorphin conjugates of the present invention for 2 hours. cAMP levels were then detected by addition of a detection mix containing a fluorescently labelled cAMP tracer (Europium-streptavidin/biotin-cAMP) and fluorescently (Alexa) labelled anti-cAMP antibody and incubating at room temperature for 24 hours. Then samples are excited at 320 nm and emitted light measured at 665 nm to determine cAMP levels. It is clear from the figure by the reduction in maximum cAMP that the OP2 receptor is activated by LC/A-CPDY-H_N/A (labelled as CPDY/A), LC/B-CPDY-H_N/B (labelled as CPDY/B), LC/C-CPDY-H_N/C (labelled as CPDY/C), and LC/D-CPDY-H_N/D (labelled as CPDY/D). The concentration required to achieve 50% reduction in cAMP with LC/A-CPDY-H_N/A, LC/B-CPDY-H_N/B, LC/C-CPDY-H_N/C (labelled as CPDY/), and LC/D-CPDY-H_N/D is 10.47 nM, 14.79 nM, 14.79 nM and 23.99 nM, respectively. Dynorphin peptide containing amino acids 1-17 of dynorphin A (labelled as dynorphin (1-17)) was more potent than the fusions; 0.15 nM concentration required to achieve 50% reduction of cAMP.

DETAILED DESCRIPTION OF THE INVENTION

[0087] The present inventors have found that the WO 98/07864 fusion protein system is not optimal for TMs requiring a N-terminal domain for interaction with a binding site on a nociceptive sensory afferent. This problem is particularly acute with TMs that require a specific N-terminus amino acid residue or a specific sequence of amino acid residues including the N-terminus amino acid residue for interaction with a binding site on a nociceptive sensory afferent.

[0088] In contrast to WO98/07864, the present invention provides a system for preparing non-cytotoxic conjugates, wherein the TM component of the fusion includes the relevant binding domain in an intra domain or an amino acid sequence located towards the middle (i.e. of the linear peptide sequence) of the TM, or preferably located towards the N-terminus of the TM, or more preferably at or near to the N-terminus. The N-terminal domain is capable of binding to a

Binding Site on a nociceptive sensory afferent, and the TM preferably has a requirement for a specific and defined sequence of amino acid residue(s) to be free at its N-terminus.

[0089] The non-cytotoxic protease component of the present invention is a non-cytotoxic protease, or a fragment thereof, which protease or protease fragment is capable of cleaving different but specific peptide bonds in one of three substrate proteins, namely synaptobrevin, syntaxin or SNAP-25, of the exocytic fusion apparatus in a nociceptive sensory afferent. These substrates are important components of the neurosecretory machinery. The non-cytotoxic protease component of the present invention is preferably a neisserial IgA protease or a fragment thereof or a clostridial neurotoxin L-chain or a fragment thereof. A particularly preferred non-cytotoxic protease component is a botulinum neurotoxin (BoNT) L-chain or a fragment thereof.

[0090] The translocation component of the present invention enables translocation of the non-cytotoxic protease (or fragment thereof) into the target cell such that functional expression of protease activity occurs within the cytosol of the target cell. The translocation component is preferably capable of forming ion-permeable pores in lipid membranes under conditions of low pH. Preferably it has been found to use only those portions of the protein molecule capable of pore-formation within the endosomal membrane. The translocation component may be obtained from a microbial protein source, in particular from a bacterial or viral protein source. Hence, in one embodiment, the translocation component is a translocating domain of an enzyme, such as a bacterial toxin or viral protein. The translocation component of the present invention is preferably a clostridial neurotoxin H-chain or a fragment thereof. Most preferably it is the H_N domain (or a functional component thereof), wherein H_N means a portion or fragment of the H-chain of a clostridial neurotoxin approximately equivalent to the amino-terminal half of the H-chain, or the domain corresponding to that fragment in the intact H-chain.

[0091] The TM component of the present invention is responsible for binding the fusion protein of the present invention to a Binding Site on a target cell. Thus, the TM component is simply a ligand through which a fusion protein of the present invention binds to a selected target cell.

[0092] In the context of the present invention, the target cell is a nociceptive sensory afferent, preferably a primary nociceptive afferent (e.g. an A-fibre such as an Aδ-fibre or a C-fibre). Thus, the fusion proteins of the present invention are capable of inhibiting neurotransmitter or neuromodulator [e.g. glutamate, substance P, calcitonin-gene related peptide (CGRP), and/or neuropeptide Y] release from discrete populations of nociceptive sensory afferent neurons. In use, the fusion proteins reduce or prevent the transmission of sensory afferent signals (e.g. neurotransmitters or neuromodulators) from peripheral to central pain fibres, and therefore have application as therapeutic molecules for the treatment of pain, in particular chronic pain.

[0093] It is routine to confirm that a TM binds to a nociceptive sensory afferent. For example, a simple radioactive displacement experiment may be employed in which tissue or cells representative of the nociceptive sensory afferent (for example DRGs) are exposed to labelled (e.g. tritiated) ligand in the presence of an excess of unlabelled ligand. In such an experiment, the relative proportions of non-specific and specific binding may be assessed, thereby allowing confirmation that the ligand binds to the nociceptive sensory afferent target

cell. Optionally, the assay may include one or more binding antagonists, and the assay may further comprise observing a loss of ligand binding. Examples of this type of experiment can be found in Hulme, E. C. (1990), *Receptor-binding studies*, a brief outline, pp. 303-311, In *Receptor biochemistry, A Practical Approach*, Ed. E. C. Hulme, Oxford University Press.

[0094] The fusion proteins of the present invention generally demonstrate a reduced binding affinity (in the region of up to 100-fold) for nociceptive sensory afferent target cells when compared with the corresponding 'free' TM. However, despite this observation, the fusion proteins of the present invention surprisingly demonstrate good efficacy. This can be attributed to two principal features. First, the non-cytotoxic protease component is catalytic—thus, the therapeutic effect of a few such molecules is rapidly amplified. Secondly, the receptors present on the nociceptive sensory afferents need only act as a gateway for entry of the therapeutic, and need not necessarily be stimulated to a level required in order to achieve a ligand-receptor mediated pharmacological response. Accordingly, the fusion proteins of the present invention may be administered at a dosage that is much lower than would be employed for other types of analgesic molecules such as NSAIDs, morphine, and gabapentin. The latter molecules are typically administered at high microgram to milligram (even up to hundreds of milligram) quantities, whereas the fusion proteins of the present invention may be administered at much lower dosages, typically at least 10-fold lower, and more typically at 100-fold lower.

[0095] The TM preferably comprises a maximum of 50 amino acid residues, more preferably a maximum of 40 amino acid residues, particularly preferably a maximum of 30 amino acid residues, and most preferably a maximum of 20 amino acid residues.

[0096] Opioids represent a preferred group of TMs of the present invention. Within this family of peptides is included enkephalins (met and leu), endomorphins 1 and 2, β -endorphin and dynorphin. Opioid peptides are frequently used in the clinic to modify the activity to nociceptors, and other cells involved in the pain response. As exemplified by the three-step World Health Organisation Analgesic Ladder, opioids have entry points into the pharmacological treatment of chronic cancer and non-cancer pain at all three stages, underlining their importance to the treatment of pain. Reference to opioids embraces fragments, variants and derivatives thereof, which retain the ability to bind to nociceptive sensory afferents.

[0097] The TM of the invention can also be a molecule that acts as an "agonist" at one or more of the receptors present on a nociceptive sensory afferent, more particularly on a primary nociceptive afferent. Conventionally, an agonist has been considered any molecule that can either increase or decrease activities within a cell, namely any molecule that simply causes an alteration of cell activity. For example, the conventional meaning of an agonist would include a chemical substance capable of combining with a receptor on a cell and initiating a reaction or activity, or a drug that induces an active response by activating receptors, whether the response is an increase or decrease in cellular activity.

[0098] However, for the purposes of this invention, an agonist is more specifically defined as a molecule that is capable of stimulating the process of exocytic fusion in a target cell, which process is susceptible to inhibition by a protease (or fragment thereof) capable of cleaving a protein of the exocytic fusion apparatus in said target cell.

[0099] Accordingly, the particular agonist definition of the present invention would exclude many molecules that would be conventionally considered as agonists. For example, nerve growth factor (NGF) is an agonist in respect of its ability to promote neuronal differentiation via binding to a TrkA receptor. However, NGF is not an agonist when assessed by the above criteria because it is not a principal inducer of exocytic fusion. In addition, the process that NGF stimulates (i.e. cell differentiation) is not susceptible to inhibition by the protease activity of a non-cytotoxic toxin molecule.

[0100] The agonist properties of a TM that binds to a receptor on a nociceptive afferent can be confirmed using the methods described in Example 10.

[0101] In a preferred embodiment of the invention, the target for the TM is the ORL₁ receptor. This receptor is a member of the G-protein-coupled class of receptors, and has a seven transmembrane domain structure. The properties of the ORL₁ receptor are discussed in detail in Mogil & Pasternak (2001), *Pharmacological Reviews*, Vol. 53, No. 3, pages 381-415.

[0102] In one embodiment, the TM is a molecule that binds (preferably that specifically binds) to the ORL₁ receptor. More preferably, the TM is an "agonist" of the ORL₁ receptor. The term "agonist" in this context is defined as above.

[0103] The agonist properties of a TM that binds to an ORL₁ receptor can be confirmed using the methods described in Example 10. These methods are based on previous experiments [see Inoue et al. 1998 [Proc. Natl. Acad. Sci., 95, 10949-10953]], which confirm that the natural agonist of the ORL₁ receptor, nociceptin, causes the induction of substance P release from nociceptive primary afferent neurons. This is supported by the fact that:

[0104] the nociceptin-induced responses are abolished by specific NK1 receptor (the substance P receptor) antagonists; and

[0105] pre-treatment of the cells with capsaicin (which depletes substance P from small diameter primary afferent neurons) attenuates the nociceptin-induced responses.

[0106] Similarly, Inoue et al. confirm that an intraplantar injection of botulinum neurotoxin type A abolishes the nociceptin-induced responses. Since it is known that BoNT inhibits the release of substance P from primary afferent neurons (Welch et al., 2000, *Toxicon*, 38, 245-258), this confirms the link between nociceptin-ORL₁ interaction and subsequent release of substance P.

[0107] Thus, a TM can be said to have agonist activity at the ORL₁ receptor if the TM causes an induction in the release of substance P from a nociceptive sensory afferent neuron (see Example 10).

[0108] In a particularly preferred embodiment of the invention, the TM is nociceptin—the natural ligand for the ORL₁ receptor. Nociceptin targets the ORL₁ receptor with high affinity. Examples of other preferred TMs include:

Code	Sequence	Ref.	SEQ ID No.
Nociceptin 1-17	FGGFTGARKSARKLANQ	[1]	37, 38
Nociceptin 1-11	FGGFTGARKSA	[1]	39, 40
Nociceptin [Y10]1-11	FGGFTGARKYA	[1]	41, 42
Nociceptin [Y11]1-11	FGGFTGARKSY	[1]	43, 44
Nociceptin [Y14]1-17	FGGFTGARKSARKYANQ	[1]	45, 46
Nociceptin 1-13	FGGFTGARKSARK	[2]	47, 48
Nociceptin [R14K15] 1-17 (also known in this specification as "variant" nociceptin)	FGGFTGARKSARKRKNQ	[3, 4]	49, 50
Peptide agonist	Peptide agonists from [5] combinatorial library approach	—	

[1] Mogil & Pasternak, 2001, Pharmacol. Rev., 53, 381-415

[2] Maile et al., 2003, Neurosci. Lett., 350, 190-192

[3] Rizzi et al., 2002, J. Pharmacol. Exp. Therap., 300, 57-63

[4] Okada et al., 2000, Biochem. Biophys. Res. Commun., 278, 493-498

[5] Dooley et al., 1997, J Pharmacol Exp Ther. 283(2), 735-41.

[0109] The above-identified "variant" TM demonstrates particularly good binding affinity (when compared with natural nociceptin) for nociceptive sensory afferents. This is surprising as the amino acid modifications occur at a position away from the N-terminus of the TM. Moreover, the modifications are almost at the C-terminus of the TM, which in turn is attached to a large polypeptide sequence (i.e. the translocation domain). Generally speaking, a TM-containing fusion protein will demonstrate an approximate 100-fold reduction in binding ability vis-à-vis the TM per se. The above-mentioned "variant" TM per se demonstrates an approximate 3- to 10-fold increase in binding ability for a nociceptive sensory afferent (e.g. via the ORL1 receptor) vis-à-vis natural nociceptin. Thus, a "variant" TM-containing fusion might be expected to demonstrate an approximate 10-fold reduction in binding ability for a nociceptive sensory afferent (e.g. via the ORL1 receptor) vis-à-vis 'free' nociceptin. However, the present inventors have demonstrated that such "variant" TM-containing fusion proteins demonstrate a binding ability that (most surprisingly) closely mirrors that of 'free' nociceptin—see FIG. 14.

[0110] In the context of the present invention, the term opioid or an agonist of the ORL₁ receptor (such as nociceptin, or any one of the peptides listed in the table above) embraces molecules having at least 70%, preferably at least 80%, more preferably at least 90%, and most preferably at least 95% homology with said opioid or agonist. The agonist homologues retain the agonist properties of nociceptin at the ORL₁ receptor, which may be tested using the methods provided in Example 10. Similarly, an opioid homologue substantially retains the binding function of the opioid with which it shows high homology.

[0111] The invention also encompasses fragments, variants, and derivatives of any one of the TMs described above. These fragments, variants, and derivatives substantially retain the properties that are ascribed to said TMs.

[0112] In addition to the above-mentioned opioid and non-opioid classes of TMs, a variety of other polypeptides are suitable for targeting the fusion proteins of the present inven-

tion to nociceptive sensory afferents (e.g. to nociceptors). In this regard, particular reference is made to galanin and derivatives of galanin. Galanin receptors are found pre- and post-synaptically in DRGs (Liu & Hokfelt, (2002), Trends Pharm. Sci., 23(10), 468-74), and are enhanced in expression during neuropathic pain states. Proteinase-activated receptors (PARs) are also a preferred group of TMs of the present invention, most particularly PAR-2. It is known that agonists of PAR-2 induce/elicite acute inflammation, in part via a neurogenic mechanism. PAR2 is expressed by primary spinal afferent neurons, and PAR2 agonists stimulate release of substance P(SP) and calcitonin gene-related peptide (CGRP) in peripheral tissues

[0113] A particularly preferred set of TMs of the present invention includes:

Ligand	Reference
Nociceptin	Guerrini, et al., (1997) J. Med. Chem., 40, pp. 1789-1793
β-endorphin	Blanc, et al., (1983) J. Biol. Chem., 258(13), pp. 8277-8284
Endomorphin-1; Endomorphin-2 Dynorphin	Zadina, et al., (1997). Nature, 386, pp. 499-502 Fields & Basbaum (2002) Chapter 11, In The Textbook of Pain, Wall & Melzack eds.
Met-enkephalin	Fields & Basbaum (2002) Chapter 11, In The Textbook of Pain, Wall & Melzack eds.
Leu-enkephalin	Fields & Basbaum (2002) Chapter 11, In The Textbook of Pain, Wall & Melzack eds.
Galanin	Xu et al., (2000) Neuropeptides, 34 (3&4), 137-147
PAR-2 peptide	Vergnolle et al., (2001) Nat. Med., 7(7), 821-826

[0114] The protease cleavage site of the present invention allows cleavage (preferably controlled cleavage) of the fusion

protein at a position between the non-cytotoxic protease component and the TM component. It is this cleavage reaction that converts the fusion protein from a single chain polypeptide into a disulphide-linked, di-chain polypeptide.

[0115] According to a preferred embodiment of the present invention, the TM binds via a domain or amino acid sequence that is located away from the C-terminus of the TM. For example, the relevant binding domain may include an intra domain or an amino acid sequence located towards the middle (i.e. of the linear peptide sequence) of the TM. Preferably, the relevant binding domain is located towards the N-terminus of the TM, more preferably at or near to the N-terminus.

[0116] In one embodiment, the single chain polypeptide fusion may include more than one proteolytic cleavage site. However, where two or more such sites exist, they are different, thereby substantially preventing the occurrence of multiple cleavage events in the presence of a single protease. In another embodiment, it is preferred that the single chain polypeptide fusion has a single protease cleavage site.

[0117] The protease cleavage sequence(s) may be introduced (and/or any inherent cleavage sequence removed) at the DNA level by conventional means, such as by site-directed mutagenesis. Screening to confirm the presence of cleavage sequences may be performed manually or with the assistance of computer software (e.g. the MapDraw program by DNASTAR, Inc.).

[0118] Whilst any protease cleavage site may be employed, the following are preferred:

Enterokinase	(DDDDK↓)
Factor Xa	(IEGR↓/IDGR↓)
TEV(Tobacco Etch virus)	(ENLYFQ↓G)
Thrombin	(LVPR↓GS)
PreScission	(LEVLFQ↓GP).

[0119] Also embraced by the term protease cleavage site is an intein, which is a self-cleaving sequence. The self-splicing reaction is controllable, for example by varying the concentration of reducing agent present.

[0120] In use, the protease cleavage site is cleaved and the N-terminal region (preferably the N-terminus) of the TM becomes exposed. The resulting polypeptide has a TM with an N-terminal domain or an intra domain that is substantially free from the remainder of the fusion protein. This arrangement ensures that the N-terminal component (or intra domain) of the TM may interact directly with a Binding Site on a target cell.

[0121] In a preferred embodiment, the TM and the protease cleavage site are distanced apart in the fusion protein by at most 10 amino acid residues, more preferably by at most 5 amino acid residues, and most preferably by zero amino acid residues. Thus, following cleavage of the protease cleavage site, a fusion is provided with a TM that has an N-terminal domain that is substantially free from the remainder of the fusion. This arrangement ensures that the N-terminal component of the Targeting Moiety may interact directly with a Binding Site on a target cell.

[0122] One advantage associated with the above-mentioned activation step is that the TM only becomes susceptible to N-terminal degradation once proteolytic cleavage of the fusion protein has occurred. In addition, the selection of a specific protease cleavage site permits selective activation of the polypeptide fusion into a di-chain conformation.

[0123] Construction of the single-chain polypeptide fusion of the present invention places the protease cleavage site between the TM and the non-cytotoxic protease component.

[0124] It is preferred that, in the single-chain fusion, the TM is located between the protease cleavage site and the translocation component. This ensures that the TM is attached to the translocation domain (i.e. as occurs with native clostridial holotoxin), though in the case of the present invention the order of the two components is reversed vis-à-vis native holotoxin. A further advantage with this arrangement is that the TM is located in an exposed loop region of the fusion protein, which has minimal structural effects on the conformation of the fusion protein. In this regard, said loop is variously referred to as the linker, the activation loop, the inter-domain linker, or just the surface exposed loop (Schiavo et al 2000, Phys. Rev., 80, 717-766; Turton et al., 2002, Trends Biochem. Sci., 27, 552-558).

[0125] In one embodiment, in the single chain polypeptide, the non-cytotoxic protease component and the translocation component are linked together by a disulphide bond. Thus, following cleavage of the protease cleavage site, the polypeptide assumes a di-chain conformation, wherein the protease and translocation components remain linked together by the disulphide bond. To this end, it is preferred that the protease and translocation components are distanced apart from one another in the single chain fusion protein by a maximum of 100 amino acid residues, more preferably a maximum of 80 amino acid residues, particularly preferably by a maximum of 60 amino acid residues, and most preferably by a maximum of 50 amino acid residues.

[0126] In one embodiment, the non-cytotoxic protease component forms a disulphide bond with the translocation component of the fusion protein. For example, the amino acid residue of the protease component that forms the disulphide bond is located within the last 20, preferably within the last 10 C-terminal amino acid residues of the protease component. Similarly, the amino acid residue within the translocation component that forms the second part of the disulphide bond may be located within the first 20, preferably within the first 10 N-terminal amino acid residues of the translocation component.

[0127] Alternatively, in the single chain polypeptide, the non-cytotoxic protease component and the TM may be linked together by a disulphide bond. In this regard, the amino acid residue of the TM that forms the disulphide bond is preferably located away from the N-terminus of the TM, more preferably towards to C-terminus of the TM.

[0128] In one embodiment, the non-cytotoxic protease component forms a disulphide bond with the TM component of the fusion protein. In this regard, the amino acid residue of the protease component that forms the disulphide bond is preferably located within the last 20, more preferably within the last 10 C-terminal amino acid residues of the protease component. Similarly, the amino acid residue within the TM component that forms the second part of the disulphide bond is preferably located within the last 20, more preferably within the last 10 C-terminal amino acid residues of the TM.

[0129] The above disulphide bond arrangements have the advantage that the protease and translocation components are arranged in a manner similar to that for native clostridial neurotoxin. By way of comparison, referring to the primary amino acid sequence for native clostridial neurotoxin, the respective cysteine amino acid residues are distanced apart by between 8 and 27 amino acid residues—taken from Popoff,

MR & Marvaud, J-C, 1999, Structural & genomic features of clostridial neurotoxins, Chapter 9, in The Comprehensive Sourcebook of Bacterial Protein Toxins. Ed. Alouf & Freer:

Serotype ¹	Sequence	'Native' length between C-C
BoNT/A1	CVRGIITSKTKS----LDKGYNKALNDLC	23
BoNT/A2	CVRGIIPFKTKS----LDEGYNKALNDLC	23
BoNT/B	CKSVKAPG-----IC	8
BoNT/C	CHKAIIDGRS-----LYNKTLDLC	15
BoNT/D	CLRLTK-----NSRDDSTC	12
BoNT/E	CKN-IVSVK-----GIRK---SIC	13
BoNT/F	CKS-VIPRK-----GTKAPP-RLC	15
BoNT/G	CKPVMYKNT-----GKSE----QC	13
TeNT	CKKIIPPTNIRENLYNRTASLTDLGELC	27

¹Information from proteolytic strains only

[0130] The fusion protein may comprise one or more purification tags, which are located N-terminal to the protease component and/or C-terminal to the translocation component.

[0131] Whilst any purification tag may be employed, the following are preferred:

His-tag (e.g. 6× histidine), preferably as a C-terminal and/or N-terminal tag

MBP-tag (maltose binding protein), preferably as an N-terminal tag

GST-tag (glutathione-S-transferase), preferably as an N-terminal tag

His-MBP-tag, preferably as an N-terminal tag

GST-MBP-tag, preferably as an N-terminal tag

Thioredoxin-tag, preferably as an N-terminal tag

CBD-tag (Chitin Binding Domain), preferably as an N-terminal tag.

[0132] According to a further embodiment of the present invention, one or more peptide spacer molecules may be included in the fusion protein. For example, a peptide spacer may be employed between a purification tag and the rest of the fusion protein molecule (e.g. between an N-terminal purification tag and a protease component of the present invention; and/or between a C-terminal purification tag and a translocation component of the present invention). A peptide spacer may be also employed between the TM and translocation components of the present invention.

[0133] A variety of different spacer molecules may be employed in any of the fusion proteins of the present invention. Examples of such spacer molecules include those illustrated in FIGS. 28 and 29. Particular mention here is made to GS15, GS20, GS25, and Hx27—see FIGS. 28 and 29.

[0134] The present inventors have unexpectedly found that the fusion proteins (e.g. CPNv/A) of the present invention may demonstrate an improved binding activity for nociceptive sensory afferents when the size of the spacer is selected so that (in use) the C-terminus of the TM and the N-terminus of the translocation component are separated from one another by 40-105 angstroms, preferably by 50-100 angstroms, and

more preferably by 50-90 angstroms. In another embodiment, the preferred spacers have an amino acid sequence of 11-29 amino acid residues, preferably 15-27 amino acid residues, and more preferably 20-27 amino acid residues. Suitable spacers may be routinely identified and obtained according to Crasto, C. J. and Feng, J. A. (2000) May, 13(5), pp. 309-312—see also <http://www.fccc.edu/research/labs/feng/linker.html>.

[0135] In accordance with a second aspect of the present invention, there is provided a DNA sequence that encodes the above-mentioned single chain polypeptide. In a preferred aspect of the present invention, the DNA sequence is prepared as part of a DNA vector, wherein the vector comprises a promoter and terminator.

[0136] In a preferred embodiment, the vector has a promoter selected from:

Promoter	Induction Agent	Typical Induction Condition
Tac (hybrid)	IPTG	0.2 mM (0.05-2.0 mM)
AraBAD	L-arabinose	0.2% (0.002-0.4%)
T7-lac operator	IPTG	0.2 mM (0.05-2.0 mM)

[0137] The DNA construct of the present invention is preferably designed in silico, and then synthesised by conventional DNA synthesis techniques.

[0138] The above-mentioned DNA sequence information is optionally modified for codon-biasing according to the ultimate host cell (e.g. *E. coli*) expression system that is to be employed.

[0139] The DNA backbone is preferably screened for any inherent nucleic acid sequence, which when transcribed and translated would produce an amino acid sequence corresponding to the protease cleave site encoded by the second peptide-coding sequence. This screening may be performed manually or with the assistance of computer software (e.g. the MapDraw program by DNASTAR, Inc.).

[0140] According to a further embodiment of the present invention, there is provided a method of preparing a non-cytotoxic agent, comprising:

[0141] a. contacting a single-chain polypeptide fusion protein of the invention with a protease capable of cleaving the protease cleavage site;

[0142] b. cleaving the protease cleavage site, and thereby forming a di-chain fusion protein.

[0143] This aspect provides a di-chain polypeptide, which generally mimics the structure of clostridial holotoxin. In more detail, the resulting di-chain polypeptide typically has a structure wherein:

[0144] a. the first chain comprises the non-cytotoxic protease, or a fragment thereof, which protease or protease fragment is capable of cleaving a protein of the exocytic fusion apparatus of a nociceptive sensory afferent;

[0145] b. the second chain comprises the TM and the translocation domain that is capable of translocating the protease or protease fragment from within an endosome, across the endosomal membrane and into the cytosol of the nociceptive sensory afferent; and

[0146] the first and second chains are disulphide linked together.

[0147] In use, the single chain or di-chain polypeptide of the invention treat, prevent or ameliorate pain.

[0148] In use, a therapeutically effective amount of a single chain or di-chain polypeptide of the invention is administered to a patient.

[0149] According to a further aspect of the present invention, there is provided use of a single chain or di-chain polypeptide of the invention, for the manufacture of a medicament for treating, preventing or ameliorating pain.

[0150] According to a related aspect, there is provided a method of treating, preventing or ameliorating pain in a subject, comprising administering to said patient a therapeutically effective amount of a single chain or di-chain polypeptide of the invention.

[0151] The compounds described here may be used to treat a patient suffering from one or more types of chronic pain including neuropathic pain, inflammatory pain, headache pain, somatic pain, visceral pain, and referred pain.

[0152] To “treat,” as used here, means to deal with medically. It includes, for example, administering a compound of the invention to prevent pain or to lessen its severity.

[0153] The term “pain,” as used here, means any unpleasant sensory experience, usually associated with a physical disorder. The physical disorder may or may not be apparent to a clinician. Pain is of two types: chronic and acute. An “acute pain” is a pain of short duration having a sudden onset. One type of acute pain, for example, is cutaneous pain felt on injury to the skin or other superficial tissues, such as caused by a cut or a burn. Cutaneous nociceptors terminate just below the skin, and due to the high concentration of nerve endings, produce a well-defined, localized pain of short duration. “Chronic pain” is a pain other than an acute pain. Chronic pain includes neuropathic pain, inflammatory pain, headache pain, somatic pain visceral pain and referred pain.

I. Neuropathic Pain

[0154] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the following neuropathic pain conditions. “Neuropathic pain” means abnormal sensory input, resulting in discomfort, from the peripheral nervous system, central nervous systems, or both.

A. Symptoms of Neuropathic Pain

[0155] Symptoms of neuropathic pain can involve persistent, spontaneous pain, as well as allodynia (a painful response to a stimulus that normally is not painful), hyperalgesia (an accentuated response to a painful stimulus that usually causes only a mild discomfort, such as a pin prick), or hyperpathia (where a short discomfort becomes a prolonged severe pain).

B. Causes of Neuropathic Pain

[0156] Neuropathic pain may be caused by any of the following.

1. A traumatic insult, such as, for example, a nerve compression injury (e.g., a nerve crush, a nerve stretch, a nerve entrapment or an incomplete nerve transection); a spinal cord injury (e.g., a hemisection of the spinal cord); a limb amputation; a contusion; an inflammation (e.g., an inflammation of the spinal cord); or a surgical procedure.
2. An ischemic event, including, for example, a stroke and heart attack.
3. An infectious agent
4. Exposure to a toxic agent, including, for example, a drug, an alcohol, a heavy metal (e.g., lead, arsenic, mercury), an industrial agent (e.g., a solvent, fumes from a glue) or nitrous oxide.

5. A disease, including, for example, an inflammatory disorder, a neoplastic tumor, an acquired immune deficiency syndrome (AIDS), Lyme’s disease, a leprosy, a metabolic disease, a peripheral nerve disorder, like neuroma, a mononeuropathy or a polyneuropathy.

C. Types of Neuropathic Pain

1. Neuralgia.

[0157] A neuralgia is a pain that radiates along the course of one or more specific nerves usually without any demonstrable pathological change in the nerve structure. The causes of neuralgia are varied. Chemical irritation, inflammation, trauma (including surgery), compression by nearby structures (for instance, tumors), and infections may all lead to neuralgia. In many cases, however, the cause is unknown or unidentifiable. Neuralgia is most common in elderly persons, but it may occur at any age. A neuralgia, includes, without limitation, a trigeminal neuralgia, a post-herpetic neuralgia, a postherpetic neuralgia, a glossopharyngeal neuralgia, a sciatica and an atypical facial pain.

[0158] Neuralgia is pain in the distribution of a nerve or nerves. Examples are trigeminal neuralgia, atypical facial pain, and postherpetic neuralgia (caused by shingles or herpes). The affected nerves are responsible for sensing touch, temperature and pressure in the facial area from the jaw to the forehead. The disorder generally causes short episodes of excruciating pain, usually for less than two minutes and on only one side of the face. The pain can be described in a variety of ways such as “stabbing,” “sharp,” “like lightning,” “burning,” and even “itchy”. In the atypical form of TN, the pain can also present as severe or merely aching and last for extended periods. The pain associated with TN is recognized as one the most excruciating pains that can be experienced.

[0159] Simple stimuli such as eating, talking, washing the face, or any light touch or sensation can trigger an attack (even the sensation of a gentle breeze). The attacks can occur in clusters or as an isolated attack.

[0160] Symptoms include sharp, stabbing pain or constant, burning pain located anywhere, usually on or near the surface of the body, in the same location for each episode; pain along the path of a specific nerve; impaired function of affected body part due to pain, or muscle weakness due to concomitant motor nerve damage; increased sensitivity of the skin or numbness of the affected skin area (feeling similar to a local anesthetic such as a Novacaine shot); and any touch or pressure is interpreted as pain. Movement may also be painful.

[0161] Trigeminal neuralgia is the most common form of neuralgia. It affects the main sensory nerve of the face, the trigeminal nerve (“trigeminal” literally means “three origins”, referring to the division of the nerve into 3 branches). This condition involves sudden and short attacks of severe pain on the side of the face, along the area supplied by the trigeminal nerve on that side. The pain attacks may be severe enough to cause a facial grimace, which is classically referred to as a painful tic (tic douloureux). Sometimes, the cause of trigeminal neuralgia is a blood vessel or small tumor pressing on the nerve. Disorders such as multiple sclerosis (an inflammatory disease affecting the brain and spinal cord), certain forms of arthritis, and diabetes (high blood sugar) may also cause trigeminal neuralgia, but a cause is not always identified. In this condition, certain movements such as chewing, talking, swallowing, or touching an area of the face may trigger a spasm of excruciating pain.

[0162] A related but rather uncommon neuralgia affects the glosso-pharyngeal nerve, which provides sensation to the throat. Symptoms of this neuralgia are short, shock-like episodes of pain located in the throat.

[0163] Neuralgia may occur after infections such as shingles, which is caused by the varicella-zoster virus, a type of herpesvirus. This neuralgia produces a constant burning pain after the shingles rash has healed. The pain is worsened by movement of or contact with the affected area. Not all of those diagnosed with shingles go on to experience postherpetic neuralgia, which can be more painful than shingles. The pain and sensitivity can last for months or even years. The pain is usually in the form of an intolerable sensitivity to any touch but especially light touch. Postherpetic neuralgia is not restricted to the face; it can occur anywhere on the body but usually occurs at the location of the shingles rash. Depression is not uncommon due to the pain and social isolation during the illness.

[0164] Postherpetic neuralgia may be debilitating long after signs of the original herpes infection have disappeared. Other infectious diseases that may cause neuralgia are syphilis and Lyme disease.

[0165] Diabetes is another common cause of neuralgia. This very common medical problem affects almost 1 out of every 20 Americans during adulthood. Diabetes damages the tiny arteries that supply circulation to the nerves, resulting in nerve fiber malfunction and sometimes nerve loss. Diabetes can produce almost any neuralgia, including trigeminal neuralgia, carpal tunnel syndrome (pain and numbness of the hand and wrist), and meralgia paresthetica (numbness and pain in the thigh due to damage to the lateral femoral cutaneous nerve). Strict control of blood sugar may prevent diabetic nerve damage and may accelerate recovery in patients who do develop neuralgia.

[0166] Other medical conditions that may be associated with neuralgias are chronic renal insufficiency and porphyria—a hereditary disease in which the body cannot rid itself of certain substances produced after the normal breakdown of blood in the body. Certain drugs may also cause this problem.

2. Deafferentation.

[0167] Deafferentation indicates a loss of the sensory input from a portion of the body, and can be caused by interruption of either peripheral sensory fibres or nerves from the central nervous system. A deafferentation pain syndrome, includes, without limitation, an injury to the brain or spinal cord, a post-stroke pain, a phantom pain, a paraplegia, a brachial plexus avulsion injuries, lumbar radiculopathies.

3. Complex Regional Pain Syndromes (CRPSs)

[0168] CRPS is a chronic pain syndrome resulting from sympathetically-maintained pain, and presents in two forms. CRPS 1 currently replaces the term “reflex sympathetic dystrophy syndrome”. It is a chronic nerve disorder that occurs most often in the arms or legs after a minor or major injury. CRPS 1 is associated with severe pain; changes in the nails, bone, and skin; and an increased sensitivity to touch in the affected limb. CRPS 2 replaces the term causalgia, and results from an identified injury to the nerve. A CRPS, includes, without limitation, a CRPS Type I (reflex sympathetic dystrophy) and a CRPS Type II (causalgia).

4. Neuropathy.

[0169] A neuropathy is a functional or pathological change in a nerve and is characterized clinically by sensory or motor neuron abnormalities.

[0170] Central neuropathy is a functional or pathological change in the central nervous system.

[0171] Peripheral neuropathy is a functional or pathological change in one or more peripheral nerves. The peripheral nerves relay information from your central nervous system (brain and spinal cord) to muscles and other organs and from your skin, joints, and other organs back to your brain. Peripheral neuropathy occurs when these nerves fail to carry information to and from the brain and spinal cord, resulting in pain, loss of sensation, or inability to control muscles. In some cases, the failure of nerves that control blood vessels, intestines, and other organs results in abnormal blood pressure, digestion problems, and loss of other basic body processes. Risk factors for neuropathy include diabetes, heavy alcohol use, and exposure to certain chemicals and drugs. Some people have a hereditary predisposition for neuropathy. Prolonged pressure on a nerve is another risk for developing a nerve injury. Pressure injury may be caused by prolonged immobility (such as a long surgical procedure or lengthy illness) or compression of a nerve by casts, splints, braces, crutches, or other devices. Polyneuropathy implies a widespread process that usually affects both sides of the body equally. The symptoms depend on which type of nerve is affected. The three main types of nerves are sensory, motor, and autonomic. Neuropathy can affect any one or a combination of all three types of nerves. Symptoms also depend on whether the condition affects the whole body or just one nerve (as from an injury). The cause of chronic inflammatory polyneuropathy is an abnormal immune response. The specific antigens, immune processes, and triggering factors are variable and in many cases are unknown. It may occur in association with other conditions such as HIV, inflammatory bowel disease, lupus erythematosus, chronic active hepatitis, and blood cell abnormalities.

[0172] Peripheral neuropathy may involve a function or pathological change to a single nerve or nerve group (mononeuropathy) or a function or pathological change affecting multiple nerves (polyneuropathy).

Peripheral Neuropathies

Hereditary Disorders

- [0173]** Charcot-Marie-Tooth disease
- [0174]** Friedreich's ataxia

Systemic or Metabolic Disorders

- [0175]** Diabetes (diabetic neuropathy)
- [0176]** Dietary deficiencies (especially vitamin B-12)
- [0177]** Excessive alcohol use (alcoholic neuropathy)
- [0178]** Uremia (from kidney failure)
- [0179]** Cancer

Infectious or Inflammatory Conditions

- [0180]** AIDS
- [0181]** Hepatitis
- [0182]** Colorado tick fever
- [0183]** diphtheria
- [0184]** Guillain-Barre syndrome
- [0185]** HIV infection without development of AIDS
- [0186]** leprosy
- [0187]** Lyme
- [0188]** polyarteritis nodosa
- [0189]** rheumatoid arthritis

- [0190] sarcoidosis
- [0191] Sjogren syndrome
- [0192] syphilis
- [0193] systemic lupus erythematosus
- [0194] amyloid

Exposure to Toxic Compounds

- [0195] sniffing glue or other toxic compounds
- [0196] nitrous oxide
- [0197] industrial agents—especially solvents
- [0198] heavy metals (lead, arsenic, mercury, etc.)
- [0199] Neuropathy secondary to drugs like analgesic nephropathy

Miscellaneous Causes

- [0200] ischemia (decreased oxygen/decreased blood flow)
- [0201] prolonged exposure to cold temperature
- [0202] a. Polyneuropathy
- [0203] Polyneuropathy is a peripheral neuropathy involving the loss of movement or sensation to an area caused by damage or destruction to multiple peripheral nerves. Polyneuropathic pain, includes, without limitation, post-polio syndrome, postmastectomy syndrome, diabetic neuropathy, alcohol neuropathy, amyloid, toxins, AIDS, hypothyroidism, uremia, vitamin deficiencies, chemotherapy-induced pain, 2',3'-didexocytidine (ddC) treatment, Guillain-Barré syndrome or Fabry's disease.
- [0204] b. Mononeuropathy
- [0205] Mononeuropathy is a peripheral neuropathy involving loss of movement or sensation to an area caused by damage or destruction to a single peripheral nerve or nerve group. Mononeuropathy is most often caused by damage to a local area resulting from injury or trauma, although occasionally systemic disorders may cause isolated nerve damage (as with mononeuritis multiplex). The usual causes are direct trauma, prolonged pressure on the nerve, and compression of the nerve by swelling or injury to nearby body structures. The damage includes destruction of the myelin sheath (covering) of the nerve or of part of the nerve cell (the axon). This damage slows or prevents conduction of impulses through the nerve. Mononeuropathy may involve any part of the body. Mononeuropathic pain, includes, without limitation, a sciatic nerve dysfunction, a common peroneal nerve dysfunction, a radial nerve dysfunction, an ulnar nerve dysfunction, a cranial mononeuropathy VI, a cranial mononeuropathy VII, a cranial mononeuropathy III (compression type), a cranial mononeuropathy III (diabetic type), an axillary nerve dysfunction, a carpal tunnel syndrome, a femoral nerve dysfunction, a tibial nerve dysfunction, a Bell's palsy, a thoracic outlet syndrome, a carpal tunnel syndrome and a sixth (abducent) nerve palsy
- [0206] c. Generalized Peripheral Neuropathies
- [0207] Generalized peripheral neuropathies are symmetrical, and usually due to various systematic illnesses and disease processes that affect the peripheral nervous system in its entirety. They are further subdivided into several categories:
- [0208] i. Distal axonopathies are the result of some metabolic or toxic derangement of neurons. They may be caused by metabolic diseases such as diabetes, renal failure, deficiency syndromes such as malnutrition and alcoholism, or the effects of toxins or drugs. Distal axonopathy (aka dying back neuropathy) is a type of peripheral neuropathy that results from some metabolic or toxic derangement of peripheral

nervous system (PNS) neurons. It is the most common response of nerves to metabolic or toxic disturbances, and as such may be caused by metabolic diseases such as diabetes, renal failure, deficiency syndromes such as malnutrition and alcoholism, or the effects of toxins or drugs. The most common cause of distal axonopathy is diabetes, and the most common distal axonopathy is diabetic neuropathy.

[0209] ii. Myelinopathies are due to a primary attack on myelin causing an acute failure of impulse conduction. The most common cause is acute inflammatory demyelinating polyneuropathy (AIDP; aka Guillain-Barré syndrome), though other causes include chronic inflammatory demyelinating syndrome (CIDP), genetic metabolic disorders (e.g., leukodystrophy), or toxins. Myelinopathy is due to primary destruction of myelin or the myelinating Schwann cells, which leaves the axon intact, but causes an acute failure of impulse conduction. This demyelination slows down or completely blocks the conduction of electrical impulses through the nerve. The most common cause is acute inflammatory demyelinating polyneuropathy (AIDP, better known as Guillain-Barré syndrome), though other causes include chronic inflammatory demyelinating polyneuropathy (CIDP), genetic metabolic disorders (e.g., leukodystrophy or Charcot-Marie-Tooth disease), or toxins.

[0210] iii. Neuronopathies are the result of destruction of peripheral nervous system (PNS) neurons. They may be caused by motor neurone diseases, sensory neuronopathies (e.g., Herpes zoster), toxins or autonomic dysfunction. Neurotoxins may cause neuronopathies, such as the chemotherapeutic agent vincristine. Neuronopathy is dysfunction due to damage to neurons of the peripheral nervous system (PNS), resulting in a peripheral neuropathy. It may be caused by motor neurone diseases, sensory neuronopathies (e.g., Herpes zoster), toxic substances or autonomic dysfunction. A person with neuronopathy may present in different ways, depending on the cause, the way it affects the nerve cells, and the type of nerve cell that is most affected.

[0211] iv. Focal entrapment neuropathies (e.g., carpal tunnel syndrome).

II. Inflammatory Pain

[0212] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the following inflammatory conditions

A. Arthritic Disorder

[0213] Arthritic disorders include, for example, a rheumatoid arthritis; a juvenile rheumatoid arthritis; a systemic lupus erythematosus (SLE); a gouty arthritis; a scleroderma; an osteoarthritis; a psoriatic arthritis; an ankylosing spondylitis; a Reiter's syndrome (reactive arthritis); an adult Still's disease; an arthritis from a viral infection; an arthritis from a bacterial infection, such as, e.g., a gonococcal arthritis and a non-gonococcal bacterial arthritis (septic arthritis); a Tertiary Lyme disease; a tuberculous arthritis; and an arthritis from a fungal infection, such as, e.g. a blastomycosis

B. Autoimmune Diseases

[0214] Autoimmune diseases include, for example, a Guillain-Barré syndrome, a Hashimoto's thyroiditis, a pernicious anemia, an Addison's disease, a type I diabetes, a systemic lupus erythematosus, a dermatomyositis, a Sjogren's syn-

drome, a lupus erythematosus, a multiple sclerosis, a myasthenia gravis, a Reiter's syndrome and a Grave's disease.

C. Connective Tissue Disorder

[0215] Connective tissue disorders include, for example, a spondyloarthritis a dermatomyositis, and a fibromyalgia.

D. Injury

[0216] Inflammation caused by injury, including, for example, a crush, puncture, stretch of a tissue or joint, may cause chronic inflammatory pain.

E. Infection

[0217] Inflammation caused by infection, including, for example, a tuberculosis or an interstitial keratitis may cause chronic inflammatory pain.

F. Neuritis

[0218] Neuritis is an inflammatory process affecting a nerve or group of nerves. Symptoms depend on the nerves involved, but may include pain, paresthesias, paresis, or hypesthesia (numbness).

[0219] Examples include:

[0220] a. Brachial neuritis

[0221] b. Retrobulbar neuropathy, an inflammatory process affecting the part of the optic nerve lying immediately behind the eyeball.

[0222] c. Optic neuropathy, an inflammatory process affecting the optic nerve causing sudden, reduced vision in the affected eye. The cause of optic neuritis is unknown. The sudden inflammation of the optic nerve (the nerve connecting the eye and the brain) leads to swelling and destruction of the myelin sheath. The inflammation may occasionally be the result of a viral infection, or it may be caused by autoimmune diseases such as multiple sclerosis. Risk factors are related to the possible causes.

[0223] d. Vestibular neuritis, a viral infection causing an inflammatory process affecting the vestibular nerve.

G. Joint Inflammation

[0224] Inflammation of the joint, such as that caused by bursitis or tendonitis, for example, may cause chronic inflammatory pain.

III. Headache Pain

[0225] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the following headache conditions. A headache (medically known as cephalgia) is a condition of mild to severe pain in the head; sometimes neck or upper back pain may also be interpreted as a headache. It may indicate an underlying local or systemic disease or be a disorder in itself.

A. Muscular/Myogenic Headache

[0226] Muscular/myogenic headaches appear to involve the tightening or tensing of facial and neck muscles; they may radiate to the forehead. Tension headache is the most common form of myogenic headache.

[0227] A tension headache is a condition involving pain or discomfort in the head, scalp, or neck, usually associated with muscle tightness in these areas. Tension headaches result

from the contraction of neck and scalp muscles. One cause of this muscle contraction is a response to stress, depression or anxiety. Any activity that causes the head to be held in one position for a long time without moving can cause a headache. Such activities include typing or use of computers, fine work with the hands, and use of a microscope. Sleeping in a cold room or sleeping with the neck in an abnormal position may also trigger this type of headache. A tension-type headache, includes, without limitation, an episodic tension headache and a chronic tension headache.

B. Vascular Headache

[0228] The most common type of vascular headache is migraine. Other kinds of vascular headaches include cluster headaches, which cause repeated episodes of intense pain, and headaches resulting from high blood pressure

[0229] 1. Migraine

[0230] A migraine is a heterogeneous disorder that generally involves recurring headaches. Migraines are different from other headaches because they occur with other symptoms, such as, e.g., nausea, vomiting, or sensitivity to light. In most people, a throbbing pain is felt only on one side of the head. Clinical features such as type of aura symptoms, presence of prodromes, or associated symptoms such as vertigo, may be seen in subgroups of patients with different underlying pathophysiological and genetic mechanisms. A migraine headache, includes, without limitation, a migraine without aura (common migraine), a migraine with aura (classic migraine), a menstrual migraine, a migraine equivalent (acephalic headache), a complicated migraine, an abdominal migraine and a mixed tension migraine.

[0231] 2. Cluster Headache

[0232] Cluster headaches affect one side of the head (unilateral) and may be associated with tearing of the eyes and nasal congestion. They occurs in clusters, happening repeatedly every day at the same time for several weeks and then remitting.

D. High Blood Pressure Headache

E. Traction and Inflammatory Headache

[0233] Traction and inflammatory headaches are usually symptoms of other disorders, ranging from stroke to sinus infection.

F. Hormone Headache

G. Rebound Headache

[0234] Rebound headaches, also known as medication overuse headaches, occur when medication is taken too frequently to relieve headache. Rebound headaches frequently occur daily and can be very painful.

H. Chronic Sinusitis Headache

[0235] Sinusitis is inflammation, either bacterial, fungal, viral, allergic or autoimmune, of the paranasal sinuses. Chronic sinusitis is one of the most common complications of the common cold. Symptoms include: Nasal congestion; facial pain; headache; fever; general malaise; thick green or yellow discharge; feeling of facial 'fullness' worsening on bending over. In a small number of cases, chronic maxillary sinusitis can also be brought on by the spreading of bacteria

from a dental infection. Chronic hyperplastic eosinophilic sinusitis is a noninfective form of chronic sinusitis.

I. An Organic Headache

J. Ictal Headaches

[0236] Ictal headaches are headaches associated with seizure activity.

IV. Somatic Pain

[0237] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the following somatic pain conditions. Somatic pain originates from ligaments, tendons, bones, blood vessels, and even nerves themselves. It is detected with somatic nociceptors. The scarcity of pain receptors in these areas produces a dull, poorly-localized pain of longer duration than cutaneous pain; examples include sprains and broken bones. Additional examples include the following.

A. Excessive Muscle Tension

[0238] Excessive muscle tension can be caused, for example, by a sprain or a strain.

B. Repetitive Motion Disorders

[0239] Repetitive motion disorders can result from overuse of the hands, wrists, elbows, shoulders, neck, back, hips, knees, feet, legs, or ankles.

C. Muscle Disorders

[0240] Muscle disorders causing somatic pain include, for example, a polymyositis, a dermatomyositis, a lupus, a fibromyalgia, a polymyalgia rheumatica, and a rhabdomyolysis.

D. Myalgia

[0241] Myalgia is muscle pain and is a symptom of many diseases and disorders. The most common cause for myalgia is either overuse or over-stretching of a muscle or group of muscles. Myalgia without a traumatic history is often due to viral infections. Longer-term myalgias may be indicative of a metabolic myopathy, some nutritional deficiencies or chronic fatigue syndrome.

E. Infection

[0242] Infection can cause somatic pain. Examples of such infection include, for example, an abscess in the muscle, a trichinosis, an influenza, a Lyme disease, a malaria, a Rocky Mountain spotted fever, Avian influenza, the common cold, community-acquired pneumonia, meningitis, monkeypox, Severe Acute Respiratory Syndrome, toxic shock syndrome, trichinosis, typhoid fever, and upper respiratory tract infection.

F. Drugs

[0243] Drugs can cause somatic pain. Such drugs include, for example, cocaine, a statin for lowering cholesterol (such as atorvastatin, simvastatin, and lovastatin), and an ACE inhibitor for lowering blood pressure (such as enalapril and captopril)

V. Visceral Pain

[0244] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the

following visceral pain conditions. Visceral pain originates from body's viscera, or organs. Visceral nociceptors are located within body organs and internal cavities. The even greater scarcity of nociceptors in these areas produces pain that is usually more aching and of a longer duration than somatic pain. Visceral pain is extremely difficult to localise, and several injuries to visceral tissue exhibit "referred" pain, where the sensation is localised to an area completely unrelated to the site of injury. Examples of visceral pain include the following.

A. Functional Visceral Pain

[0245] Functional visceral pain includes, for example, an irritable bowel syndrome and a chronic functional abdominal pain (CFAP), a functional constipation and a functional dyspepsia, a non-cardiac chest pain (NCCP) and a chronic abdominal pain.

B. Chronic Gastrointestinal Inflammation

[0246] Chronic gastrointestinal inflammation includes, for example, a gastritis, an inflammatory bowel disease, like, e.g., a Crohn's disease, an ulcerative colitis, a microscopic colitis, a diverticulitis and a gastroenteritis; an interstitial cystitis; an intestinal ischemia; a cholecystitis; an appendicitis; a gastroesophageal reflux; an ulcer, a nephrolithiasis, an urinary tract infection, a pancreatitis and a hernia.

C. Autoimmune Pain

[0247] Autoimmune pain includes, for example, a sarcoidosis and a vasculitis.

D. Organic Visceral Pain

[0248] Organic visceral pain includes, for example, pain resulting from a traumatic, inflammatory or degenerative lesion of the gut or produced by a tumor impinging on sensory innervation.

E. Treatment-Induced Visceral Pain

[0249] Treatment-induced visceral pain includes, for example, a pain attendant to chemotherapy therapy or a pain attendant to radiation therapy.

VI. Referred Pain

[0250] The compounds of the invention may be used to treat pain caused by or otherwise associated with any of the following referred pain conditions.

[0251] Referred pain arises from pain localized to an area separate from the site of pain stimulation. Often, referred pain arises when a nerve is compressed or damaged at or near its origin. In this circumstance, the sensation of pain will generally be felt in the territory that the nerve serves, even though the damage originates elsewhere. A common example occurs in intervertebral disc herniation, in which a nerve root arising from the spinal cord is compressed by adjacent disc material. Although pain may arise from the damaged disc itself, pain will also be felt in the region served by the compressed nerve (for example, the thigh, knee, or foot). Relieving the pressure on the nerve root may ameliorate the referred pain, provided that permanent nerve damage has not occurred. Myocardial ischaemia (the loss of blood flow to a part of the heart muscle tissue) is possibly the best known example of referred pain;

the sensation can occur in the upper chest as a restricted feeling, or as an ache in the left shoulder, arm or even hand.

[0252] The present invention addresses a wide range of pain conditions, in particular chronic pain conditions. Preferred conditions include cancerous and non-cancerous pain, inflammatory pain and neuropathic pain. The opioid-fusions of the present application are particularly suited to addressing inflammatory pain, though may be less suited to addressing neuropathic pain. The galanin-fusions are more suited to addressing neuropathic pain.

[0253] In use, the polypeptides of the present invention are typically employed in the form of a pharmaceutical composition in association with a pharmaceutical carrier, diluent and/or excipient, although the exact form of the composition may be tailored to the mode of administration. Administration is preferably to a mammal, more preferably to a human.

[0254] The polypeptides may, for example, be employed in the form of a sterile solution for intra-articular administration or intra-cranial administration. Spinal injection (e.g. epidural or intrathecal) is preferred.

[0255] The dosage ranges for administration of the polypeptides of the present invention are those to produce the desired therapeutic effect. It will be appreciated that the dosage range required depends on the precise nature of the components, the route of administration, the nature of the formulation, the age of the patient, the nature, extent or severity of the patient's condition, contraindications, if any, and the judgement of the attending physician.

[0256] Suitable daily dosages are in the range 0.0001-1 mg/kg, preferably 0.0001-0.5 mg/kg, more preferably 0.002-0.5 mg/kg, and particularly preferably 0.004-0.5 mg/kg. The unit dosage can vary from less than 1 microgram to 30 mg, but typically will be in the region of 0.01 to 1 mg per dose, which may be administered daily or preferably less frequently, such as weekly or six monthly.

[0257] A particularly preferred dosing regimen is based on 2.5 ng of fusion protein (e.g. CPNv/A) as the $1\times$ dose. In this regard, preferred dosages are in the range $1\times$ - $100\times$ (i.e. 2.5-250 ng). This dosage range is significantly lower (i.e. at least 10-fold, typically 100-fold lower) than would be employed with other types of analgesic molecules such as NSAIDs, morphine, and gabapentin. Moreover, the above-mentioned difference is considerably magnified when the same comparison is made on a molar basis—this is because the fusion proteins of the present invention have a considerably greater Mw than do conventional 'small' molecule therapeutics.

[0258] Wide variations in the required dosage, however, are to be expected depending on the precise nature of the components, and the differing efficiencies of various routes of administration.

[0259] Variations in these dosage levels can be adjusted using standard empirical routines for optimisation, as is well understood in the art.

[0260] Compositions suitable for injection may be in the form of solutions, suspensions or emulsions, or dry powders which are dissolved or suspended in a suitable vehicle prior to use.

[0261] Fluid unit dosage forms are typically prepared utilising a pyrogen-free sterile vehicle. The active ingredients, depending on the vehicle and concentration used, can be either dissolved or suspended in the vehicle.

[0262] In preparing administrable solutions, the polypeptides can be dissolved in a vehicle, the solution being made isotonic if necessary by addition of sodium chloride and

sterilised by filtration through a sterile filter using aseptic techniques before filling into suitable sterile vials or ampoules and sealing. Alternatively, if solution stability is adequate, the solution in its sealed containers may be sterilised by autoclaving.

[0263] Advantageously additives such as buffering, solubilising, stabilising, preservative or bactericidal, suspending or emulsifying agents may be dissolved in the vehicle.

[0264] Dry powders which are dissolved or suspended in a suitable vehicle prior to use may be prepared by filling pre-sterilised drug substance and other ingredients into a sterile container using aseptic technique in a sterile area.

[0265] Alternatively the polypeptides and other ingredients may be dissolved in an aqueous vehicle, the solution is sterilized by filtration and distributed into suitable containers using aseptic technique in a sterile area. The product is then freeze dried and the containers are sealed aseptically.

[0266] Parenteral suspensions, suitable for intramuscular, subcutaneous or intradermal injection, are prepared in substantially the same manner, except that the sterile components are suspended in the sterile vehicle, instead of being dissolved and sterilisation cannot be accomplished by filtration. The components may be isolated in a sterile state or alternatively it may be sterilised after isolation, e.g. by gamma irradiation.

[0267] Advantageously, a suspending agent for example polyvinylpyrrolidone is included in the composition/s to facilitate uniform distribution of the components.

Definitions Section

[0268] Targeting Moiety (TM) means any chemical structure associated with an agent that functionally interacts with a Binding Site to cause a physical association between the agent and the surface of a target cell. In the context of the present invention, the target cell is a nociceptive sensory afferent. The term TM embraces any molecule (i.e. a naturally occurring molecule, or a chemically/physically modified variant thereof) that is capable of binding to a Binding Site on the target cell, which Binding Site is capable of internalisation (e.g. endosome formation)—also referred to as receptor-mediated endocytosis. The TM may possess an endosomal membrane translocation function, in which case separate TM and Translocation Domain components need not be present in an agent of the present invention.

[0269] The TM of the present invention binds (preferably specifically binds) to a nociceptive sensory afferent (e.g. a primary nociceptive afferent). In this regard, specifically binds means that the TM binds to a nociceptive sensory afferent (e.g. a primary nociceptive afferent) with a greater affinity than it binds to other neurons such as non-nociceptive afferents, and/or to motor neurons (i.e. the natural target for clostridial neurotoxin holotoxin). The term "specifically binding" can also mean that a given TM binds to a given receptor, for example the ORL₁ receptor, with a binding affinity (K_a) of 10^6 M^{-1} or greater, preferably 10^7 M^{-1} or greater, more preferably 10^8 M^{-1} or greater, and most preferably, 10^9 M^{-1} or greater.

[0270] For the purposes of this invention, an agonist is defined as a molecule that is capable of stimulating the process of exocytic fusion in a target cell, which process is susceptible to inhibition by a protease (or fragment thereof) capable of cleaving a protein of the exocytic fusion apparatus in said target cell.

[0271] Accordingly, the particular agonist definition of the present invention would exclude many molecules that would be conventionally considered as agonists.

[0272] For example, nerve growth factor (NGF) is an agonist in respect of its ability to promote neuronal differentiation via binding to a TrkA receptor. However, NGF is not an agonist when assessed by the above criteria because it is not a principal inducer of exocytic fusion. In addition, the process that NGF stimulates (i.e. cell differentiation) is not susceptible to inhibition by the protease activity of a non-cytotoxic toxin molecule.

[0273] The term “fragment”, when used in relation to a protein, means a peptide having at least thirty-five, preferably at least twenty-five, more preferably at least twenty, and most preferably at least ten amino acid residues of the protein in question.

[0274] The term “variant”, when used in relation to a protein, means a peptide or peptide fragment of the protein that contains one or more analogues of an amino acid (e.g. an unnatural amino acid), or a substituted linkage.

[0275] The term “derivative”, when used in relation to a protein, means a protein that comprises the protein in question, and a further peptide sequence. The further peptide sequence should preferably not interfere with the basic folding and thus conformational structure of the original protein. Two or more peptides (or fragments, or variants) may be joined together to form a derivative. Alternatively, a peptide (or fragment, or variant) may be joined to an unrelated molecule (e.g. a second, unrelated peptide). Derivatives may be chemically synthesized, but will be typically prepared by recombinant nucleic acid methods. Additional components such as lipid, and/or polysaccharide, and/or polyketide components may be included.

[0276] Throughout this specification, reference to the “ORL₁ receptor” embraces all members of the ORL₁ receptor family. Members of the ORL₁ receptor family typically have a seven transmembrane domain structure and are coupled to G-proteins of the G_i and G_o families. A method for determining the G-protein-stimulating activity of ligands of the ORL₁ receptor is given in Example 12. A method for measuring reduction in cellular cAMP levels following ORL₁ activation is given in Example 11. A further characteristic of members of the ORL₁ receptor family is that they are typically able to bind nociceptin (the natural ligand of ORL₁). As an example, all alternative splice variants of the ORL₁ receptor, are members of the ORL₁ receptor family.

[0277] The term non-cytotoxic means that the protease molecule in question does not kill the target cell to which it has been re-targeted.

[0278] The protease of the present invention embraces all naturally-occurring non-cytotoxic proteases that are capable of cleaving one or more proteins of the exocytic fusion apparatus in eukaryotic cells.

[0279] The protease of the present invention is preferably a bacterial protease (or fragment thereof). More preferably the bacterial protease is selected from the genera *Clostridium* or *Neisseria* (e.g. a clostridial L-chain, or a neisserial IgA protease preferably from *N. gonorrhoeae*).

[0280] The present invention also embraces modified non-cytotoxic proteases, which include amino acid sequences that do not occur in nature and/or synthetic amino acid residues, so long as the modified proteases still demonstrate the above-mentioned protease activity.

[0281] The protease of the present invention preferably demonstrates a serine or metalloprotease activity (e.g. endopeptidase activity). The protease is preferably specific for a SNARE protein (e.g. SNAP-25, synaptobrevin/VAMP, or syntaxin).

[0282] Particular mention is made to the protease domains of neurotoxins, for example the protease domains of bacterial neurotoxins. Thus, the present invention embraces the use of neurotoxin domains, which occur in nature, as well as recombinantly prepared versions of said naturally-occurring neurotoxins.

[0283] Exemplary neurotoxins are produced by clostridia, and the term clostridial neurotoxin embraces neurotoxins produced by *C. tetani* (TeNT), and by *C. botulinum* (BoNT) serotypes A-G, as well as the closely related BoNT-like neurotoxins produced by *C. baratii* and *C. butyricum*. The above-mentioned abbreviations are used throughout the present specification. For example, the nomenclature BoNT/A denotes the source of neurotoxin as BoNT (serotype A). Corresponding nomenclature applies to other BoNT serotypes.

[0284] The term L-chain fragment means a component of the L-chain of a neurotoxin, which fragment demonstrates a metalloprotease activity and is capable of proteolytically cleaving a vesicle and/or plasma membrane associated protein involved in cellular exocytosis.

[0285] A Translocation Domain is a molecule that enables translocation of a protease (or fragment thereof) into a target cell such that a functional expression of protease activity occurs within the cytosol of the target cell. Whether any molecule (e.g. a protein or peptide) possesses the requisite translocation function of the present invention may be confirmed by any one of a number of conventional assays.

[0286] For example, Shone C. (1987) describes an in vitro assay employing liposomes, which are challenged with a test molecule. Presence of the requisite translocation function is confirmed by release from the liposomes of K⁺ and/or labelled NAD, which may be readily monitored [see Shone C. (1987) Eur. J. Biochem; vol. 167(1): pp. 175-180].

[0287] A further example is provided by Blaustein R. (1987), which describes a simple in vitro assay employing planar phospholipid bilayer membranes. The membranes are challenged with a test molecule and the requisite translocation function is confirmed by an increase in conductance across said membranes [see Blaustein (1987) FEBS Letts; vol. 226, no. 1: pp. 115-120].

[0288] Additional methodology to enable assessment of membrane fusion and thus identification of Translocation Domains suitable for use in the present invention are provided by Methods in Enzymology Vol 220 and 221, Membrane Fusion Techniques, Parts A and B, Academic Press 1993.

[0289] The Translocation Domain is preferably capable of formation of ion-permeable pores in lipid membranes under conditions of low pH. Preferably it has been found to use only those portions of the protein molecule capable of pore-formation within the endosomal membrane.

[0290] The Translocation Domain may be obtained from a microbial protein source, in particular from a bacterial or viral protein source. Hence, in one embodiment, the Translocation Domain is a translocating domain of an enzyme, such as a bacterial toxin or viral protein.

[0291] It is well documented that certain domains of bacterial toxin molecules are capable of forming such pores. It is also known that certain translocation domains of virally

expressed membrane fusion proteins are capable of forming such pores. Such domains may be employed in the present invention.

[0292] The Translocation Domain may be of a clostridial origin, namely the H_N domain (or a functional component thereof). H_N means a portion or fragment of the H-chain of a clostridial neurotoxin approximately equivalent to the amino-terminal half of the H-chain, or the domain corresponding to that fragment in the intact H-chain. It is preferred that the H-chain substantially lacks the natural binding function of the H_C component of the H-chain. In this regard, the H_C function may be removed by deletion of the H_C amino acid sequence (either at the DNA synthesis level, or at the post-synthesis level by nuclease or protease treatment). Alternatively, the H_C function may be inactivated by chemical or biological treatment. Thus, the H-chain is preferably incapable of binding to the Binding Site on a target cell to which native clostridial neurotoxin (i.e. holotoxin) binds.

[0293] In one embodiment, the translocation domain is a H_N domain (or a fragment thereof) of a clostridial neurotoxin. Examples of suitable clostridial Translocation

[0294] Domains include:

[0295] Botulinum type A neurotoxin—amino acid residues (449-871)

[0296] Botulinum type B neurotoxin—amino acid residues (441-858)

[0297] Botulinum type C neurotoxin—amino acid residues (442-866)

[0298] Botulinum type D neurotoxin—amino acid residues (446-862)

[0299] Botulinum type E neurotoxin—amino acid residues (423-845)

[0300] Botulinum type F neurotoxin—amino acid residues (440-864)

[0301] Botulinum type G neurotoxin—amino acid residues (442-863)

[0302] Tetanus neurotoxin—amino acid residues (458-879)

[0303] For further details on the genetic basis of toxin production in *Clostridium botulinum* and *C. tetani*, we refer to Henderson et al (1997) in *The Clostridia: Molecular Biology and Pathogenesis*, Academic press.

[0304] The term H_N embraces naturally-occurring neurotoxin H_N portions, and modified H_N portions having amino acid sequences that do not occur in nature and/or synthetic amino acid residues, so long as the modified H_N portions still demonstrate the above-mentioned translocation function.

[0305] Alternatively, the Translocation Domain may be of a non-clostridial origin (see Table 4). Examples of non-clostridial Translocation Domain origins include, but not be restricted to, the translocation domain of diphtheria toxin [O=Keefe et al., Proc. Natl. Acad. Sci. USA (1992) 89, 6202-6206; Silverman et al., J. Biol. Chem. (1993) 269, 22524-22532; and London, E. (1992) *Biochem. Biophys. Acta.*, 1112, pp. 25-51], the translocation domain of *Pseudomonas* exotoxin type A [Prior et al. Biochemistry (1992) 31, 3555-3559], the translocation domains of anthrax toxin [Blanke et al. Proc. Natl. Acad. Sci. USA (1996) 93, 8437-8442], a variety of fusogenic or hydrophobic peptides of translocating function [Plank et al. J. Biol. Chem. (1994) 269, 12918-12924; and Wagner et al (1992) *PNAS*, 89, pp. 7934-7938], and amphiphilic peptides [Murata et al (1992) *Biochem.*, 31, pp. 1986-1992]. The Translocation Domain may mirror the Translocation Domain present in a naturally-occurring protein, or may include amino acid variations so long as the variations do not destroy the translocating ability of the Translocation Domain.

[0306] Particular examples of viral Translocation Domains suitable for use in the present invention include certain translocating domains of virally expressed membrane fusion proteins. For example, Wagner et al. (1992) and Murata et al. (1992) describe the translocation (i.e. membrane fusion and vesiculation) function of a number of fusogenic and amphiphilic peptides derived from the N-terminal region of influenza virus haemagglutinin. Other virally expressed membrane fusion proteins known to have the desired translocating activity are a translocating domain of a fusogenic peptide of Semliki Forest Virus (SFV), a translocating domain of vesicular stomatitis virus (VSV) glycoprotein G, a translocating domain of SER virus F protein and a translocating domain of Foamy virus envelope glycoprotein. Virally encoded Aspike proteins have particular application in the context of the present invention, for example, the E1 protein of SFV and the G protein of the G protein of VSV.

[0307] Use of the Translocation Domains listed in Table (below) includes use of sequence variants thereof. A variant may comprise one or more conservative nucleic acid substitutions and/or nucleic acid deletions or insertions, with the proviso that the variant possesses the requisite translocating function. A variant may also comprise one or more amino acid substitutions and/or amino acid deletions or insertions, so long as the variant possesses the requisite translocating function.

Translocation domain source	Amino acid residues	References
Diphtheria toxin	194-380	Silverman et al., 1994, J. Biol. Chem. 269, 22524-22532 London E., 1992, Biochem. Biophys. Acta., 1113, 25-51
Domain II of <i>pseudomonas</i> exotoxin	405-613	Prior et al., 1992, Biochemistry 31, 3555-3559 Kihara & Pastan, 1994, Bioconj Chem. 5, 532-538
Influenza virus haemagglutinin	GLFGAIAGFIENGWE GMIDGWYG, and Variants thereof	Plank et al., 1994, J. Biol. Chem. 269, 12918-12924 Wagner et al., 1992, PNAS, 89, 7934-7938 Murata et al., 1992, Biochemistry 31, 1986-1992

-continued

Translocation domain source	Amino acid residues	References
Semliki Forest virus fusogenic protein	Translocation domain	Kielian et al., 1996, J Cell Biol. 134(4), 863-872
Vesicular Stomatitis virus glycoprotein G	118-139	Yao et al., 2003, Virology 310(2), 319-332
SER virus F protein	Translocation domain	Seth et al., 2003, J Virol 77(11) 6520-6527
Foamy virus envelope glycoprotein	Translocation domain	Picard-Maureau et al., 2003, J Virol. 77(8), 4722-4730

SEQ ID NOs

[0308] Where an initial Met amino acid residue or a corresponding initial codon is indicated in any of the following SEQ ID NOs, said residue/codon is optional.

SEQ ID1 DNA sequence of the LC/A
 SEQ ID2 DNA sequence of the H_N/A
 SEQ ID3 DNA sequence of the LC/B
 SEQ ID4 DNA sequence of the H_N/B
 SEQ ID5 DNA sequence of the LC/C
 SEQ ID6 DNA sequence of the H_N/C
 SEQ ID7 DNA sequence of the CPN-A linker
 SEQ ID8 DNA sequence of the A linker
 SEQ ID9 DNA sequence of the N-terminal presentation nociceptin insert
 SEQ ID10 DNA sequence of the CPN-C linker
 SEQ ID11 DNA sequence of the CPBE-A linker
 SEQ ID12 DNA sequence of the CPNvar-A linker
 SEQ ID13 DNA sequence of the LC/A-CPN-H_N/A fusion
 SEQ ID14 Protein sequence of the LC/A-CPN-H_N/A fusion
 SEQ ID15 DNA sequence of the N-LC/A-H_N/A fusion
 SEQ ID16 Protein sequence of the N-LC/A-H_N/A fusion
 SEQ ID17 DNA sequence of the LC/C-CPN-H_N/C fusion
 SEQ ID18 Protein sequence of the LC/C-CPN-H_N/C fusion
 SEQ ID19 DNA sequence of the LC/C-CPN-H_N/C (A-linker) fusion
 SEQ ID20 Protein sequence of the LC/C-CPN-H_N/C (A-linker) fusion
 SEQ ID21 DNA sequence of the LC/A-CPME-H_N/A fusion
 SEQ ID22 Protein sequence of the LC/A-CPME-H_N/A fusion
 SEQ ID23 DNA sequence of the LC/A-CPBE-H_N/A fusion
 SEQ ID24 Protein sequence of the LC/A-CPBE-H_N/A fusion
 SEQ ID25 DNA sequence of the LC/A-CPNv-H_N/A fusion
 SEQ ID26 Protein sequence of the LC/A-CPNv-H_N/A fusion
 SEQ ID27 DNA sequence of the LC/A-CPN[1-11]-HN/A fusion
 SEQ ID28 Protein sequence of the LC/A-CPN[1-11]-HN/A fusion
 SEQ ID29 DNA sequence of the LC/A-CPN[[Y10]1-11]-HN/A fusion
 SEQ ID30 Protein sequence of the LC/A-CPN[[Y10]1-11]-HN/A fusion
 SEQ ID31 DNA sequence of the LC/A-CPN[[Y11]1-11]-HN/A fusion
 SEQ ID32 Protein sequence of the LC/A-CPN[[Y11]1-11]-HN/A fusion

SEQ ID33 DNA sequence of the LC/A-CPN[[Y14]1-17]-HN/A fusion
 SEQ ID34 Protein sequence of the LC/A-CPN[[Y14]1-17]-HN/A fusion
 SEQ ID35 DNA sequence of the LC/A-CPN[1-13]-HN/A fusion
 SEQ ID36 Protein sequence of the LC/A-CPN[1-13]-HN/A fusion
 SEQ ID37 DNA sequence of CPN[1-17]

SEQ ID38 Protein Sequence of CPN[1-17]

[0309] SEQ ID39 DNA sequence of CPN[1-11]
 SEQ ID40 Protein sequence of CPN[1-11]
 SEQ ID41 DNA sequence of CPN[[Y10]1-11]
 SEQ ID42 Protein sequence of CPN[[Y10]1-11]
 SEQ ID43 DNA sequence of CPN[[Y11]1-11]
 SEQ ID44 Protein sequence of CPN[[Y11]1-11]
 SEQ ID45 DNA sequence of CPN[[Y14]1-17]
 SEQ ID46 Protein sequence of CPN[[Y14]1-17]
 SEQ ID47 DNA sequence of CPN[1-13]
 SEQ ID48 Protein sequence of CPN[1-13]
 SEQ ID49 DNA sequence of CPNv (also known as N[[R14K15]1-17])
 SEQ ID50 Protein sequence of CPNv (also known as N[[R14K15]1-17])
 SEQ ID51 DNA sequence of the nociceptin-spacer-LC/A-H_N/A fusion
 SEQ ID52 Protein sequence of the nociceptin-spacer-LC/A-H_N/A fusion
 SEQ ID53 DNA sequence of the CPN-A GS10 linker
 SEQ ID54 DNA sequence of the CPN-A GS15 linker
 SEQ ID55 DNA sequence of the CPN-A GS25 linker
 SEQ ID56 DNA sequence of the CPN-A GS30 linker
 SEQ ID57 DNA sequence of the CPN-A HX27 linker
 SEQ ID58 DNA sequence of the LC/A-CPN(GS15)-H_N/A fusion
 SEQ ID59 Protein sequence of the LC/A-CPN(GS15)-H_N/A fusion
 SEQ ID60 DNA sequence of the LC/A-CPN(GS25)-H_N/A fusion
 SEQ ID61 Protein sequence of the LC/A-CPN(GS25)-H_N/A fusion
 SEQ ID62 DNA sequence of the CPNvar-A Enterokinase activatable linker
 SEQ ID63 DNA sequence of the LC/A-CPNv(Ek)-H_N/A fusion

SEQ ID64 Protein sequence of the LC/A-CPNV(Ek)-H_N/A fusion
 SEQ ID65 DNA sequence of the CPNvar-A linker
 SEQ ID66 DNA sequence of the LC/C-CPNV-H_N/C fusion (act. A)
 SEQ ID67 Protein sequence of the LC/C-CPNV-H_N/C fusion (act. A)
 SEQ ID68 DNA sequence of the LC/A-CPLE-H_N/A fusion
 SEQ ID69 Protein sequence of the LC/A-CPLE-H_N/A fusion
 SEQ ID70 DNA sequence of the LC/A-CPOP-H_N/A fusion
 SEQ ID71 Protein sequence of the LC/A-CPOP-H_N/A fusion
 SEQ ID72 DNA sequence of the LC/A-CPOPv-H_N/A fusion
 SEQ ID73 Protein sequence of the LC/A-CPOPv-H_N/A fusion
 SEQ ID74 DNA sequence of the IgA protease
 SEQ ID75 DNA sequence of the IgA-CPNV-H_N/A fusion
 SEQ ID76 Protein sequence of the IgA-CPNV-H_N/A fusion
 SEQ ID77 DNA sequence of the FXa-HT
 SEQ ID78 DNA sequence of the CPNV-A-FXa-HT
 SEQ ID79 Protein sequence of the CPNV-A-FXa-HT fusion
 SEQ ID80 DNA sequence of the DT translocation domain
 SEQ ID81 DNA sequence of the CPLE-DT-A
 SEQ ID82 Protein sequence of the CPLE-DT-A fusion
 SEQ ID83 DNA sequence of the TeNT LC
 SEQ ID84 DNA sequence of the CPNV-TeNT LC
 SEQ ID85 Protein sequence of the CPNV-TeNT LC fusion
 SEQ ID86 DNA sequence of the CPNvar-C linker
 SEQ ID87 DNA sequence of the LC/C-CPNV-H_N/C fusion (act. C)
 SEQ ID88 Protein sequence of the LC/C-CPNV-H_N/C fusion (act. C)
 SEQ ID89 Protein sequence of dynorphin
 SEQ ID90 DNA sequence of LC/A-CPDY-HN/A fusion
 SEQ ID91 Protein sequence of LC/A-CPDY-HN/A fusion
 SEQ ID92 Protein sequence of LC/A-CPDY(GS10)-H_N/A fusion
 SEQ ID93 Protein sequence of LC/A-CPDY(GS15)-H_N/A fusion
 SEQ ID94 Protein sequence of LC/A-CPDY(GS25)-H_N/A fusion
 SEQ ID95 Protein sequence of LC/C-CPDY-HN/C fusion
 SEQ ID96 Protein sequence of IgA-CPDY-HN/A fusion
 SEQ ID97 Protein sequence of CPDY-TeNT LC fusion

EXAMPLES

Example 1

Preparation of a LC/A and H_N/A Backbone Clones

[0310] The following procedure creates the LC and H_N fragments for use as the component backbone for multidomain fusion expression. This example is based on preparation of a serotype A based clone (SEQ ID1 and SEQ ID2), though the procedures and methods are equally applicable to the other serotypes [illustrated by the sequence listing for serotype B (SEQ ID3 and SEQ ID4) and serotype C (SEQ ID5 and SEQ ID6)].

Preparation of Cloning and Expression Vectors

[0311] pCR 4 (Invitrogen) is the chosen standard cloning vector, selected due to the lack of restriction sequences within the vector and adjacent sequencing primer sites for easy construct confirmation. The expression vector is based on the pMAL (NEB) expression vector, which has the desired

restriction sequences within the multiple cloning site in the correct orientation for construct insertion (BamHI-Sall-PstI-HindIII). A fragment of the expression vector has been removed to create a non-mobilisable plasmid and a variety of different fusion tags have been inserted to increase purification options.

Preparation of Protease (e.g. LC/A) Insert

[0312] The LC/A (SEQ ID1) is created by one of two ways:

[0313] The DNA sequence is designed by back translation of the LC/A amino acid sequence [obtained from freely available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO)] using one of a variety of reverse translation software tools (for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)). BamHI/Sall recognition sequences are incorporated at the 5' and 3' ends respectively of the sequence, maintaining the correct reading frame. The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence containing the LC/A open reading frame (ORF) is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0314] The alternative method is to use PCR amplification from an existing DNA sequence with BamHI and Sall restriction enzyme sequences incorporated into the 5' and 3' PCR primers respectively. Complementary oligonucleotide primers are chemically synthesised by a supplier (for example MWG or Sigma-Genosys), so that each pair has the ability to hybridize to the opposite strands (3' ends pointing "towards" each other) flanking the stretch of *Clostridium* target DNA, one oligonucleotide for each of the two DNA strands. To generate a PCR product the pair of short oligonucleotide primers specific for the *Clostridium* DNA sequence are mixed with the *Clostridium* DNA template and other reaction components and placed in a machine (the 'PCR machine') that can change the incubation temperature of the reaction tube automatically, cycling between approximately 94° C. (for denaturation), 55° C. (for oligonucleotide annealing), and 72° C. (for synthesis). Other reagents required for amplification of a PCR product include a DNA polymerase (such as Taq or Pfu polymerase), each of the four nucleotide dNTP building blocks of DNA in equimolar amounts (50-200 μM) and a buffer appropriate for the enzyme optimised for Mg²⁺ concentration (0.5-5 mM).

[0315] The amplification product is cloned into pCR 4 using either, TOPO TA cloning for Taq PCR products or Zero Blunt TOPO cloning for Pfu PCR products (both kits commercially available from Invitrogen). The resultant clone is checked by sequencing. Any additional restriction sequences which are not compatible with the cloning system are then removed using site directed mutagenesis [for example, using Quickchange (Stratagene Inc.)].

Preparation of Translocation (e.g. H_N) Insert

[0316] The H_N/A (SEQ ID2) is created by one of two ways:

[0317] The DNA sequence is designed by back translation of the H_N/A amino acid sequence [obtained from freely available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO)] using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)]. A PstI restriction sequence added to the N-terminus and XbaI-stop codon-HindIII to the C-terminus ensuring the correct reading frame is maintained. The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0318] The alternative method is to use PCR amplification from an existing DNA sequence with PstI and XbaI-stop codon-HindIII restriction enzyme sequences incorporated into the 5' and 3' PCR primers respectively. The PCR amplification is performed as described above. The PCR product is inserted into pCR 4 vector and checked by sequencing. Any additional restriction sequences which are not compatible with the cloning system are then removed using site directed mutagenesis [for example using Quickchange (Stratagene Inc.)].

Example 2

Preparation of a LC/A-nociceptin- H_N/A Fusion Protein (Nociceptin is N-Terminal of the H_N -Chain)

Preparation of Linker-nociceptin-Spacer Insert

[0319] The LC- H_N linker can be designed from first principle, using the existing sequence information for the linker as the template. For example, the serotype A linker (in this case defined as the inter-domain polypeptide region that exists between the cysteines of the disulphide bridge between LC and H_N) is 23 amino acids long and has the sequence VRGIITSKTKSLDKGYNKALNDL. Within this sequence, it is understood that proteolytic activation in nature leads to an H_N domain that has an N-terminus of the sequence ALNDL. This sequence information is freely available from available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO). Into this linker a Factor Xa site, nociceptin and spacer are incorporated; and using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)], the DNA sequence encoding the linker-ligand-spacer region is determined. Restriction sites are then incorporated into the DNA sequence and can be arranged as BamHI-Sall-linker-protease site-nociceptin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID7). It is important to ensure the correct reading frame is maintained for the spacer, nociceptin and restriction sequences and that the XbaI

sequence is not preceded by the bases, TC, which would result on DAM methylation. The DNA sequence is screened for restriction sequence incorporation, and any additional sequences are removed manually from the remaining sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example, GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

Preparation of the LC/A-nociceptin- H_N/A Fusion

[0320] In order to create the LC-linker-nociceptin-spacer- H_N construct (SEQ ID13), the pCR 4 vector encoding the linker (SEQ ID7) is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient vector for insertion and ligation of the LC/A DNA (SEQ ID1) cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with PstI+XbaI restriction enzymes and serves as the recipient vector for the insertion and ligation of the H_N/A DNA (SEQ ID2) cleaved with PstI+XbaI. The final construct contains the LC-linker-nociceptin-spacer- H_N ORF (SEQ ID13) for transfer into expression vectors for expression to result in a fusion protein of the sequence illustrated in SEQ ID14.

Example 3

Preparation of a Nociceptin-LC/A- H_N/A Fusion Protein (Nociceptin is N-Terminal of the LC-Chain)

[0321] The LC/A- H_N/A backbone is constructed as described in Example 2 using the synthesised A serotype linker with the addition of a Factor Xa site for activation, arranged as BamHI-Sall-linker-protease site-linker-PstI-XbaI-stop codon-HindIII (SEQ ID8). The LC/A- H_N/A backbone and the synthesised N-terminal presentation nociceptin insert (SEQ ID9) are cleaved with BamHI+HindIII restriction enzymes, gel purified and ligated together to create a nociceptin-spacer-LC-linker- H_N . The ORF (SEQ ID15) is then cut out using restriction enzymes Aval+XbaI for transfer into expression vectors for expression to result in a fusion protein of the sequence illustrated in SEQ ID16.

Example 4

Preparation of a LC/C-nociceptin- H_N/C Fusion Protein

[0322] Following the methods used in Examples 1 and 2, the LC/C (SEQ ID5) and H_N/C (SEQ ID6) are created and inserted into the C serotype linker arranged as BamHI-Sall-linker-protease site-nociceptin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID10). The final construct contains the LC-linker-nociceptin-spacer- H_N ORF (SEQ ID17) for expression as a protein of the sequence illustrated in SEQ ID18.

Example 5

Preparation of a LC/C-nociceptin- H_N/C Fusion Protein with a Serotype A Activation Sequence

[0323] Following the methods used in Examples 1 and 2, the LC/C (SEQ ID5) and H_N/C (SEQ ID6) are created and

inserted into the A serotype linker arranged as BamHI-Sall-linker-protease site-nociceptin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID7). The final construct contains the LC-linker-nociceptin-spacer- H_N ORF (SEQ ID19) for expression as a protein of the sequence illustrated in SEQ ID20.

Example 6

Preparation of a LC/A-met Enkephalin- H_N /A Fusion Protein

[0324] Due to the small, five-amino acid, size of the met-enkephalin ligand the LC/A-met enkephalin- H_N /A fusion is created by site directed mutagenesis [for example using Quickchange (Stratagene Inc.)] using the LC/A-nociceptin- H_N /A fusion (SEQ ID13) as a template. Oligonucleotides are designed encoding the YGGFM met-enkephalin peptide, ensuring standard *E. coli* codon usage is maintained and no additional restriction sites are incorporated, flanked by sequences complimentary to the linker region of the LC/A-nociceptin- H_N /A fusion (SEQ ID13) either side on the nociceptin section. The SDM product is checked by sequencing and the final construct containing the LC-linker-met enkephalin-spacer- H_N ORF (SEQ ID21) for expression as a protein of the sequence illustrated in SEQ ID22.

Example 7

Preparation of a LC/A- β Endorphin- H_N /A Fusion Protein

[0325] Following the methods used in Examples 1 and 2, the LC/A (SEQ ID1) and H_N /A (SEQ ID2) are created and inserted into the A serotype β endorphin linker arranged as BamHI-Sall-linker-protease site- β endorphin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID11). The final construct contains the LC-linker- β endorphin-spacer- H_N ORF (SEQ ID23) for expression as a protein of the sequence illustrated in SEQ ID24.

Example 8

Preparation of a LC/A-nociceptin Variant- H_N /A Fusion Protein

[0326] Following the methods used in Examples 1 and 2, the LC/A (SEQ ID1) and H_N /A (SEQ ID2) are created and inserted into the A serotype nociceptin variant linker arranged as BamHI-Sall-linker-protease site-nociceptin variant-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID12). The final construct contains the LC-linker-nociceptin variant-spacer- H_N ORF (SEQ ID25) for expression as a protein of the sequence illustrated in SEQ ID26.

Example 9

Purification Method for LC/A-nociceptin- H_N /A Fusion Protein

[0327] Defrost falcon tube containing 25 ml 50 mM HEPES pH 7.2, 200 mM NaCl and approximately 10 g of *E. coli* BL21 cell paste. Make the thawed cell paste up to 80 ml with 50 mM HEPES pH 7.2, 200 mM NaCl and sonicate on ice 30 seconds on, 30 seconds off for 10 cycles at a power of 22 microns ensuring the sample remains cool. Spin the lysed cells at 18 000 rpm, 4° C. for 30 minutes. Load the supernatant onto a 0.1 M NiSO_4 charged Chelating column (20-30 ml column is sufficient) equilibrated with 50 mM HEPES pH

7.2, 200 mM NaCl. Using a step gradient of 10 and 40 mM imidazol, wash away the non-specific bound protein and elute the fusion protein with 100 mM imidazol. Dialyse the eluted fusion protein against 5 L of 50 mM HEPES pH 7.2, 200 mM NaCl at 4° C. overnight and measure the OD of the dialysed fusion protein. Add 1 unit of factor Xa per 100 μg fusion protein and incubate at 25° C. static overnight. Load onto a 0.1 M NiSO_4 charged Chelating column (20-30 ml column is sufficient) equilibrated with 50 mM HEPES pH 7.2, 200 mM NaCl. Wash column to baseline with 50 mM HEPES pH 7.2, 200 mM NaCl. Using a step gradient of 10 and 40 mM imidazol, wash away the non-specific bound protein and elute the fusion protein with 100 mM imidazol. Dialyse the eluted fusion protein against 5 L of 50 mM HEPES pH 7.2, 200 mM NaCl at 4° C. overnight and concentrate the fusion to about 2 mg/ml, aliquot sample and freeze at -20° C. Test purified protein using OD, BCA, purity analysis and SNAP-25 assessments.

Example 10

Confirmation of TM Agonist Activity by Measuring Release of Substance P from Neuronal Cell Cultures

Materials

[0328] Substance P EIA is obtained from R&D Systems, UK.

Methods

[0329] Primary neuronal cultures of eDRG are established as described previously (Duggan et al., 2002). Substance P release from the cultures is assessed by EIA, essentially as described previously (Duggan et al., 2002). The TM of interest is added to the neuronal cultures (established for at least 2 weeks prior to treatment); control cultures are performed in parallel by addition of vehicle in place of TM. Stimulated (100 mM KCl) and basal release, together with total cell lysate content, of substance P are obtained for both control and TM treated cultures. Substance P immunoreactivity is measured using Substance P Enzyme Immunoassay Kits (Cayman Chemical Company, USA or R&D Systems, UK) according to manufacturers' instructions.

[0330] The amount of Substance P released by the neuronal cells in the presence of the TM of interest is compared to the release obtained in the presence and absence of 100 mM KCl. Stimulation of Substance P release by the TM of interest above the basal release, establishes that the TM of interest is an "agonist ligand" as defined in this specification. If desired the stimulation of Substance P release by the TM of interest can be compared to a standard Substance P release-curve produced using the natural ORL-1 receptor ligand, nociceptin (Tocris).

Example 11

Confirmation of ORL₁ Receptor Activation by Measuring Forskolin-Stimulated cAMP Production

[0331] Confirmation that a given TM is acting via the ORL₁ receptor is provided by the following test, in which the TMs ability to inhibit forskolin-stimulated cAMP production is assessed.

Materials

[0332] [³H]adenine and [¹⁴C]cAMP are obtained from GE Healthcare

Methods

[0333] The test is conducted essentially as described previously by Meunier et al. [Isolation and structure of the

endogenous agonist of opioid receptor-like ORL₁ receptor. Nature 377: 532-535, 1995] in intact transfected-CHO cells plated on 24-well plastic plates.

[0334] To the cells is added [3H]adenine (1.0 μ Ci) in 0.4 ml of culture medium. The cells remain at 37° C. for 2 h to allow the adenine to incorporate into the intracellular ATP pool. After 2 h, the cells are washed once with incubation buffer containing: 130 mM NaCl, 4.8 mM KCl, 1.2 mM KH₂PO₄, 1.3 mM CaCl₂, 1.2 mM MgSO₄, 10 mM glucose, 1 mg/ml bovine serum albumin and 25 mM HEPES pH 7.4, and replaced with buffer containing forskolin (10 μ M) and isobutylmethylxanthine (50 μ M) with or without the TM of interest. After 10 min, the medium is aspirated and replaced with 0.5 ml, 0.2 M HCl. Approximately 1000 cpm of [¹⁴C]cAMP is added to each well and used as an internal standard. The contents of the wells are then transferred to columns of 0.65 g dry alumina powder. The columns are eluted with 4 ml of 5 mM HCl, 0.5 ml of 0.1 M ammonium acetate, then two additional milliliters of ammonium acetate. The final eluate is collected into scintillation vials and counted for ¹⁴C and tritium. Amounts collected are corrected for recovery of [¹⁴C]cAMP. TMs that are agonists at the ORL₁ receptor cause a reduction in the level of cAMP produced in response to forskolin.

Example 12

Confirmation of ORL₁ Receptor Activation Using a GTP γ S Binding Functional Assay

[0335] Confirmation that a given TM is acting via the ORL₁ receptor is also provided by the following test, a GTP γ S binding functional assay.

Materials

[0336] [³⁵S]GTP γ S is obtained from GE Healthcare. Wheatgerm agglutinin-coated (SPA) beads are obtained from GE Healthcare.

Methods

[0337] This assay is carried out essentially as described by Traynor and Nahorski [Modulation by μ -opioid agonists of guanosine-5'-O-(3-[³⁵S]thio)triphosphate binding to membranes from human neuroblastoma SH-SY5Y cells. Mol. Pharmacol. 47: 848-854, 1995].

[0338] Cells are scraped from tissue culture dishes into 20 mM HEPES, 1 mM ethylenediaminetetraacetic acid, then centrifuged at 500 $\times$ g for 10 min. Cells are resuspended in this buffer and homogenized with a Polytron Homogenizer.

[0339] The homogenate is centrifuged at 27,000 $\times$ g for 15 min, and the pellet resuspended in buffer A, containing: 20 mM HEPES, 10 mM MgCl₂, 100 mM NaCl, pH 7.4. The suspension is recentrifuged at 20,000 $\times$ g and suspended once more in buffer A. For the binding assay, membranes (8-15 μ g protein) are incubated with [³⁵S]GTP S (50 μ M), GDP (10 μ M), with and without the TM of interest, in a total volume of 1.0 ml, for 60 min at 25° C. Samples are filtered over glass fibre filters and counted as described for the binding assays.

Example 13

Preparation of a LC/A-nociceptin-H_N/A Fusion Protein (Nociceptin is N-Terminal of the H_N-Chain)

[0340] The linker-nociceptin-spacer insert is prepared as described in Example 2.

Preparation of the LC/A-nociceptin-H_N/A Fusion

[0341] In order to create the LC-linker-nociceptin-spacer-H_N construct (SEQ ID13), the pCR 4 vector encoding the

linker (SEQ ID7) is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient for insertion and ligation of the LC/A DNA (SEQ ID1) also cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with BamHI+HindIII restriction enzymes and the LC/A-linker fragment inserted into a similarly cleaved vector containing a unique multiple cloning site for BamHI, Sall, PstI, and HindIII such as the pMAL vector (NEB). The H_N/A DNA (SEQ ID2) is then cleaved with PstI+HindIII restriction enzymes and inserted into the similarly cleaved pMAL-LC/A-linker construct. The final construct contains the LC-linker-nociceptin-spacer-H_N ORF (SEQ ID13) for expression as a protein of the sequence illustrated in SEQ ID14.

Example 14

Preparation of a Nociceptin-LC/A-H_N/A Fusion Protein (Nociceptin is N-Terminal of the LC-Chain)

[0342] In order to create the nociceptin-spacer-LC/A-H_N/A construct, an A serotype linker with the addition of a Factor Xa site for activation, arranged as BamHI-Sall-linker-protease site-linker-PstI-XbaI-stop codon-HindIII (SEQ ID8) is synthesised as described in Example 13. The pCR 4 vector encoding the linker is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient for insertion and ligation of the LC/A DNA (SEQ ID1) also cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with BamHI+HindIII restriction enzymes and the LC/A-linker fragment inserted into a similarly cleaved vector containing the synthesised N-terminal presentation nociceptin insert (SEQ ID9). This construct is then cleaved with Aval+HindIII and inserted into an expression vector such as the pMAL plasmid (NEB). The H_N/A DNA (SEQ ID2) is then cleaved with PstI+HindIII restriction enzymes and inserted into the similarly cleaved pMAL-nociceptin-LC/A-linker construct. The final construct contains the nociceptin-spacer-LC/A-H_N/A ORF (SEQ ID51) for expression as a protein of the sequence illustrated in SEQ ID52.

Example 15

Preparation and Purification of an LC/A-nociceptin-H_N/A Fusion Protein Family with Variable Spacer Length

[0343] Using the same strategy as employed in Example 2, a range of DNA linkers were prepared that encoded nociceptin and variable spacer content. Using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)], the DNA sequence encoding the linker-ligand-spacer region is determined. Restriction sites are then incorporated into the DNA sequence and can be arranged as BamHI-Sall-linker-protease site-nociceptin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID53 to SEQ ID57). It is important to ensure the correct reading frame is maintained for the spacer, nociceptin and restriction sequences and that the XbaI sequence is not preceded by the bases, TC which would result on DAM methylation. The DNA sequence is screened for restriction sequence incorporation and any additional sequences are removed manually from the remaining sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage

Analysers (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0344] The spacers that were created included:

TABLE 1

Code	Protein sequence of the linker	SEQ ID of the linker DNA
GS10	ALAGGGGSALVLQ	53
GS15	ALAGGGSGGGGSALVLQ	54
GS25	ALAGGGSGGGSGGGSGGGGSALVLQ	55
GS30	ALAGGGSGGGSGGGSGGGSGGGGSALVLQ	56
HX27	ALAAEAAAKEAAAKEAAKAGGGGSALVLQ	57

[0345] By way of example, in order to create the LC/A-CPN(GS15)-H_N/A fusion construct (SEQ ID58), the pCR 4 vector encoding the linker (SEQ ID54) is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient vector for insertion and ligation of the LC/A DNA (SEQ ID1) also cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with BamHI+HindIII restriction enzymes and the LC/A-linker fragment inserted into a similarly cleaved vector containing a unique multiple cloning site for BamHI, Sall, PstI, and HindIII such as the pMAL vector (NEB). The H_N/A DNA (SEQ ID2) is then cleaved with PstI+HindIII restriction enzymes and inserted into the similarly cleaved pMAL-LC/A-linker construct. The final construct contains the LC/A-CPN(GS15)-H_N/A ORF (SEQ ID58) for expression as a protein of the sequence illustrated in SEQ ID59.

[0346] As a further example, to create the LC/A-CPN(GS25)-H_N/A fusion construct (SEQ ID60), the pCR 4 vector encoding the linker (SEQ ID55) is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient vector for insertion and ligation of the LC/A DNA (SEQ ID1) cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with BamHI+HindIII restriction enzymes and the LC/A-linker fragment inserted into a similarly cleaved vector containing a unique multiple cloning site for BamHI, Sall, PstI, and HindIII such as the pMAL vector (NEB). The H_N/A DNA (SEQ ID2) is then cleaved with PstI+HindIII restriction enzymes and inserted into the similarly cleaved pMAL-LC/A-linker construct. The final construct contains the LC/A-CPN(GS25)-H_N/A ORF (SEQ ID60) for expression as a protein of the sequence illustrated in SEQ ID61.

[0347] Variants of the LC/A-CPN-H_N/A fusion consisting of GS10, GS30 and HX27 are similarly created. Using the purification methodology described in Example 9, fusion protein is purified from *E. coli* cell paste. FIG. 9 illustrates the purified product obtained in the case of LC/A-CPN(GS10)-H_N/A, LC/A-CPN(GS15)-H_N/A, LC/A-CPN(GS25)-H_N/A, LC/A-CPN(GS30)-H_N/A and LC/A-CPN(HX27)-H_N/A.

Example 16

Assessment of in Vitro Efficacy of an LC/A-nociceptin-H_N/A Fusion

[0348] Fusion protein prepared according to Examples 2 and 9 was assessed in the eDRG neuronal cell model.

[0349] Assays for the inhibition of substance P release and cleavage of SNAP-25 have been previously reported (Duggan et al., 2002, *J. Biol. Chem.*, 277, 34846-34852). Briefly, dorsal root ganglia neurons are harvested from 15-day-old fetal Sprague-Dawley rats and dissociated cells plated onto 24-well plates coated with Matrigel at a density of 1×10⁶ cells/well. One day post-plating the cells are treated with 10 μM cytosine β-D-arabinofuranoside for 48 h. Cells are maintained in Dulbecco's minimal essential medium supplemented with 5% heat-inactivated fetal bovine serum, 5 mM L-glutamine, 0.6% D-glucose, 2% B27 supplement, and 100 ng/ml 2.5 S mouse nerve growth factor. Cultures are maintained for 2 weeks at 37° C. in 95% air/5% CO₂ before addition of test materials.

[0350] Release of substance P from eDRG is assessed by enzyme-linked immunosorbent assay. Briefly, eDRG cells are washed twice with low potassium-balanced salt solution (BSS: 5 mM KCl, 137 mM NaCl, 1.2 mM MgCl₂, 5 mM glucose, 0.44 mM KH₂PO₄, 20 mM HEPES, pH 7.4, 2 mM CaCl₂). Basal samples are obtained by incubating each well for 5 min. with 1 ml of low potassium BSS. After removal of this buffer, the cells are stimulated to release by incubation with 1 ml of high potassium buffer (BSS as above with modification to include 100 mM KCl isotonicity balanced with NaCl) for 5 min. All samples are removed to tubes on ice prior to assay of substance P. Total cell lysates are prepared by addition of 250 μl of 2 M acetic acid/0.1% trifluoroacetic acid to lyse the cells, centrifugal evaporation, and resuspension in 500 μl of assay buffer. Diluted samples are assessed for substance P content. Substance P immunoreactivity is measured using Substance P Enzyme Immunoassay Kits (Cayman Chemical Company or R&D Systems) according to manufacturers' instructions. Substance P is expressed in pg/ml relative to a standard substance P curve run in parallel.

[0351] SDS-PAGE and Western blot analysis were performed using standard protocols (Novex). SNAP-25 proteins were resolved on a 12% Tris/glycine polyacrylamide gel (Novex) and subsequently transferred to nitrocellulose membrane. The membranes were probed with a monoclonal antibody (SMI-81) that recognises cleaved and intact SNAP-25. Specific binding was visualised using peroxidase-conjugated secondary antibodies and a chemiluminescent detection system. Cleavage of SNAP-25 was quantified by scanning densitometry (Molecular Dynamics Personal SI, ImageQuant data analysis software). Percent SNAP-25 cleavage was calculated according to the formula: (Cleaved SNAP-25/(Cleaved+Intact SNAP-25))×100.

[0352] Following exposure of eDRG neurons to an LC/A-nociceptin-H_N/A fusion (termed CPN-A), both inhibition of substance P release and cleavage of SNAP-25 are observed (FIG. 10). After 24 h exposure to the fusion, 50% of maximal SNAP-25 cleavage is achieved by a fusion concentration of 6.3±2.5 nM.

[0353] The effect of the fusion is also assessed at defined time points following a 16 h exposure of eDRG to CPN-A. FIG. 11 illustrates the prolonged duration of action of the CPN-A fusion protein, with measurable activity still being observed at 28 days post exposure.

Example 17

Assessment of in Vitro Efficacy of an LC/A-nociceptin Variant-H_N/A Fusion

[0354] Fusion protein prepared according to Examples 8 and 9 was assessed in the eDRG neuronal cell mode using the method described in Example 16.

[0355] Following exposure of eDRG neurons to an LC/A-nociceptin variant- H_N/A fusion (termed CPNv-A), both inhibition of substance P release and cleavage of SNAP-25 are observed. After 24 h exposure to the fusion, 50% of maximal SNAP-25 cleavage is achieved by a fusion concentration of 1.4 ± 0.4 nM (FIG. 12).

[0356] The effect of the fusion is also assessed at defined time points following a 16 h exposure of eDRG to CPN-A. FIG. 13 illustrates the prolonged duration of action of the CPN-A fusion protein, with measurable activity still being observed at 24 days post exposure.

[0357] The binding capability of the CPNv-A fusion protein is also assessed in comparison to the CPN-A fusion. FIG. 14 illustrates the results of a competition experiment to determine binding efficacy at the ORL-1 receptor. CPNv-A is demonstrated to displace $[3H]$ -nociceptin, thereby confirming that access to the receptor is possible with the ligand in the central presentation format.

Example 18

Preparation of an LC/A-nociceptin Variant- H_N/A Fusion Protein that is Activated by Treatment with Enterokinase

[0358] Following the methods used in Examples 1 and 2, the LC/A (SEQ ID1) and H_N/A (SEQ ID2) are created and inserted into the A serotype nociceptin variant linker arranged as BamHI-Sall-linker-enterokinase protease site-nociceptin variant-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID62). The final construct contains the LC-linker-nociceptin variant-spacer- H_N ORF sequences (SEQ ID63) for expression as a protein of the sequence illustrated in SEQ ID64. The fusion protein is termed CPNv(Ek)-A. FIG. 15 illustrates the purification of CPNv(Ek)-A from *E. coli* following the methods used in Example 9 but using Enterokinase for activation at $0.00064 \mu\text{g}$ per $100 \mu\text{g}$ of fusion protein.

Example 19

Assessment of in Vitro Efficacy of a LC/A-nociceptin Variant- H_N/A Fusion that has been Activated by Treatment with Enterokinase

[0359] The CPNv(Ek)-A prepared in Example 18 is obtained in a purified form and applied to the eDRG cell model to assess cleavage of SNAP-25 (using methodology from Example 16). FIG. 16 illustrates the cleavage of SNAP-25 following 24 h exposure of eDRG to CPNv(Ek)-A. The efficiency of cleavage is observed to be similar to that achieved with the Factor Xa-cleaved material, as recorded in Example 17.

Example 20

Preparation of an LC/C-nociceptin Variant- H_N/C Fusion Protein with a Factor Xa Activation Linker Derived from Serotype A

[0360] Following the methods used in Example 4, the LC/C (SEQ ID5) and H_N/C (SEQ ID6) are created and inserted into the A serotype nociceptin variant linker arranged as BamHI-Sall-linker-nociceptin variant-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID65). The final construct contains the LC-linker-nociceptin variant-spacer- H_N ORF sequences (SEQ ID66) for expression as a protein of the sequence illustrated in SEQ ID67. The fusion protein is termed CPNv-C

(act. A). FIG. 17 illustrates the purification of CPNv-C (act. A) from *E. coli* following the methods used in Example 9.

Example 21

Assessment of in Vitro Efficacy of an LC/C-nociceptin Variant- H_N/C Fusion Protein

[0361] Following the methods used in Example 9, the CPNv-C (act. A) prepared in Example 20 is obtained in a purified form and applied to the eDRG cell model to assess cleavage of SNAP-25 (using methodology from Example 16). After 24 h exposure to the fusion, 50% of maximal syntaxin cleavage is achieved by a fusion concentration of 3.1 ± 2.0 nM. FIG. 18 illustrates the cleavage of syntaxin following 24 h exposure of eDRG to CPNv-C (act. A).

Example 22

Assessment of in Vivo Efficacy of an LC/A-nociceptin- H_N/A Fusion

[0362] The ability of an LC/A-nociceptin- H_N/A fusion (CPN/A) to inhibit acute capsaicin-induced mechanical allodynia is evaluated following subcutaneous intraplantar injection in the rat hind paw. Test animals are evaluated for paw withdrawal frequency (PWF %) in response to a 10 g Von Frey filament stimulus series (10 stimuli $\times$ 3 trials) prior to recruitment into the study, after subcutaneous treatment with CPN/A but before capsaicin, and following capsaicin challenge post-injection of CPN/A (average of responses at 15' and 30'). Capsaicin challenge is achieved by injection of 10 μL of a 0.3% solution. Sample dilutions are prepared in 0.5% BSA/saline. FIG. 19 illustrates the reversal of mechanical allodynia that is achieved by pre-treatment of the animals with a range of concentrations of LC/A-nociceptin- H_N/A fusion.

[0363] The ability of an LC/A-nociceptin- H_N/A fusion (CPN/A) to inhibit streptozotocin (STZ)-induced mechanical (tactile) allodynia in rats is evaluated. STZ-induced mechanical allodynia in rats is achieved by injection of streptozotocin (i.p. or i.v.) which yields destruction of pancreatic β -cells leading to loss of insulin production, with concomitant metabolic stress (hyperglycemia and hyperlipidemia). As such, STZ induces Type I diabetes. In addition, STZ treatment leads to progressive development of neuropathy, which serves as a model of chronic pain with hyperalgesia and allodynia that may reflect signs observed in diabetic humans (peripheral diabetic neuropathy).

[0364] Male Sprague-Dawley rats (250-300 g) are treated with 65 mg/kg STZ in citrate buffer (I.V.) and blood glucose and lipid are measured weekly to define the readiness of the model. Paw Withdrawal Threshold (PWT) is measured in response to a Von Frey filament stimulus series over a period of time. Allodynia is said to be established when the PWT on two consecutive test dates (separated by 1 week) measures below 6 g on the scale. At this point, rats are randomized to either a saline group (negative efficacy control), gabapentin group (positive efficacy control) or a test group (CPN/A). Test materials (20-25 μL are injected subcutaneously as a single injection (except gabapentin) and the PWT is measured at 1 day post-treatment and periodically thereafter over a 2-week period. Gabapentin (30 mg/kg i.p. @ 3 ml/kg injection volume) is injected daily, 2 hours prior to the start of PWT testing. FIG. 20 illustrates the reversal of allodynia achieved

by pre-treatment of the animals with 750 ng of CPN/A. Data were obtained over a 2-week period after a single injection of CPN/A

Example 23

Assessment of in Vivo Efficacy of an LC/A-nociceptin Variant-H_N/A Fusion

[0365] The ability of an LC/A-nociceptin variant-H_N/A fusion (CPNv/A) to inhibit capsaicin-induced mechanical allodynia is evaluated following subcutaneous intraplantar injection in the rat hind paw. Test animals are evaluated for paw withdrawal frequency (PWF %) in response to a 10 g Von Frey filament stimulus series (10 stimuli×3 trials) prior to recruitment into the study (Pre-Treat); after subcutaneous intraplantar treatment with CPNv/A but before capsaicin (Pre-CAP); and following capsaicin challenge post-injection of CPNv/A (average of responses at 15' and 30'; CAP). Capsaicin challenge is achieved by injection of 10 μL of a 0.3% solution. Sample dilutions are prepared in 0.5% BSA/saline.

[0366] FIG. 21 illustrates the reversal of allodynia that is achieved by pre-treatment of the animals with a range of concentrations of LC/A-nociceptin variant-H_N/A fusion in comparison to the reversal achieved with the addition of LC/A-nociceptin-H_N/A fusion. These data are expressed as a normalized paw withdrawal frequency differential, in which the difference between the peak response (post-capsaicin) and the baseline response (pre-capsaicin) is expressed as a percentage. With this analysis, it can be seen that CPNv/A is more potent than CPN/A since a lower dose of CPNv/A is required to achieve similar analgesic effect to that seen with CPN/A.

Example 24

Preparation of an LC/A-leu Enkephalin-H_N/A Fusion Protein

[0367] Due to the small, five-amino acid, size of the leu-enkephalin ligand the LC/A-leu enkephalin-H_N/A fusion is created by site directed mutagenesis [for example using Quickchange (Stratagene Inc.)] using the LC/A-nociceptin-H_N/A fusion (SEQ ID13) as a template. Oligonucleotides are designed encoding the YGGFL leu-enkephalin peptide, ensuring standard *E. coli* codon usage is maintained and no additional restriction sites are incorporated, flanked by sequences complimentary to the linker region of the LC/A-nociceptin-H_N/A fusion (SEQ ID13) either side on the nociceptin section. The SDM product is checked by sequencing and the final construct containing the LC-linker-leu enkephalin-spacer-H_N ORF (SEQ ID68) for expression as a protein of the sequence illustrated in SEQ ID69. The fusion protein is termed CPLE-A. FIG. 22 illustrates the purification of CPLE-A from *E. coli* following the methods used in Example 9.

Example 25

Expression and Purification of an LC/A-beta-endorphin-H_N/A Fusion Protein

[0368] Following the methods used in Example 9, and with the LC/A-beta-endorphin-H_N/A fusion protein (termed

CPBE-A) created in Example 7, the CPBE-A is purified from *E. coli*. FIG. 23 illustrates the purified protein as analysed by SDS-PAGE.

Example 26

Preparation of an LC/A-nociceptin Mutant-H_N/A Fusion Protein

[0369] Due to the single amino acid modification necessary to mutate the nociceptin sequence at position 1 from a Phe to a Tyr, the LC/A-nociceptin mutant-H_N/A fusion is created by site directed mutagenesis [for example using Quickchange (Stratagene Inc.)] using the LC/A-nociceptin-H_N/A fusion (SEQ ID13) as a template. Oligonucleotides are designed encoding tyrosine at position 1 of the nociceptin sequence, ensuring standard *E. coli* codon usage is maintained and no additional restriction sites are incorporated, flanked by sequences complimentary to the linker region of the LC/A-nociceptin-H_N/A fusion (SEQ ID13) either side on the nociceptin section. The SDM product is checked by sequencing and the final construct containing the LC/A-nociceptin mutant-spacer-H_N/A fusion ORF (SEQ ID70) for expression as a protein of the sequence illustrated in SEQ ID71. The fusion protein is termed CPOP-A. FIG. 24 illustrates the purification of CPOP-A from *E. coli* following the methods used in Example 9.

Example 27

Preparation and Assessment of an LC/A-nociceptin Variant Mutant-H_N/A Fusion Protein

[0370] Due to the single amino acid modification necessary to mutate the nociceptin sequence at position 1 from a Phe to a Tyr, the LC/A-nociceptin variant mutant-H_N/A fusion is created by site directed mutagenesis [for example using Quickchange (Stratagene Inc.)] using the LC/A-nociceptin variant-H_N/A fusion (SEQ ID25) as a template. Oligonucleotides are designed encoding tyrosine at position 1 of the nociceptin sequence, ensuring standard *E. coli* codon usage is maintained and no additional restriction sites are incorporated, flanked by sequences complimentary to the linker region of the LC/A-nociceptin variant-H_N/A fusion (SEQ ID25) either side on the nociceptin section. The SDM product is checked by sequencing and the final construct containing the LC/A-nociceptin mutant-spacer-H_N/A fusion ORF (SEQ ID72) for expression as a protein of the sequence illustrated in SEQ ID73. The fusion protein is termed CPOPv-A. FIG. 25 illustrates the purification of CPOPv-A from *E. coli* following the methods used in Example 9.

[0371] Using methodology described in Example 16, CPOPv-A is assessed for its ability to cleave SNAP-25 in the eDRG cell model. FIG. 26 illustrates that CPOPv-A is able to cleave SNAP-25 in the eDRG model, achieving cleavage of 50% of the maximal SNAP-25 after exposure of the cells to approximately 5.9 nM fusion for 24 h.

Example 28

Preparation of an IgA Protease-nociceptin Variant-H_N/A Fusion Protein

[0372] The IgA protease amino acid sequence was obtained from freely available database sources such as GenBank (accession number P09790). Information regarding the structure of the *N. Gonorrhoeae* IgA protease gene is available in the

literature (Pohlner et al., Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease, Nature, 1987, 325(6103), 458-62). Using Backtranslation tool v2.0 (Entelechon), the DNA sequence encoding the IgA protease modified for *E. coli* expression was determined. A BamHI recognition sequence was incorporated at the 5' end and a codon encoding a cysteine amino acid and SalI recognition sequence were incorporated at the 3' end of the IgA DNA. The DNA sequence was screened using MapDraw, (DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required for cloning were removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage was assessed Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables. This optimised DNA sequence (SEQ ID74) containing the IgA open reading frame (ORF) is then commercially synthesized.

[0373] The IgA (SEQ ID74) is inserted into the LC-linker-nociceptin variant-spacer-H_N ORF (SEQ ID25) using BamHI and SalI restriction enzymes to replace the LC with the IgA protease DNA. The final construct contains the IgA-linker-nociceptin variant-spacer-H_N ORF (SEQ ID75) for expression as a protein of the sequence illustrated in SEQ ID76.

Example 29

Preparation and Assessment of a Nociceptin Targeted Endopeptidase Fusion Protein with a Removable Histidine Purification Tag

[0374] DNA was prepared that encoded a Factor Xa removable his-tag (his6), although it is clear that alternative proteases site such as Enterokinase and alternative purification tags such as longer histidine tags are also possible. Using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)], the DNA sequence encoding the Factor Xa removable his-tag region is determined. Restriction sites are then incorporated into the DNA sequence and can be arranged as NheI-linker-SpeI-PstI-H_N/A-XbaI-LEIEGRSGHHHHHStop codon-HindIII (SEQ ID77). The DNA sequence is screened for restriction sequence incorporated and any additional sequences are removed manually from the remaining sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector. In order to create CPNV-A-FXa-HT (SEQ ID78, removable his-tag construct) the pCR 4 vector encoding the removable his-tag is cleaved with NheI and HindIII. The NheI-HindIII fragment is then inserted into the LC/A-CPNV-H_N/A vector (SEQ ID25) that has also been cleaved by NheI and HindIII. The final construct contains the LC/A-linker-nociceptin variant-spacer-H_N-FXa-Histag-HindIII ORF sequences (SEQ ID78) for expression as a protein of the

sequence illustrated in SEQ ID79. FIG. 27 illustrates the purification of CPNV-A-FXa-HT from *E. coli* following the methods used in Example 9.

Example 30

Preparation of a Leu-Enkephalin Targeted Endopeptidase Fusion Protein Containing a Translocation Domain Derived from Diphtheria Toxin

[0375] The DNA sequence is designed by back translation of the amino acid sequence of the translocation domain of the diphtheria toxin (obtained from freely available database sources such as GenBank (accession number 1xDTT) using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)]. Restriction sites are then incorporated into the DNA sequence and can be arranged as NheI-Linker-SpeI-PstI-diphtheria translocation domain-XbaI-stop codon-HindIII (SEQ ID80). PstI/XbaI recognition sequences are incorporated at the 5' and 3' ends of the translocation domain respectively of the sequence maintaining the correct reading frame. The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence containing the diphtheria translocation domain is then commercially synthesized as NheI-Linker-SpeI-PstI-diphtheria translocation domain-XbaI-stop codon-HindIII (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector (Invitrogen). The pCR 4 vector encoding the diphtheria translocation domain is cleaved with NheI and XbaI. The NheI-XbaI fragment is then inserted into the LC/A-CPLE-H_N/A vector (SEQ ID68) that has also been cleaved by NheI and XbaI. The final construct contains the LC/A-leu-enkephalin-spacer-diphtheria translocation domain ORF sequences (SEQ ID81) for expression as a protein of the sequence illustrated in SEQ ID82.

Example 31

Preparation of a Nociceptin Variant Targeted Endopeptidase Fusion Protein Containing a LC Domain Derived from Tetanus Toxin

[0376] The DNA sequence is designed by back translation of the tetanus toxin LC amino acid sequence (obtained from freely available database sources such as GenBank (accession number X04436) using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)]. BamHI/SalI recognition sequences are incorporated at the 5' and 3' ends respectively of the sequence maintaining the correct reading frame (SEQ ID83). The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required by

the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence containing the tetanus toxin LC open reading frame (ORF) is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector (Invitrogen). The pCR 4 vector encoding the TeNT LC is cleaved with BamHI and SalI. The BamHI-SalI fragment is then inserted into the LC/A-CPNV-H_NA vector (SEQ ID25) that has also been cleaved by BamHI and SalI. The final construct contains the TeNT LC-linker-nociceptin variant-spacer-H_N ORF sequences (SEQ ID84) for expression as a protein of the sequence illustrated in SEQ ID85.

Example 32

Preparation of an LC/C-nociceptin Variant-H_N/C Fusion Protein with a Native Serotype C Linker that is Susceptible to Factor Xa Cleavage

[0377] Following the methods used in Example 4, the LC/C (SEQ ID5) and H_N/C (SEQ ID6) are created and inserted into the C serotype nociceptin variant linker arranged as BamHI-SalI-linker-nociceptin variant-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII (SEQ ID86). The final construct contains the LC-linker-nociceptin variant-spacer-H_N ORF sequences (SEQ ID87) for expression as a protein of the sequence illustrated in SEQ ID88. The fusion protein is termed CPNV-C (act. C).

Example 33

Construction of CHO-K1 OP2 Receptor Activation Assay and SNAP-25 Cleavage Assay

Cell-Line Creation

[0378] CHO OP2 cell line was purchased from Perkin Elmer (ES-541-C, lot 451-719-A). Cells were transfected with SNAP-25 DNA using Lipofectamine™ 2000 and incubated for 4 hours before media replacement. After 24 hours, cells were transferred to a T175 flask. 100 µg/ml Zeocin was added after a further 24 hours to begin selection of SNAP-25 expressing cells, and 5 µg/ml Blasticidin added to maintain selective pressure for the receptor. Cells were maintained in media containing selection agents for two weeks, passaging cells every two to three days to maintain 30-70% confluence. Cells were then diluted in selective media to achieve 0.5 cell per well in a 96 well microplate. After a few days, the plates were examined under a microscope, and those containing single colonies were marked. Media in these wells was changed weekly. As cells became confluent in the wells, they were transferred to T25 flasks. When they had expanded sufficiently each clone was seeded to 24 wells of a 96 well plate, plus a frozen stock vial created. LC/A-CPDY-H_NA fusion and LC/A-H_NA were applied to the cells for 24 hours, and then western blots performed to detect SNAP-25 cleavage. Clones from which SNAP-25 bands were strong and cleavage levels were high with fusion were maintained for further investigation. Full dose curves were run on these, and

the clone (D30) with the highest differential between LC/A-CPDY-H_NA fusion and LC/A-H_NA cleavage levels was selected.

OP2 Receptor Activation Assay

[0379] The OP2 receptor activation measures the potency and intrinsic efficacy of ligands at OP2 receptor in transfected CHO-K1 cells by quantifying the reduction of forskolin-stimulated intracellular cAMP using a FRET-based cAMP (Perkin Elmer LANCE cAMP kit). After stimulation, a fluorescently labelled cAMP tracer (Europium-streptavidin/biotin-cAMP) and fluorescently (Alexa) labelled anti-cAMP antibody are added to the cells in a lysis buffer. cAMP from the cells competes with the cAMP tracer for antibody binding sites. When read, a light pulse at 320 nm excites the fluorescent portion (Europium) of the cAMP tracer. The energy emitted from the europium is transferred to the Alexa fluor-labelled antibodies bound to the tracer, generating a TR-FRET signal at 665 nm (Time-resolved fluorescence resonance energy transfer is based on the proximity of the donor label, europium, and the acceptor label, Alexa fluor, which have been brought together by a specific binding reaction). Residual energy from the europium produces light at 615 nm. In agonist treated cells there will be less cAMP to compete with the tracer so a dose dependant increase in signal at 665 nm will be observed compared with samples treated with forskolin alone. The signal at 665 nm signal is converted to cAMP concentration by interpolation to a cAMP standard curve which is included in each experiment.

Culture of Cells for Receptor Activation Assay:

[0380] Cells were seeded and cultured in T175 flasks containing Ham F12 with Glutamax, 10% Foetal bovine serum, 5 µg ml⁻¹ Blasticidin and 100 µg ml⁻¹ Zeocin. The flasks were incubated at 37° C. in a humidified environment containing 5% CO₂ until 60-80% confluent. On the day of harvest the media was removed and the cells washed twice with 25 ml PBS. The cells were removed from the flask by addition of 10 ml of Tryple Express, and incubation at 37° C. for 10 min followed by gentle tapping of the flask. The dislodged cells were transferred to a 50 ml centrifuge tube and the flask washed twice with 10 ml media which was added to the cell suspension. The tube was centrifuged at 1300×g for 3 min and the supernatant removed. Cells were gently re-suspended in 10 ml media (if freezing cells) or assay buffer (if using 'fresh' cells in assay), and a sample was removed for counting using a nucleocounter (ChemoMetec). Cells for use 'fresh' in an assay were diluted further in assay buffer to the appropriate concentration. Cells harvested for freezing were re-centrifuged (1300×g; 3 min), the supernatant removed and cells re-suspended in Synth-a-freeze at 4° C. to 3×10⁶ cells/ml. Cryovials containing 1 ml suspension each were placed in a chilled Nalgene Mr Frosty freezing container (-1° C./minute cooling rate), and left overnight in a -80° C. freezer. The following day vials were transferred to the vapour phase of a liquid nitrogen storage tank.

Dilution of Test Materials and Cell Assay

[0381] Using Gilson pipettes and Sigmacoted or lo-bind tips, test materials and standards were diluted to the appropriate concentrations in the wells of the first two columns of an eppendorf 500 µl deep-well lo-bind plate, in assay buffer containing 10 µM forskolin. The chosen concentrations in

columns one and two were half a log unit apart. From these, serial 1:10 dilutions were made across the plate (using an electronic eight channel pipette with sigmacote or lo-bind tips) until eleven concentrations at half log intervals had been created. In the twelfth column, assay buffer only was added as a 'basal'. Using a 12 channel digital pipette, 10 μ M of sample from the lo-bind plate was transferred to the optiplate 96 well microplate.

[0382] To wells containing the standard curve, 10 μ l of assay buffer was added using a multichannel digital pipette. To wells containing the test materials, 10 μ l of cells in assay buffer at the appropriate concentration were added. Plates were sealed and incubated for 120 min at room temperature, for the first hour on an IKA MTS 2/4 orbital shaker set to maximum speed.

Detection

[0383] LANCE Eu-W8044 labelled streptavidin (Eu-SA) and Biotin-cAMP (b-cAMP) were diluted in cAMP Detection Buffer (both from Perkin Elmer LANCE cAMP kit) to create sub-stocks, at dilution ratios of 1:17 and 1:5, respectively. The final detection mix was prepared by diluting from the two sub stocks into detection buffer at a ratio of 1:125. The mixture was incubated for 15-30 min at room temperature before addition of 1:200 Alexa Fluor® 647-anti cAMP Antibody (Alexa-Fluor Ab). After briefly vortex mixing, 20 μ l was immediately added to each well using a digital multichannel pipette. Microplate sealers were applied and plates incubated for 24 h at room temperature (for the first hour on an IKA MTS 2/4 orbital shaker set to maximum speed). Plate sealers were removed prior to reading on the Envision.

[0384] FIGS. 33 and 34 show that dynorphin conjugates with LC/A-H_N/A, LC/B-H_N/B, LC/C-H_N/C and LC/D-H_N/D backbones active the OP2 receptor.

CHO-K1 OP2 SNAP-25 Cleavage Assay

[0385] Cultures of cells were exposed to varying concentrations of fusion protein for 24 hours. Cellular proteins were separated by SDS-PAGE and western blotted with anti-SNAP-25 antibody to facilitate assessment of SNAP-25 cleavage. SNAP-25 cleavage calculated by densitometric analysis (Syngene).

Plating Cells

[0386] Prepare cells at 2×10⁵ cells/ml and seed 125 μ l per well of 96 well plate. Use the following media: 500 ml Gibco Ham F12 with Glutamax (product code 31765068), 50 ml FBS, 5 μ g/ml Blasticidin (250 μ l aliquot from box in freezer, G13) (Calbiochem #203351, 10 ml at 10 mg/ml), 100 μ g/ml Zeocin (500 μ l from box in freezer, G35). (Invitrogen from Fisher, 1 g in 8×1.25 ml tubes at 100 mg/ml product code VXR25001). Allow cells to grow for 24 hrs (37° C., 5 CO₂, humidified atmosphere).

Cell Treatment

[0387] Prepare dilutions of test protein for a dose range of each test proteins (make up double (2×) the desired final concentrations because 125 μ l will be applied directly onto 125 μ l of media already in each well). Filter sterilize CHO KOR D30 feeding medium (20 ml syringe, 0.2 μ m syringe filter) to make the dilutions. Add the filtered medium into 5 labelled bijoux's (7 ml tubes), 0.9 ml each using a Gilson pipette or multi-stepper. Dilute the stock test protein to 2000

nM (working stock solution 1) and 600 nM (working stock solution 2). Using a Gilson pipette prepare 10-fold serial dilutions of each working stock, by adding 100 μ l to the next concentration in the series. Pipette up and down to mix thoroughly. Repeat to obtain 4 serial dilutions for solution 1, and 3 serial dilutions for solution 2. A 0 nM control (filtered feeding medium only) should also be prepared as a negative control for each plate. Repeat the above for each test protein. In each experiment a 'standard' batch of material must be included as control/reference material, this is unliganded LC/A-H_N/A.

Apply Diluted Sample to CHO KOR D30 Plates

[0388] Apply 125 μ l of test sample (double concentration) per well. Each test sample should be applied to triplicate wells and each dose range should include a 0 nM control. Incubate for 24 hrs (37° C., 5% CO₂, humidified atmosphere).

Cell Lysis

[0389] Prepare fresh lysis buffer (20 mls per plate) with 25% (4×) NuPAGE LDS sample buffer, 65% dH₂O and 10% 1 M DTT. Remove medium from the CHO KOR D30 plate by inverting over a waste receptacle. Drain the remaining media from each well using a fine-tipped pipette. Lyse the cells by adding 125 μ l of lysis buffer per well using a multi-stepper pipette. After a minimum of 20 mins, remove the buffer from each well to a 1.5 ml microcentrifuge tube. Tubes must be numbered to allowing tracking of the CHO KOR treatments throughout the blotting procedure. A1-A3 down to H1-H3 numbered 1-24, A4-A6 down to H4-H6 numbered 25-48, A7-A9 down to H7-H9 numbered 49-72, A10-A12 down to H10-H12 numbered 73-96. Vortex each sample and heat at 90° C. for 5-10 mins in a prewarmed heat block. Store at -20° C. or use on the same day on an SDS gel.

Gel Electrophoresis

[0390] If the sample has been stored o/n or longer, put in a heat block prewarmed to 90° C. for 5-10 mins. Set up SDS page gels, use 1 gel per 12 samples, prepare running buffer (1×, Invitrogen NuPAGE MOPS SDS Running Buffer (20×) (NP0001))≈800 ml/gel tank. Add 500 μ l of NuPAGE antioxidant to the upper buffer chamber. Load 15 μ l samples onto gel lanes from left to right as and load 2.5 μ l of Invitrogen Magic Marker XP and 5 μ l Invitrogen See Blue Plus 2 pre-stained standard and 15 μ l of non-treated control. It is important to maximize the resolution of separation during SDS_PAGE. This can be achieved by running 12% bis-tris gels at 200 V for 1 hour and 25 minutes (until the pink (17 kDa) marker reaches the bottom of the tank).

Western Blotting

[0391] Complete a Semi-dry transfer: using an Invitrogen iBlot (use iBlot Programme 3 for 6 minutes). Put the nitrocellulose membranes in individual small trays. Incubate the membranes with blocking buffer solution (5 g Marvel milk powder per 100 ml 0.1% PBS/Tween) at room temperature, on a rocker, for 1 hour. Apply primary antibody (Anti-SNAP-25 1:1000 dilution) and incubate the membranes with primary antibody (diluted in blocking buffer) for 1 hour on a rocker at room temperature. Wash the membranes by rinsing 3 times with PBS/Tween (0.1%). Then apply the secondary (Anti-Rabbit-HRP conjugate diluted 1:1000) and incubate the membranes with secondary antibody (diluted in blocking

buffer) at room temperature, on a rocker, for 1 hour. Wash the membranes by rinsing 3 times with PBS/Tween (0.1%), leave membrane a minimum of 20 mins for the last wash. Detect the bound antibody using Syngene: Drain blots of PBS/Tween, mix WestDura reagents 1:1 and add to blots for 5 minutes. Ensure enough solution is added to the membranes to completely cover them. Place membrane in Syngene tray, set up Syngene software for 5 min expose time.

[0392] FIG. 31 clearly shows that LC/A-CPDY-H_N/A conjugates effectively cleave SNAP-25.

Example 34

Construction and Activation of Dynorphin Conjugates

Preparation of a LC/A and H_N/A Backbone Clones

[0393] The following procedure creates the LC and H_N fragments for use as the component backbone for multidomain fusion expression. This example is based on preparation of a serotype A based clone (SEQ ID1 and SEQ ID2), though the procedures and methods are equally applicable to the other serotypes [illustrated by the sequence listing for serotype B (SEQ ID3 and SEQ ID4) and serotype C (SEQ ID5 and SEQ ID6)].

Preparation of Cloning and Expression Vectors

[0394] pCR 4 (Invitrogen) is the chosen standard cloning vector, selected due to the lack of restriction sequences within the vector and adjacent sequencing primer sites for easy construct confirmation. The expression vector is based on the pMAL (NEB) expression vector, which has the desired restriction sequences within the multiple cloning site in the correct orientation for construct insertion (BamHI-SalI-PstI-HindIII). A fragment of the expression vector has been removed to create a non-mobilisable plasmid and a variety of different fusion tags have been inserted to increase purification options.

Preparation of Protease (e.g. LC/A) Insert

[0395] The LC/A (SEQ ID1) is created by one of two ways:

[0396] The DNA sequence is designed by back translation of the LC/A amino acid sequence [obtained from freely available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO)] using one of a variety of reverse translation software tools (for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)). BamHI/SalI recognition sequences are incorporated at the 5' and 3' ends respectively of the sequence, maintaining the correct reading frame. The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence containing the LC/A open reading frame (ORF) is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0397] The alternative method is to use PCR amplification from an existing DNA sequence with BamHI and SalI restriction enzyme sequences incorporated into the 5' and 3' PCR primers respectively. Complementary oligonucleotide primers are chemically synthesised by a supplier (for example MWG or Sigma-Genosys), so that each pair has the ability to hybridize to the opposite strands (3' ends pointing "towards" each other) flanking the stretch of *Clostridium* target DNA, one oligonucleotide for each of the two DNA strands. To generate a PCR product the pair of short oligonucleotide primers specific for the *Clostridium* DNA sequence are mixed with the *Clostridium* DNA template and other reaction components and placed in a machine (the 'PCR machine') that can change the incubation temperature of the reaction tube automatically, cycling between approximately 94° C. (for denaturation), 55° C. (for oligonucleotide annealing), and 72° C. (for synthesis). Other reagents required for amplification of a PCR product include a DNA polymerase (such as Taq or Pfu polymerase), each of the four nucleotide dNTP building blocks of DNA in equimolar amounts (50-200 μM) and a buffer appropriate for the enzyme optimised for Mg²⁺ concentration (0.5-5 mM).

[0398] The amplification product is cloned into pCR 4 using either, TOPO TA cloning for Taq PCR products or Zero Blunt TOPO cloning for Pfu PCR products (both kits commercially available from Invitrogen). The resultant clone is checked by sequencing. Any additional restriction sequences which are not compatible with the cloning system are then removed using site directed mutagenesis [for example, using Quickchange (Stratagene Inc.)].

Preparation of Translocation (e.g. H_N) Insert

[0399] The H_N/A (SEQ ID2) is created by one of two ways:

[0400] The DNA sequence is designed by back translation of the H_N/A amino acid sequence [obtained from freely available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO)] using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)]. A PstI restriction sequence added to the N-terminus and XbaI-stop codon-HindIII to the C-terminus ensuring the correct reading frame is maintained. The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0401] The alternative method is to use PCR amplification from an existing DNA sequence with PstI and XbaI-stop codon-HindIII restriction enzyme sequences incorporated into the 5' and 3' PCR primers respectively. The PCR amplification is performed as described above. The PCR product is inserted into pCR 4 vector and checked by sequencing. Any additional restriction sequences which are not compatible

with the cloning system are then removed using site directed mutagenesis [for example using Quickchange (Stratagene Inc.)].

Preparation of Linker-Dynorphin-Spacer Insert

[0402] The LC- H_N linker can be designed from first principle, using the existing sequence information for the linker as the template. For example, the serotype A linker (in this case defined as the inter-domain polypeptide region that exists between the cysteines of the disulphide bridge between LC and H_N) is 23 amino acids long and has the sequence VRGIITSKTKSLDKGYNKALNDL. Within this sequence, it is understood that proteolytic activation in nature leads to an H_N domain that has an N-terminus of the sequence ALNDL. This sequence information is freely available from available database sources such as GenBank (accession number P10845) or Swissprot (accession locus BXA1_CLOBO). Into this linker an enterokinase site, dynorphin and spacer are incorporated; and using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)], the DNA sequence encoding the linker-ligand-spacer region is determined. Restriction sites are then incorporated into the DNA sequence and can be arranged as BamH I-Sall-linker-protease site-dynorphin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII. It is important to ensure the correct reading frame is maintained for the spacer, dynorphin and restriction sequences and that the XbaI sequence is not preceded by the bases, TC, which would result on DAM methylation. The DNA sequence is screened for restriction sequence incorporation, and any additional sequences are removed manually from the remaining sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example, GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

Preparation of the LC/A-dynorphin- H_N /A Fusion

[0403] In order to create the LC-linker-dynorphin-spacer- H_N construct (SEQ ID90), the pCR 4 vector encoding the linker is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient vector for insertion and ligation of the LC/A DNA (SEQ ID1) cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with PstI+XbaI restriction enzymes and serves as the recipient vector for the insertion and ligation of the H_N /A DNA (SEQ ID2) cleaved with PstI+XbaI. The final construct contains the LC-linker-dynorphin-spacer- H_N ORF (SEQ ID90) for transfer into expression vectors for expression to result in a fusion protein of the sequence illustrated in SEQ ID91.

Examples 35

Preparation and Purification of an LC/A-dynorphin- H_N /A Fusion Protein Family with Variable Spacer Length

[0404] Using the same strategy as employed in Example 34, a range of DNA linkers were prepared that encoded

dynorphin and variable spacer content. Using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)], the DNA sequence encoding the linker-ligand-spacer region is determined. Restriction sites are then incorporated into the DNA sequence and can be arranged as BamHI-Sall-linker-protease site-dynorphin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII. It is important to ensure the correct reading frame is maintained for the spacer, dynorphin and restriction sequences and that the XbaI sequence is not preceded by the bases, TC which would result on DAM methylation. The DNA sequence is screened for restriction sequence incorporation and any additional sequences are removed manually from the remaining sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector.

[0405] The spacers that were created included:

Code	Protein sequence of the linker	SEQ ID of the linker DNA
GS10	ALAGGGGSALVLQ	92
GS15	ALAGGGSGGGGSALVLQ	93
GS25	ALAGGGSGGGSGGGSGGGGSALVLQ	94

[0406] By way of example, in order to create the LC/A-CPDY(GS25)- H_N /A fusion construct (SEQ ID94), the pCR 4 vector encoding the linker is cleaved with BamHI+Sall restriction enzymes. This cleaved vector then serves as the recipient vector for insertion and ligation of the LC/A DNA (SEQ ID1) also cleaved with BamHI+Sall. The resulting plasmid DNA is then cleaved with BamHI+HindIII restriction enzymes and the LC/A-linker fragment inserted into a similarly cleaved vector containing a unique multiple cloning site for BamHI, Sall, PstI, and HindIII such as the pMAL vector (NEB). The H_N /A DNA (SEQ ID2) is then cleaved with PstI+HindIII restriction enzymes and inserted into the similarly cleaved pMAL-LC/A-linker construct. The final construct contains the LC/A-CPDY(GS25)- H_N /A ORF for expression as a protein of the sequence illustrated in SEQ ID94.

Example 36

Purification Method for LC/A-Dynorphin- H_N /A Fusion Protein

[0407] Defrost falcon tube containing 25 ml 50 mM HEPES pH 7.2, 200 mM NaCl and approximately 10 g of *E. coli* BL21 cell paste. Make the thawed cell paste up to 80 ml with 50 mM HEPES pH 7.2, 200 mM NaCl and sonicate on ice 30 seconds on, 30 seconds off for 10 cycles at a power of 22 microns ensuring the sample remains cool. Spin the lysed cells at 18 000 rpm, 4° C. for 30 minutes. Load the supernatant onto a 0.1 M NiSO₄ charged Chelating column (20-30 ml column is sufficient) equilibrated with 50 mM HEPES pH

7.2, 200 mM NaCl. Using a step gradient of 10 and 40 mM imidazol, wash away the non-specific bound protein and elute the fusion protein with 100 mM imidazol. Dialyse the eluted fusion protein against 5 L of 50 mM HEPES pH 7.2, 200 mM NaCl at 4° C. overnight and measure the OD of the dialysed fusion protein. Add 3.2 µl of enterokinase (2 µg/ml) per 1 mg fusion protein and incubate at 25° C. static overnight. Load onto a 0.1 M NiSO₄ charged Chelating column (20-30 ml column is sufficient) equilibrated with 50 mM HEPES pH 7.2, 200 mM NaCl. Wash column to baseline with 50 mM HEPES pH 7.2, 200 mM NaCl. Using a step gradient of 10 and 40 mM imidazol, wash away the non-specific bound protein and elute the fusion protein with 100 mM imidazol. Dialyse the eluted fusion protein against 5 L of 50 mM HEPES pH 7.2, 200 mM NaCl at 4° C. overnight and concentrate the fusion to about 2 mg/ml, aliquot sample and freeze at -20° C. Test purified protein using OD, BCA, purity analysis and SNAP-25 assessments.

Example 37

Preparation of a LC/C-dynorphin-H_N/C Fusion Protein with a Serotype A Activation Sequence

[0408] Following the methods used in Examples 1 and 2, the LC/C (SEQ ID5) and H_N/C (SEQ ID6) are created and inserted into the A serotype linker arranged as BamHI-Sall-linker-protease site-dynorphin-NheI-spacer-SpeI-PstI-XbaI-stop codon-HindIII. The final construct contains the LC-linker-dynorphin-spacer-H_N ORF for expression as a protein of the sequence illustrated in SEQ ID95.

Example 38

Preparation of an IgA Protease-Dynorphin Variant-H_N/A Fusion Protein

[0409] The IgA protease amino acid sequence was obtained from freely available database sources such as GenBank (accession number P09790). Information regarding the structure of the *N. Gonorrhoeae* IgA protease gene is available in the literature (Pohlner et al., Gene structure and extracellular secretion of *Neisseria gonorrhoeae* IgA protease, Nature, 1987, 325(6103), 458-62). Using Backtranslation tool v2.0 (Entelechon), the DNA sequence encoding the IgA protease modified for *E. coli* expression was determined. A BamHI recognition sequence was incorporated at the 5' end and a codon encoding a cysteine amino acid and Sall recognition sequence were incorporated at the 3' end of the IgA DNA. The DNA sequence was screened using MapDraw, (DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required for cloning were removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage was assessed Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables. This optimised DNA sequence (SEQ ID74) containing the IgA open reading frame (ORF) is then commercially synthesized.

[0410] The IgA (SEQ ID74) is inserted into the LC-linker-dynorphin-spacer-H_N ORF (SEQ ID90) using BamHI and Sall restriction enzymes to replace the LC with the IgA protease DNA. The final construct contains the IgA-linker-

dynorphin-spacer-H_N ORF for expression as a protein of the sequence illustrated in SEQ ID96.

Example 39

Preparation of a Dynorphin Targeted Endopeptidase Fusion Protein Containing a LC Domain Derived from Tetanus Toxin

[0411] The DNA sequence is designed by back translation of the tetanus toxin LC amino acid sequence (obtained from freely available database sources such as GenBank (accession number X04436) using one of a variety of reverse translation software tools [for example EditSeq best *E. coli* reverse translation (DNASTAR Inc.), or Backtranslation tool v2.0 (Entelechon)]. BamHI/Sall recognition sequences are incorporated at the 5' and 3' ends respectively of the sequence maintaining the correct reading frame (SEQ ID83). The DNA sequence is screened (using software such as MapDraw, DNASTAR Inc.) for restriction enzyme cleavage sequences incorporated during the back translation. Any cleavage sequences that are found to be common to those required by the cloning system are removed manually from the proposed coding sequence ensuring common *E. coli* codon usage is maintained. *E. coli* codon usage is assessed by reference to software programs such as Graphical Codon Usage Analyser (Geneart), and the % GC content and codon usage ratio assessed by reference to published codon usage tables (for example GenBank Release 143, 13 Sep. 2004). This optimised DNA sequence containing the tetanus toxin LC open reading frame (ORF) is then commercially synthesized (for example by Entelechon, Geneart or Sigma-Genosys) and is provided in the pCR 4 vector (Invitrogen). The pCR 4 vector encoding the TeNT LC is cleaved with BamHI and Sall. The BamHI-Sall fragment is then inserted into the LC/A-dynorphin-H_N/A vector (SEQ ID90) that has also been cleaved by BamHI and Sall. The final construct contains the TeNT LC-linker-dynorphin-spacer-H_N ORF sequences for expression as a protein of the sequence illustrated in SEQ ID97.

Example 40

[0412] A method of treating, preventing or ameliorating pain in a subject, comprising administration to said patient a therapeutic effective amount of fusion protein, wherein said pain is selected from the group consisting of: chronic pain arising from malignant disease, chronic pain not caused by malignant disease (peripheral neuropathies).

Patient A

[0413] A 73 year old woman suffering from severe pain caused by posthepatic neuralgia is treated by a peripheral injection with fusion protein to reduce neurotransmitter release at the synapse of nerve terminals to reduce the pain. The patient experiences good analgesic effect within 2 hours of said injection.

Patient B

[0414] A 32 year old male suffering from phantom limb pain after having his left arm amputated following a car accident is treated by peripheral injection with fusion protein

to reduce the pain. The patient experiences good analgesic effect within 1 hour of said injection.

Patient C

[0415] A 55 year male suffering from diabetic neuropathy is treated by a peripheral injection with fusion protein to reduce neurotransmitter release at the synapse of nerve terminals to reduce the pain. The patient experiences good analgesic effect within 4 hours of said injection.

Patient D

[0416] A 63 year old woman suffering from cancer pain is treated by a peripheral injection with fusion protein to reduce

neurotransmitter release at the synapse of nerve terminals to reduce the pain. The patient experiences good analgesic effect within 4 hours of said injection.

[0417] All documents, books, manuals, papers, patents, published patent applications, guides, abstracts and other reference materials cited herein are incorporated by reference in their entirety. While the foregoing specification teaches the principles of the present invention, with examples provided for the purpose of illustration, it will be appreciated by one skilled in the art from reading this disclosure that various changes in form and detail can be made without departing from the true scope of the invention.

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 97

<210> SEQ ID NO 1

<211> LENGTH: 1302

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 1

```

ggatccatgg agttcgtaa caaacagttc aactataaag acccagttaa cgggtgtgac      60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc      120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccggga agaagggtgac      180
ctgaaccgcg caccggaagc gaaacagggt cgggtatctt actatgactc cacctacctg      240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactgtt cgagcgtatt      300
tactccacgg acctgggccc tatgctgctg actagcatcg ttgcgggtat cccgttctgg      360
ggcggttcta ccacgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt      420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct      480
gatatcatcc agttcgagtg taagagcttt ggtaacgaag ttctgaacct caccgtaac      540
ggctacgggt ccaactcagta catccgtttc tctccggact tcaccttcgg ttttgaagaa      600
tccttggaag tagacacgaa cccactgctg gccgctggta aattcgcaac tgatcctgcg      660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat      720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt      780
agcttcgaag aactgcgtac ttttggcggc cagcagccta aattcatcga ctctctgcaa      840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac      900
aaagcgaaat ccacgtggg taccactgct tctctccagt acatgaagaa cgttttttaa      960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc     1020
gataaaactt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt     1080
aaagttctga accgcaaac ctatctgaac ttgacaagg cagtattcaa aatcaacatc     1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaaac caacctggct     1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaaac     1260
ttcactggtc tgttcagatt ttacaagctg ctgtgcgtcg ac                          1302

```

-continued

<210> SEQ ID NO 2
<211> LENGTH: 1257
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 2

```
ctgcagtgta tcaagggttaa caactgggat ttattcttca gcccgagtga agacaacttc      60
accaacgacc tgaacaaagg tgaagaaatc acctcagata ctaacatcga agcagccgaa     120
gaaaacatct cgctggacct gatccagcag tactacctga cctttaattt cgacaacgag     180
ccggaaaaca ttctatcga aaacctgagc tctgatatca tcggccagct ggaactgatg     240
ccgaacatcg aacgtttccc aaacggtaaa aagtacgagc tggacaaata taccatgttc     300
cactacctgc gcgcgcagga atttgaacac ggcaaatccc gtatcgcaact gactaactcc     360
gttaacgaag ctctgctcaa cccgtcccgt gtatacacct tcttctctag cgactacgtg     420
aaaaaggtca acaaagcgac tgaagctgca atgttcttgg gttgggttga acagcttggt     480
tatgatttta ccgacgagac gtccgaagta tctactaccg acaaaattgc ggatatcact     540
atcatcatcc cgtacatcgg tccggctctg aacattggca acatgctgta caaagacgac     600
ttcgttggcg cactgatctt ctccggtgcg gtgacctgc tggagttcat cccggaaatc     660
gccatcccg tactgggac ctttgctctg gtttcttaca ttgcaaaca ggttctgact     720
gtacaaacca tcgacaacgc gctgagcaaa cgtaacgaaa aatgggatga agtttacaaa     780
tatatcgtga ccaactggct ggctaagggt aatactcaga tcgacctcat ccgcaaaaaa     840
atgaaagaag cactgaaaaa ccaggcggaa gctaccaagg caatcattaa ctaccagtac     900
aaccagtaca ccgaggaaga aaaaaacaac atcaacttca acatcgacga tctgtcctct     960
aaactgaacg aatccatcaa caaagctatg atcaacatca acaagttcct gaaccagtgc    1020
tctgtaagct atctgatgaa ctccatgatc ccgtacggtg ttaaacgtct ggaggacttc    1080
gatgctgttc tgaagacgc cctgtgaaa tacatttacg acaaccgtgg cactctgac     1140
ggtcagggtg atcgtctgaa ggacaaaagt aacaatacct tatcgaccga catccctttt    1200
cagctcagta aatatgtcga taaccaacgc cttttgtcca ctctagacta gaagctt      1257
```

<210> SEQ ID NO 3
<211> LENGTH: 1323
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 3

```
ggatccatgc cggttaccat caacaacttc aactacaacg acccgatcga caacaacaac      60
atcattatga tggaaaccgc gttcgcacgt ggtaccggac gttactacaa ggcttttaag     120
atcaccgacc gtatctggat catcccggaa cgttacacct tcggttacaa acctgaggac     180
ttcaacaaga gtagcgggat ttccaatcgt gacgtctgcg agtactatga tccagattat     240
ctgaatacca acgataagaa gaacatatc cttcagacta tgattaaact cttcaaccgt     300
atcaaaagca aaccgctcgg tgaaaaactc ctcgaaatga ttatcaacgg tatcccgtac     360
ctcggtgacc gtcgtgtccc gcttgaagag ttcaacacca acatcgcaag cgtcaccgtc     420
aacaactca tcagcaaccc aggtgaagtc gaacgtaaaa aaggtatctt cgcaaacctc     480
```

-continued

```

atcatcttcg gtccgggtcc ggtcctcaac gaaaacgaaa ccatcgacat cggatatccag 540
aaccacttcg caagccgtga aggtttcggg ggtatcatgc agatgaaatt ctgcccgga 600
tacgtcagtg ttttcaacaa cgtccaggaa aacaaagggtg caagcatctt caaccgtcgt 660
ggttacttca gcgacccggc actcatctc atgcatgaac tcatccacgt cctccacggg 720
ctctacggta tcaaagttag cgacctcccg atcgtcccg aagagaagaa attcttcatg 780
cagagcaccg acgcaatcca ggctgaggaa ctctacacct tcggtggcca agaccaagt 840
atcatacccc cgtccaccga caaaagcatc tacgacaaag tcctccagaa cttcaggggt 900
atcgtggaca gactcaacaa agtcctcgtc tgcacagcg acccgaaat caatatcaac 960
atatacaaga acaagttcaa agacaagtac aaattcgtcg aggacagcga aggcaaatac 1020
agcatcgacg tagaaagttt cgacaagctc taaaaagcc tcatgttcgg tttcaccgaa 1080
accaacatcg ccgagaacta caagatcaag acaagggcaa gttacttcag cgacagctc 1140
ccgcctgtca aaatcaagaa cctcttagac aacgagattt acacaattga agagggttc 1200
aacatcagtg acaaagacat ggagaaggaa tacagaggtc agaacaaggc tatcaacaaa 1260
caggcatacg aggagatcag caaagaacac ctgcgagtct acaagatcca gatgtgcgtc 1320
gac 1323

```

```

<210> SEQ ID NO 4
<211> LENGTH: 1260
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 4

```

```

ctgcagtgca tcgacgttga caacgaagac ctgttcttca tcgctgacaa aaacagcttc 60
agtgacgacc tgagcaaaaa cgaacgtatc gaatacaaca cccagagcaa ctacatcgaa 120
aacgacttcc cgatcaacga actgatcctg gacaccgacc tgataagtaa aatcgaaactg 180
ccgagcgaaa acaccgaaaag tctgaccgac ttcaacgttg acgttccggg ttacgaaaaa 240
cagccgggcta tcaagaaaat cttcaccgac gaaaacacca tcttcagta cctgtacagc 300
cagaccttcc cgctggacat ccgtgacatc agtctgacca gcagtttcga cgacgctctg 360
ctgttcagca acaaagttta cagtttcttc agcatggact acatcaaaac cgctaacaaa 420
gttgttgaag cagggtgtgt cgctggttgg gttaaacaga tcgttaacga cttcgttatc 480
gaagctaaca aaagcaacac tatggacaaa atcgctgaca tcagtctgat cgttccgtac 540
atcgggtctgg ctctgaacgt tggtaacgaa accgctaaag gtaactttga aaacgctttc 600
gagatcgctg gtgcaagcat cctgctggag ttcaccccg aactgctgat cccggttgtt 660
ggtgctttcc tgctggaaaag ttacatcgac aacaaaaaca agatcatcaa aaccatcgac 720
aacgctctga ccaaacgtta cgaaaaatgg agtgatatgt acggtctgat cgttgctcag 780
tggtgagca ccgtcaacac ccagttctac accatcaaag aaggtatgta caaagctctg 840
aactaccagg ctcaggctct ggaagagatc atcaataacc gttacaacat ctacagttag 900
aaggaaaaga gtaacatcaa catcgacttc aacgacatca acgcaaaact gaacgaaggt 960
atcaaccagg ctatcgacaa catcaacaac ttcacaaag gttgcagtg tagctacctg 1020
atgaagaaga tgatcccgtc ggctgttgaa aaactgctgg acttcgacaa caccctgaaa 1080

```

-continued

```

aagaacctgc tgaactacat cgacgaaaac aagctgtacc tgatcggtag tgctgaatac 1140
gaaaaaagta aagtgaacaa atacctgaag accatcatgc cgttcgacct gagtatctac 1200
accaacgaca ccacctgat cgaaatgttc aacaaatata actctctaga ctagaagctt 1260

```

```

<210> SEQ ID NO 5
<211> LENGTH: 1329
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 5

```

```

ggatccgaat tcatgccgat caccatcaac aacttcaact acagcgatcc ggtggataac 60
aaaaacatcc tgtacctgga taccatctg aataccctgg cgaacgaacc ggaaaaagcg 120
tttcgtatca ccggcaacat ttgggttatt ccggatcggt tttagcgtaa cagcaaccgg 180
aatctgaata aaccgcgcgc tgttaccagc ccgaaaagcg gttattacga tccgaactat 240
ctgagcaccg atagcgataa agataccttc ctgaaagaaa tcatacaact gttcaaacgc 300
atcaacagcc gtgaaattgg cgaagaactg atctatcgcc tgagcaccga tattccggtt 360
ccgggcaaca acaacacccc gatcaacacc tttgatttcg atgtggattt caacagcggt 420
gatgttaaaa ccgcgccagg taacaattgg gtgaaaacgg gcagcattaa ccgagcggtg 480
attattaccg gtccgcgcga aaacattatt gatccggaaa ccagcacctt taaactgacc 540
aacaacacct ttgcggcgca ggaaggtttt ggcgcgctga gcattattag cattagcccg 600
cgctttatgc tgacctatag caacgcgacc aacgatgttg gtgaaggccg ttccagcaaa 660
agcgaatttt gcatggaccc gatcctgac ctgatgcatg aactgaacca tgcgatgcat 720
aacctgtatg gcatcgcgat tccgaacgat cagaccatta gcagcgtgac cagcaacatc 780
ttttacagcc agtacaacgt gaaactggaa tatgcggaaa tctatgcgtt tggcgggtccg 840
accattgac tgattccgaa aagcgcgcgc aaatacttcg aagaaaaagc gctggattac 900
tatcgcagca ttgcgaaacg tctgaacagc attaccaccg cgaatccgag cagcttcaac 960
aaatatatcg gcgaatataa acagaaaactg atccgcaaat atcgctttgt ggtggaaaagc 1020
agcggcgaag ttaccgttaa ccgcaataaa ttcgtggaac tgtacaacga actgaccag 1080
atcttcaccg aatttaacta tgcgaaaatc tataacgtgc agaaccgtaa aatctacctg 1140
agcaacgtgt ataccccggt gaccgcgaat attctggatg ataacgtgta cgatatccag 1200
aacggcttta acatcccgaa aagcaacctg aacgttctgt ttatgggcca gaacctgagc 1260
cgtaatccgg cgctgcgtaa agtgaaccgg gaaaacatgc tgtacctgtt caccaaattt 1320
tgcgtcgac 1329

```

```

<210> SEQ ID NO 6
<211> LENGTH: 1263
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 6

```

```

ctgcagtgtc gtgaactgct ggtgaaaaac accgatctgc cgtttattgg cgatatcagc 60
gatgtgaaaa ccgatatctt cctgcgcaaa gatatcaacg aagaaaccga agtgatctac 120

```

-continued

taccgcgata acgtgagcgt tgatcagggtg atcctgagca aaaacaccag cgaacatggt	180
cagctggatc tgctgtatcc gagcattgat agcgaaagcg aaattctgcc gggcgaaaac	240
cagggtgtttt acgataaccg taccagaaac gtggattacc tgaacageta ttactacctg	300
gaaagccaga aactgagcga taacgtggaa gattttacct ttaccgcgag cattgaagaa	360
gcgctggata acagcgcgaa agtttacacc tttttccga cctggcgaa caaagttaat	420
gcgggtgttc agggcggtct gtttctgatg tggcgaaacg atgtggtgga agatttcacc	480
accaacatcc tgcgtaaaaga taccctggat aaaatcagcg atgttagcgc gattattccg	540
tatatgtgctc cggcgctgaa cattagcaat agcgtgcgtc gtggcaatct taccgaagcg	600
tttgcggtta ccggtgtgac cattctgctg gaagcgtttc cggaatttac cattccggcg	660
ctgggtgcgt ttgtgatcta tagcaaatg caggaacgca acgaaatcat caaaaccatc	720
gataactgcc tggaacagcg tattaacgc tggaagata gctatgaatg gatgatgggc	780
acctggctga gccgtattat caccagttc aacaacatca gctaccagat gtacgatagc	840
ctgaactatc aggcgggtgc gattaagcg aaaatcgatc tggaatacaa aaaatacagc	900
ggcagcgata aagaaaacat caaaagccag gttgaaaacc tgaaaaacag cctggatgtg	960
aaaattagcg aagcgatgaa taacatcaac aaattcatcc gcgaatgcag cgtgacctac	1020
ctgttcaaaa acatgctgcc gaaagtatc gatgaactga acgaatttga tcgcaacacc	1080
aaagcgaaac tgatcaacct gatcgatagc cacaacatta ttctggtggg cgaagtggat	1140
aaactgaaag cgaaagttaa caacagcttc cagaacacca tcccgtttaa catcttcagc	1200
tataccaaca acagcctgct gaaagatatc atcaacgaat acttcaatct agactagaag	1260
ctt	1263

<210> SEQ ID NO 7
 <211> LENGTH: 207
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 7

ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat cgaaggtcgt	60
tttggcggtt tcacggcgcg acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg	120
ggcggtggcg gtagcgcgcg tggcggtagc ggcggtggcg gtagcgcaact agtgctgcag	180
acgcacggtc tagaatgata aaagctt	207

<210> SEQ ID NO 8
 <211> LENGTH: 108
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 8

ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga	60
aacaaagcgc tgaacctgca gacgcacggt ctagaatgat aaaagctt	108

<210> SEQ ID NO 9
 <211> LENGTH: 186

-continued

```

<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 9
catatgaata acctcgggat tgagggtcgt tttggcgggt tcacggggcgc acgcaaatca      60
gcgcgtaaat tagctaacca gactagtggc ggtgggggta gtggcgggtg cggttcgggc      120
gggggtggga gccctagggg atccgtcgac ctgcagggtc tagaagcgct agcgtgataa      180
aagctt                                          186

<210> SEQ ID NO 10
<211> LENGTH: 180
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 10
ggatccacgc acgtcgacgc gattgatggt cgttttggcg gtttcacggg cgcacgcaaa      60
tcagcgcgta aattagctaa ccaggcgcta gcgggcgggt gcggtagcgg cggtggcgggt      120
agcggcggtg gcggtagcgc actagtgtcg cagacgcacg gtctagaatg ataaaagctt      180

<210> SEQ ID NO 11
<211> LENGTH: 249
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 11
ggatccacgc acgtcgacgc catcattacc tccaaaacta aatctctgat cgaaggtcgt      60
tacggtggtt tcatgacctc tgaaaaatct cagacccgcg tggttaccct gttcaaaaac      120
gctatcatca aaaacgctta caaaaaaggt gaagcgctag cgggtggtgg tggttctggt      180
ggtggtggtt ctggtggtgg tggttctgca ctagtgtgc agacgcacgg tctagaatga      240
taaaagctt                                          249

<210> SEQ ID NO 12
<211> LENGTH: 207
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 12
ggatccacgc acgtcgacgc catcattacc tccaaaacta aatctctgat cgaaggtcgt      60
tttggcgggt tcacggggcgc acgcaaatca gcgcgtaaac gtaagaacca ggcgctagcg      120
ggcgggtggcg gtagcggcgg tggcggtagc ggcgggtggcg gtagcgcact agtgcgcag      180
acgcacggtc tagaatgata aaagctt                                          207

<210> SEQ ID NO 13
<211> LENGTH: 2709
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

-continued

<400> SEQUENCE: 13

ggatccatgg agttcggttaa caaacagttc aactataaag acccagttaa cgggtgttgac	60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc	120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac	180
ctgaaccgcg caccggaagc gaaacagggt cgggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaagggtgta ctaaaactgtt cgagcgtatt	300
tactccaccg acctggggcg tatgtgctg actagcatcg ttccgggtat cccgttctgg	360
ggcggttcta ccatcgatag cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct	480
gatatcatcc agttcgatg taagagcttt ggtcacgaag ttctgaacct cacccgtaac	540
ggctacgggt ccaactcagta catccgttcc tctccggact tcaccttcgg ttttgaagaa	600
tccttggaag tagacacgaa cccactgtg ggcgtggta aatcgcaac tgatcctgcg	660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttgccggt cacgacgcta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccatcggtgg taccactgct tctctccagt acatgaagaa cgtttttaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagtctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgcgaaaag ttaactacac tatctacgat gggttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggtc tggtcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatttggc gggttcacgg gcgcacgcaa atcagcgcg	1380
aaattagcta accagggcgt agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt	1440
ggcggtagcg cactagtgtc gcagtgtatc aaggtttaaca actgggattt attcttcagc	1500
ccgagtgaag acaacttcac caacgacctg aacaaagggt aagaaatcac ctacagatact	1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc	1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc	1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg	1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaaat ttgaacacgg caaatcccg	1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccggtg atacaccttc	1860
ttctctagcg actacgtgaa aaagggtcaac aaagcgactg aagctgcaat gttcttggt	1920
tgggttgaa agcttggtta tgattttacc gacgagacgt ccgaagtatc tactaccgac	1980
aaaaattcgg atatcactat catcatcccg tacatcggtc cggctctgaa cattggcaac	2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct cgggtgcggt gatcctgctg	2100
gagttcatcc cggaaatcgc catccggta ctgggcacct ttgctctggt ttcttacatt	2160
gcaacaaggt ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa	2220

-continued

```

tgggatgaag ttacaaata tatcgtgacc aactggctgg ctaagggttaa tactcagatc 2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgacgac tgctctctaa actgaacgaa tccatcaaca aagctatgat caacatcaac 2460
aagttcctga accagtgtct tgtaagctat ctgatgaact ccatgatccc gtacgggtgtt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgetgaaata catttacgac 2580
aaccgtggca ctctgatcgg tcaggttgat cgtctgaagg acaaagtga caatacctta 2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct tttgtccact 2700
ctagactag 2709

```

```

<210> SEQ ID NO 14
<211> LENGTH: 902
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 14

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1          5          10          15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20         25         30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35         40         45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50         55         60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65         70         75         80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85         90         95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100        105        110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115        120        125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130        135        140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145        150        155        160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165        170        175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180        185        190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195        200        205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210        215        220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225        230        235        240
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245        250        255

```

-continued

Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260						265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
		275					280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290					295					300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325						330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
			340					345					350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
	355						360					365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405						410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
	435						440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465				470					475						480
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp
			485					490						495	
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys
		500						505					510		
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn
	515						520					525			
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp
	530					535					540				
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile
545					550					555					560
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys
			565						570					575	
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln
		580						585					590		
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn
	595						600					605			
Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp
	610					615					620				
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly
625					630					635					640
Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val
			645						650					655	

-continued

Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile
660 665 670

Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val
675 680 685

Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro
690 695 700

Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile
705 710 715 720

Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu Ser Lys
725 730 735

Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp
740 745 750

Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg Lys Lys Met Lys
755 760 765

Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr
770 775 780

Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn
785 790 795 800

Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met
805 810 815

Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met
820 825 830

Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala
835 840 845

Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr
850 855 860

Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu
865 870 875 880

Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg
885 890 895

Leu Leu Ser Thr Leu Asp
900

<210> SEQ ID NO 15

<211> LENGTH: 2736

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 15

```

ctcgggattg agggtcgttt tggcggtttc acgggcgcac gcaaatcagc gcgtaaatta    60
gctaaccaga ctagtggcgg tgggggtagt ggcggtggcg gttcgggcgg ggggtgggagc    120
cctaggggat ccatggagtt cgtaacaaa cagttcaact ataaagaccc agttaacggt    180
gttgacattg cttacatcaa aatccgaac gctggccaga tgcagccggt aaaggcattc    240
aaaaatccaca aaaaatctg gggtatcccg gaacgtgata cctttactaa cccggaagaa    300
ggtgacctga acccgccacc ggaagcgaaa caggtgccgg tatcttacta tgactccacc    360
tacctgtcta ccgataacga aaaggacaac tacctgaaag gtgttactaa actgttcgag    420
cgtatttact ccaccgacct gggccgtatg ctgctgacta gcacgttcg cggtatcccg    480
ttctggggcg gttctacat cgataccgaa ctgaaagtaa tcgacactaa ctgcatcaac    540

```

-continued

gttattcagc	cggaacggtc	ctatcggttc	gaagaactga	acctggtgat	catcggcccg	600
tctgctgata	tcattccagtt	cgagtgttaag	agctttggtc	acgaagttct	gaacctcacc	660
cgtaacggct	acgggtccac	tcagtacatc	cgtttctctc	cggacttcac	cttcgggtttt	720
gaagaatccc	tggaagtaga	cacgaaccca	ctgctgggcg	ctggtaaatt	cgcaactgat	780
cctgcggtta	ccctggctca	cgaactgatt	catgcaggcc	accgcctgta	cggtatcgcc	840
atcaatccga	accgtgtctt	caaagttaac	accaacgcgt	attacgagat	gtccgggtctg	900
gaagttagct	tcgaagaact	gcgtactttt	ggcggtcacg	acgctaaatt	catcgactct	960
ctgcaagaaa	acgagttccg	tctgtactac	tataacaagt	tcaaagatat	cgcatccacc	1020
ctgaacaaa	cgaaatccat	cgtgggtacc	actgcttctc	tccagtacat	gaagaacggt	1080
tttaagaaa	aatacctgct	cagcgaagac	acctccggca	aattctctgt	agacaagttg	1140
aaattcgata	aactttacaa	aatgctgact	gaaatttaca	ccgaagacaa	cttcggttaag	1200
ttctttaaag	ttctgaacgg	caaaacctat	ctgaacttcg	acaaggcagt	attcaaaatc	1260
aacatcgtgc	cgaaagttaa	ctacactatc	tacgatgggt	tcaacctgcg	taacaccaac	1320
ctggctgcta	attttaacgg	ccagaacacg	gaaatcaaca	acatgaactt	cacaaaactg	1380
aaaaacttca	ctggctgtgt	cgagttttac	aagctgctgt	gcgtcgacgg	catcattacc	1440
tccaaaacta	aatctctgat	agaaggtaga	aacaaagcgc	tgaacgacct	ctgtatcaag	1500
gttaacaact	gggatttatt	cttcagcccg	agtgaagaca	acttcaccaa	cgacctgaac	1560
aaagggtgaag	aaatcacctc	agatactaac	atcgaagcag	ccgaagaaaa	catctcgctg	1620
gacctgatcc	agcagtacta	cctgaccttt	aatttcgaca	acgagccgga	aaacatttct	1680
atcgaaaacc	tgagctctga	tatcatcgcc	cagctggaac	tgatgccgaa	catcgaaact	1740
ttccaaaacg	gtaaaaagta	cgagctggac	aaatatacca	tgttccacta	cctgcgcgcg	1800
caggaatttg	aacacggcaa	atcccgatc	gcactgacta	actccgttaa	cgaagctctg	1860
ctcaacccgt	cccgtgtata	caccttcttc	tctagcgact	acgtgaaaaa	ggtaacaaaa	1920
gcgactgaag	ctgcaatgtt	cttgggttgg	gttgaacagc	ttgtttatga	ttttaccgac	1980
gagacgtccg	aagtatctac	taccgacaaa	attgcggata	tcactatcat	catcccgtag	2040
atcggtccgg	ctctgaacat	tggcaacatg	ctgtacaaag	acgacttcgt	tggcgactctg	2100
atcttctccg	gtgcggtgat	cctgctggag	ttcatcccg	aaatcgccat	cccggtagctg	2160
ggcacctttg	ctctggtttc	ttacattgca	aacaagggtc	tgactgtaca	aaccatcgac	2220
aacgcgctga	gcaaacgtaa	cgaaaaatgg	gatgaagttt	acaaatatat	cgtgaccaac	2280
tggttggtta	aggttaatac	tcagatcgac	ctcatccgca	aaaaaatgaa	agaagcactg	2340
gaaaaccagg	cggaagctac	caaggcaatc	attaactacc	agtacaacca	gtacaccgag	2400
gaagaaaaaa	acaacatcaa	cttcaacatc	gacgatctgt	cctctaaact	gaacgaatcc	2460
atcaacaaag	ctatgatcaa	catcaacaag	ttcctgaacc	agtgcctctgt	aagctatctg	2520
atgaactcca	tgatcccgta	cgggtgttaa	cgtctggagg	acttcgatgc	gtctctgaaa	2580
gacgcctctg	tgaatacat	ttacgacaac	cgtggcactc	tgatcggtca	ggttgatcgt	2640
ctgaaggaca	aagtgaacaa	taccttatcg	accgacatcc	cttttcagct	cagtaaatat	2700
gtcgataacc	aacgcctttt	gtccactcta	gactag			2736

-continued

```

<211> LENGTH: 911
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 16

Leu Gly Ile Glu Gly Arg Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser
1           5           10           15

Ala Arg Lys Leu Ala Asn Gln Thr Ser Gly Gly Gly Gly Ser Gly Gly
20           25           30

Gly Gly Ser Gly Gly Gly Gly Ser Pro Arg Gly Ser Met Glu Phe Val
35           40           45

Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala
50           55           60

Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro Val Lys Ala Phe
65           70           75           80

Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr
85           90           95

Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Pro Glu Ala Lys Gln Val
100          105          110

Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys
115          120          125

Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser
130          135          140

Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro
145          150          155          160

Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr
165          170          175

Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr Arg Ser Glu Glu
180          185          190

Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu
195          200          205

Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr
210          215          220

Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe
225          230          235          240

Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys
245          250          255

Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala
260          265          270

Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys
275          280          285

Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe
290          295          300

Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser
305          310          315          320

Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr Tyr Asn Lys Phe Lys Asp
325          330          335

Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser Ile Val Gly Thr Thr Ala
340          345          350

Ser Leu Gln Tyr Met Lys Asn Val Phe Lys Glu Lys Tyr Leu Leu Ser
355          360          365

```

-continued

Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	Lys	Leu	Lys	Phe	Asp	Lys	370	375	380	
Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	Glu	Asp	Asn	Phe	Val	Lys	385	390	395	400
Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	Leu	Asn	Phe	Asp	Lys	Ala	405	410	415	
Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	Asn	Tyr	Thr	Ile	Tyr	Asp	420	425	430	
Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	Ala	Asn	Phe	Asn	Gly	Gln	435	440	445	
Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	Lys	Leu	Lys	Asn	Phe	Thr	450	455	460	
Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	Val	Asp	Gly	Ile	Ile	Thr	465	470	475	480
Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	Asn	Lys	Ala	Leu	Asn	Asp	485	490	495	
Leu	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	500	505	510	
Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	515	520	525	
Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	530	535	540	
Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	545	550	555	560
Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	565	570	575	
Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	580	585	590	
Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	595	600	605	
Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	610	615	620	
Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	625	630	635	640
Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	645	650	655	
Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala	660	665	670	
Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Gly	675	680	685	
Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser	Gly	690	695	700	
Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val	Leu	705	710	715	720
Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val	725	730	735	
Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu	740	745	750	
Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr	Gln	755	760	765	

-continued

Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln	Ala
	770					775					780				
Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr	Glu
	785				790					795					800
Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser	Lys
			805						810					815	
Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe	Leu
			820					825					830		
Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr	Gly
	835						840					845			
Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu	Leu
	850					855					860				
Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp	Arg
	865				870					875					880
Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Gln
				885				890						895	
Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp	
			900					905					910		

<210> SEQ ID NO 17

<211> LENGTH: 2715

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 17

ggatccgaat tcatgccgat caccatcaac aacttcaact acagcgatcc ggtggataac	60
aaaaacatcc tgtacctgga taccatctg aataccctgg cgaacgaacc ggaaaaagcg	120
tttcgtatca ccggcaacat ttgggttatt ccggatcggt ttagccgtaa cagcaaccgc	180
aatctgaata aaccgccgcg tgttaccagc ccgaaaagcg gttattacga tccgaactat	240
ctgagcaccg atagcgataa agataccttc ctgaaagaaa tcatcaaaact gttcaaacgc	300
atcaacagcc gtgaaattgg cgaagaactg atctatcgcc tgagcaccga tattccgttt	360
ccgggcaaca acaacacccc gatcaacacc tttgatttcg atgtggattt caacagcggt	420
gatgttaaaa cccgccaggg taacaattgg gtgaaaaccg gcagcattaa cccgagcgtg	480
attattaccg gtccgcgcga aaacattatt gatccggaaa ccagcacctt taaactgacc	540
aaacaacacct ttgcggcgca ggaagggttt ggccgcgtga gcattattag cattagcccg	600
cgctttatgc tgacctatag caacgcgacc aacgatgttg gtgaaggccg ttccagcaaa	660
agcgaatttt gcatggaccc gatcctgac ctgatgcatg aactgaacca tgcgatgcat	720
aacctgtatg gcatcgcat tccgaacgat cagaccatta gcagcgtgac cagcaacatc	780
ttttacagcc agtacaacgt gaaactggaa tatgcggaaa tctatgcgtt tggcgggtccg	840
accattgac tgattccgaa aagcgcgcgc aaatacttcg aagaaaaagc gctggattac	900
tatgcgagca ttgcgaaacg tctgaacagc attaccaccg cgaatccgag cagcttcaac	960
aaatatatcg gcgaataata acagaaactg atccgcaaat atcgctttgt ggtggaaagc	1020
agcggcgaag ttaccgttaa ccgcaataaa ttcgtggaac tgtacaacga actgacccag	1080
atcttcaccg aatttaacta tgcgaaaatc tataacgtgc agaaccgtaa aatctacctg	1140
agcaacgtgt ataccccggt gaccgcgaat attctggatg ataacgtgta cgatatccag	1200

-continued

```

aacggcttta acatcccgaa aagcaacctg aacgttctgt ttatgggcca gaacctgagc 1260
cgtaatccgg cgctgcgtaa agtgaacccg gaaaacatgc tgtacctgtt caccaaattt 1320
tgcgtcgacg cgatagatgg tagatttggc ggtttcacgg gcgcacgcaa atcagcgcgt 1380
aaattagcta accaggcgct agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt 1440
ggcggtagcg cactagtgtc gcagtgtcgt gaactgctgg tgaaaaacac cgatctgccg 1500
tttattggcg atatcagcga tgtgaaaacc gatattctcc tgcgcaaaga tatcaacgaa 1560
gaaaccgaag tgatctacta cccggataac gtgagcgttg atcaggtgat cctgagcaaa 1620
aacaccagcg aacatggtca gctggatctg ctgtatccga gcattgatag cgaagcgaa 1680
attctgccgg gcgaaaacca ggtgttttac gataaccgta cccagaacgt ggattacctg 1740
aacagctatt actacctgga aagccagaaa ctgagcgata acgtggaaga ttttaccttt 1800
acccgcagca ttgaagaagc gctggataac agcgcgaaag tttacaccta tttccgacc 1860
ctggcgaaca aagttaatgc ggggtgttcag ggcggctctgt ttctgatgtg ggcgaacgat 1920
gtggtggaag atttcaccac caacatctg cgtaaagata ccctggataa aatcagcgat 1980
gttagcgcga ttattccgta tattgggtccg gcgctgaaca ttagcaatag cgtgcgtcgt 2040
ggcaatttta ccgaagcggt tgccggttacc ggtgtgacca ttctgctgga agcgtttccg 2100
gaatttacca ttccggcgct ggggtgcgtt gtgatctata gcaaagtga ggaacgcaac 2160
gaaatcatca aaaccatcga taactgcctg gaacagcgta ttaaacgctg gaaagatagc 2220
tatgaatgga tgatgggcac ctggctgagc cgtattatca cccagttcaa caacatcagc 2280
taccagatgt acgatagcct gaactatcag gcgggtgcga ttaagcgaa aatcgatctg 2340
gaatacaaaa aatacagcgg cagcgataaa gaaaacatca aaagccaggt tgaaaacctg 2400
aaaaacagcc tggatgtgaa aattagcgaa gcgatgaata acatcaacaa attcatccgc 2460
gaatgcagcg tgacctacct gttcaaaaac atgctgccga aagtgatcga tgaactgaac 2520
gaatttgatc gcaacaccaa agcgaaactg atcaacctga tcgatagcca caacattatt 2580
ctggtgggcg aagtggataa actgaaacg aaagttaaca acagcttcca gaacaccatc 2640
ccgtttaaca tcttcagcta taccaacaac agcctgctga aagatatcat caacgaatac 2700
ttcaatctag actag 2715

```

<210> SEQ ID NO 18

<211> LENGTH: 904

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 18

```

Gly Ser Glu Phe Met Pro Ile Thr Ile Asn Asn Phe Asn Tyr Ser Asp
1           5           10           15

```

```

Pro Val Asp Asn Lys Asn Ile Leu Tyr Leu Asp Thr His Leu Asn Thr
20           25           30

```

```

Leu Ala Asn Glu Pro Glu Lys Ala Phe Arg Ile Thr Gly Asn Ile Trp
35           40           45

```

```

Val Ile Pro Asp Arg Phe Ser Arg Asn Ser Asn Pro Asn Leu Asn Lys
50           55           60

```

```

Pro Pro Arg Val Thr Ser Pro Lys Ser Gly Tyr Tyr Asp Pro Asn Tyr

```

-continued

65	70				75				80			
Leu Ser Thr Asp	Ser	Asp	Lys	Asp	Thr	Phe	Leu	Lys	Glu	Ile	Ile	Lys
	85					90					95	
Leu Phe Lys Arg	Ile	Asn	Ser	Arg	Glu	Ile	Gly	Glu	Glu	Leu	Ile	Tyr
	100				105					110		
Arg Leu Ser Thr	Asp	Ile	Pro	Phe	Pro	Gly	Asn	Asn	Asn	Thr	Pro	Ile
	115			120					125			
Asn Thr Phe Asp	Phe	Asp	Val	Asp	Phe	Asn	Ser	Val	Asp	Val	Lys	Thr
	130			135				140				
Arg Gln Gly Asn	Asn	Trp	Val	Lys	Thr	Gly	Ser	Ile	Asn	Pro	Ser	Val
	145		150				155					160
Ile Ile Thr Gly	Pro	Arg	Glu	Asn	Ile	Ile	Asp	Pro	Glu	Thr	Ser	Thr
	165					170					175	
Phe Lys Leu Thr	Asn	Asn	Thr	Phe	Ala	Ala	Gln	Glu	Gly	Phe	Gly	Ala
	180				185					190		
Leu Ser Ile Ile	Ser	Ile	Ser	Pro	Arg	Phe	Met	Leu	Thr	Tyr	Ser	Asn
	195			200					205			
Ala Thr Asn Asp	Val	Gly	Glu	Gly	Arg	Phe	Ser	Lys	Ser	Glu	Phe	Cys
	210			215				220				
Met Asp Pro Ile	Leu	Ile	Leu	Met	His	Glu	Leu	Asn	His	Ala	Met	His
	225		230				235					240
Asn Leu Tyr Gly	Ile	Ala	Ile	Pro	Asn	Asp	Gln	Thr	Ile	Ser	Ser	Val
	245					250					255	
Thr Ser Asn Ile	Phe	Tyr	Ser	Gln	Tyr	Asn	Val	Lys	Leu	Glu	Tyr	Ala
	260				265					270		
Glu Ile Tyr Ala	Phe	Gly	Gly	Pro	Thr	Ile	Asp	Leu	Ile	Pro	Lys	Ser
	275			280					285			
Ala Arg Lys Tyr	Phe	Glu	Glu	Lys	Ala	Leu	Asp	Tyr	Tyr	Arg	Ser	Ile
	290			295				300				
Ala Lys Arg Leu	Asn	Ser	Ile	Thr	Thr	Ala	Asn	Pro	Ser	Ser	Phe	Asn
	305		310				315					320
Lys Tyr Ile Gly	Glu	Tyr	Lys	Gln	Lys	Leu	Ile	Arg	Lys	Tyr	Arg	Phe
	325					330					335	
Val Val Glu Ser	Ser	Gly	Glu	Val	Thr	Val	Asn	Arg	Asn	Lys	Phe	Val
	340				345					350		
Glu Leu Tyr Asn	Glu	Leu	Thr	Gln	Ile	Phe	Thr	Glu	Phe	Asn	Tyr	Ala
	355			360					365			
Lys Ile Tyr Asn	Val	Gln	Asn	Arg	Lys	Ile	Tyr	Leu	Ser	Asn	Val	Tyr
	370			375				380				
Thr Pro Val Thr	Ala	Asn	Ile	Leu	Asp	Asp	Asn	Val	Tyr	Asp	Ile	Gln
	385		390				395					400
Asn Gly Phe Asn	Ile	Pro	Lys	Ser	Asn	Leu	Asn	Val	Leu	Phe	Met	Gly
	405					410					415	
Gln Asn Leu Ser	Arg	Asn	Pro	Ala	Leu	Arg	Lys	Val	Asn	Pro	Glu	Asn
	420				425					430		
Met Leu Tyr Leu	Phe	Thr	Lys	Phe	Cys	Val	Asp	Ala	Ile	Asp	Gly	Arg
	435				440				445			
Phe Gly Gly Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn
	450			455				460				
Gln Ala Leu Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
	465			470			475					480

-continued

Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Arg	Glu	Leu	Leu	Val	Lys	Asn
				485					490					495	
Thr	Asp	Leu	Pro	Phe	Ile	Gly	Asp	Ile	Ser	Asp	Val	Lys	Thr	Asp	Ile
		500						505					510		
Phe	Leu	Arg	Lys	Asp	Ile	Asn	Glu	Glu	Thr	Glu	Val	Ile	Tyr	Tyr	Pro
	515					520						525			
Asp	Asn	Val	Ser	Val	Asp	Gln	Val	Ile	Leu	Ser	Lys	Asn	Thr	Ser	Glu
	530					535					540				
His	Gly	Gln	Leu	Asp	Leu	Leu	Tyr	Pro	Ser	Ile	Asp	Ser	Glu	Ser	Glu
545					550					555					560
Ile	Leu	Pro	Gly	Glu	Asn	Gln	Val	Phe	Tyr	Asp	Asn	Arg	Thr	Gln	Asn
			565						570					575	
Val	Asp	Tyr	Leu	Asn	Ser	Tyr	Tyr	Tyr	Leu	Glu	Ser	Gln	Lys	Leu	Ser
		580						585					590		
Asp	Asn	Val	Glu	Asp	Phe	Thr	Phe	Thr	Arg	Ser	Ile	Glu	Glu	Ala	Leu
		595					600					605			
Asp	Asn	Ser	Ala	Lys	Val	Tyr	Thr	Tyr	Phe	Pro	Thr	Leu	Ala	Asn	Lys
	610					615					620				
Val	Asn	Ala	Gly	Val	Gln	Gly	Gly	Leu	Phe	Leu	Met	Trp	Ala	Asn	Asp
625					630					635					640
Val	Val	Glu	Asp	Phe	Thr	Thr	Asn	Ile	Leu	Arg	Lys	Asp	Thr	Leu	Asp
			645						650					655	
Lys	Ile	Ser	Asp	Val	Ser	Ala	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu
		660						665					670		
Asn	Ile	Ser	Asn	Ser	Val	Arg	Arg	Gly	Asn	Phe	Thr	Glu	Ala	Phe	Ala
	675						680					685			
Val	Thr	Gly	Val	Thr	Ile	Leu	Leu	Glu	Ala	Phe	Pro	Glu	Phe	Thr	Ile
	690					695					700				
Pro	Ala	Leu	Gly	Ala	Phe	Val	Ile	Tyr	Ser	Lys	Val	Gln	Glu	Arg	Asn
705						710				715					720
Glu	Ile	Ile	Lys	Thr	Ile	Asp	Asn	Cys	Leu	Glu	Gln	Arg	Ile	Lys	Arg
			725						730					735	
Trp	Lys	Asp	Ser	Tyr	Glu	Trp	Met	Met	Gly	Thr	Trp	Leu	Ser	Arg	Ile
		740						745					750		
Ile	Thr	Gln	Phe	Asn	Asn	Ile	Ser	Tyr	Gln	Met	Tyr	Asp	Ser	Leu	Asn
	755						760					765			
Tyr	Gln	Ala	Gly	Ala	Ile	Lys	Ala	Lys	Ile	Asp	Leu	Glu	Tyr	Lys	Lys
	770					775					780				
Tyr	Ser	Gly	Ser	Asp	Lys	Glu	Asn	Ile	Lys	Ser	Gln	Val	Glu	Asn	Leu
785					790					795					800
Lys	Asn	Ser	Leu	Asp	Val	Lys	Ile	Ser	Glu	Ala	Met	Asn	Asn	Ile	Asn
			805						810					815	
Lys	Phe	Ile	Arg	Glu	Cys	Ser	Val	Thr	Tyr	Leu	Phe	Lys	Asn	Met	Leu
		820						825					830		
Pro	Lys	Val	Ile	Asp	Glu	Leu	Asn	Glu	Phe	Asp	Arg	Asn	Thr	Lys	Ala
		835					840					845			
Lys	Leu	Ile	Asn	Leu	Ile	Asp	Ser	His	Asn	Ile	Ile	Leu	Val	Gly	Glu
	850					855					860				
Val	Asp	Lys	Leu	Lys	Ala	Lys	Val	Asn	Asn	Ser	Phe	Gln	Asn	Thr	Ile
865					870					875					880

-continued

Pro Phe Asn Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile
885 890 895

Ile Asn Glu Tyr Phe Asn Leu Asp
900

<210> SEQ ID NO 19

<211> LENGTH: 2742

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 19

```

ggatccgaat tcatgccgat caccatcaac aacttcaact acagcgatcc ggtggataac      60
aaaaacatcc tgtacctgga taccatctg aataccctgg cgaacgaacc gaaaaagcg      120
tttcgtatca cgggcaacat ttgggttatt ccg gatcgtt ttagccgtaa cagcaaccg      180
aatctgaata aaccgcgcg tgttaccagc ccgaaaagcg gttattacga tccgaactat      240
ctgagcaccg atagcgataa agataccttc ctgaaagaaa tcatacaact gttcaaacgc      300
atcaacagcc gtgaaattgg cgaagaactg atctatcgcc tgagcaccga tattccgttt      360
ccgggcaaca acaacacccc gatcaacacc ttgtatttcg atgtggattt caacagcggt      420
gatgttaaaa cccgcaggcg taacaattgg gtgaaaaccg gcagcattaa cccgagcggtg      480
attattaccg gtccgcgcga aaacattatt gatccggaaa ccagcacctt taaactgacc      540
aacaacacct ttgcgcgcga ggaagggttt ggccgcgtga gcattattag cattagcccg      600
cgctttatgc tgacctatag caacgcgacc aacgatgttg gtgaaggccg ttccagcaaa      660
agcgaatttt gcatggaccc gatcctgatc ctgatgcatg aactgaacca tgcgatgcat      720
aacctgtatg gcatcgcgat tccgaacgat cagaccatta gcagcgtgac cagcaacatc      780
ttttacagcc agtacaacgt gaaactggaa tatgcggaaa tctatgcgtt tggcgggtccg      840
accattgatc tgattccgaa aagcgcgcgc aaatacttcg aagaaaaagc gctggattac      900
tatcgcagca ttgcgaaacg tctgaacagc attaccaccg cgaatccgag cagcttcaac      960
aaatatatcg gcgaataata acagaaactg atccgcaaat atcgctttgt ggtggaaagc      1020
agcggcgaag ttaccgttaa ccgcaataaa ttcgtggaac tgtacaacga actgacccag      1080
atcttcaccg aatttaacta tgcgaaaatc tataacgtgc agaaccgtaa aatctacctg      1140
agcaacgtgt ataccccggt gaccgcgaat attctggatg ataacgtgta cgatatccag      1200
aacggcttta acatcccga aagcaacctg aacgttctgt ttatgggcca gaacctgagc      1260
cgtaatccgg cgctgcgtaa agtgaaccgg gaaaacatgc tgtacctgtt caccaaattt      1320
tgctgcgacg gcatcattac ctccaaaact aaatctctga tagaaggtag atttgccggg      1380
ttcacgggag caccgaaatc agcgcgtaaa ttagctaacc aggcgctagc gggcgggtggc      1440
ggtagcggcg gtggcggtag cggcggtggc ggtagcgcac tagtgctgca gtgtcgtgaa      1500
ctgtggtgta aaaacaccga tctgcggttt attggcgata tcagcgatgt gaaaaccgat      1560
atcttctctg gcaaagatat caacgaagaa accgaagtga tctactaccc ggataacgtg      1620
agcgttgatc aggtgatcct gagcaaaaac accagcgaac atggtcagct ggatctgctg      1680
tatccgagca ttgatagcga aagcgaaatt ctgccgggag aaaaccaggt gttttacgat      1740
aaccgtaacc agaactgga ttacctgaac agctattact acctggaaag ccagaaactg      1800

```

-continued

```

agcgataacg tggaagattt tacctttacc cgcagcattg aagaagcgct ggataacagc 1860
gcgaaagttt acacctattt tccgaccctg gcgaacaaag ttaatgcggg tgttcagggc 1920
gggtctgtttc tgatgtgggc gaacgatgtg gtggaagatt tcaccaccaa catcctgcgt 1980
aaagataccc tggataaaat cagcgatgtt agcgcgatta ttccgtatat tggtcggcg 2040
ctgaacatta gcaatagcgt gcgtcgtggc aattttaccg aagcgtttgc ggttaccggt 2100
gtgaccattc tgctggaagc gtttcggaa tttaccattc cggcgctggg tgcgtttgtg 2160
atctatagca aagtgcagga acgcaacgaa atcatcaaaa ccacgataa ctgcctggaa 2220
cagcgtatta aacgctggaa agatagctat gaatggatga tgggcacctg gctgagccgt 2280
attatcacc cagttcaaaa catcagctac cagatgtacg atagcctgaa ctatcagggc 2340
gggtcggatta aagcgaaaat cgatctggaa tacaaaaaat acagcggcag cgataaagaa 2400
aacatcaaaa gccaggttga aaacctgaaa aacagcctgg atgtgaaaat tagcgaagcg 2460
atgaataaca tcaacaaatt catccgcgaa tgcagcgtga cctacctgtt caaaaacatg 2520
ctgccgaaag tgatcgatga actgaacgaa tttgatcgca acaccaaagc gaaactgatc 2580
aacctgatcg atagccacaa cattattctg gtgggcgaag tggataaact gaaagcgaaa 2640
gttaacaaca gttccagaa caccatccg ttaacatct tcagctatac caacaacagc 2700
ctgctgaaag atatcatcaa cgaatacttc aatctagact ag 2742

```

<210> SEQ ID NO 20

<211> LENGTH: 913

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 20

```

Gly Ser Glu Phe Met Pro Ile Thr Ile Asn Asn Phe Asn Tyr Ser Asp
1           5           10           15

Pro Val Asp Asn Lys Asn Ile Leu Tyr Leu Asp Thr His Leu Asn Thr
20           25           30

Leu Ala Asn Glu Pro Glu Lys Ala Phe Arg Ile Thr Gly Asn Ile Trp
35           40           45

Val Ile Pro Asp Arg Phe Ser Arg Asn Ser Asn Pro Asn Leu Asn Lys
50           55           60

Pro Pro Arg Val Thr Ser Pro Lys Ser Gly Tyr Tyr Asp Pro Asn Tyr
65           70           75           80

Leu Ser Thr Asp Ser Asp Lys Asp Thr Phe Leu Lys Glu Ile Ile Lys
85           90           95

Leu Phe Lys Arg Ile Asn Ser Arg Glu Ile Gly Glu Glu Leu Ile Tyr
100          105          110

Arg Leu Ser Thr Asp Ile Pro Phe Pro Gly Asn Asn Asn Thr Pro Ile
115          120          125

Asn Thr Phe Asp Phe Asp Val Asp Phe Asn Ser Val Asp Val Lys Thr
130          135          140

Arg Gln Gly Asn Asn Trp Val Lys Thr Gly Ser Ile Asn Pro Ser Val
145          150          155          160

Ile Ile Thr Gly Pro Arg Glu Asn Ile Ile Asp Pro Glu Thr Ser Thr
165          170          175

Phe Lys Leu Thr Asn Asn Thr Phe Ala Ala Gln Glu Gly Phe Gly Ala

```

-continued

180						185						190					
Leu	Ser	Ile	Ile	Ser	Ile	Ser	Pro	Arg	Phe	Met	Leu	Thr	Tyr	Ser	Asn		
		195					200					205					
Ala	Thr	Asn	Asp	Val	Gly	Glu	Gly	Arg	Phe	Ser	Lys	Ser	Glu	Phe	Cys		
	210					215					220						
Met	Asp	Pro	Ile	Leu	Ile	Leu	Met	His	Glu	Leu	Asn	His	Ala	Met	His		
	225				230					235					240		
Asn	Leu	Tyr	Gly	Ile	Ala	Ile	Pro	Asn	Asp	Gln	Thr	Ile	Ser	Ser	Val		
			245						250					255			
Thr	Ser	Asn	Ile	Phe	Tyr	Ser	Gln	Tyr	Asn	Val	Lys	Leu	Glu	Tyr	Ala		
		260						265					270				
Glu	Ile	Tyr	Ala	Phe	Gly	Gly	Pro	Thr	Ile	Asp	Leu	Ile	Pro	Lys	Ser		
	275						280					285					
Ala	Arg	Lys	Tyr	Phe	Glu	Glu	Lys	Ala	Leu	Asp	Tyr	Tyr	Arg	Ser	Ile		
	290					295					300						
Ala	Lys	Arg	Leu	Asn	Ser	Ile	Thr	Thr	Ala	Asn	Pro	Ser	Ser	Phe	Asn		
	305				310					315					320		
Lys	Tyr	Ile	Gly	Glu	Tyr	Lys	Gln	Lys	Leu	Ile	Arg	Lys	Tyr	Arg	Phe		
			325						330					335			
Val	Val	Glu	Ser	Ser	Gly	Glu	Val	Thr	Val	Asn	Arg	Asn	Lys	Phe	Val		
		340						345					350				
Glu	Leu	Tyr	Asn	Glu	Leu	Thr	Gln	Ile	Phe	Thr	Glu	Phe	Asn	Tyr	Ala		
	355						360					365					
Lys	Ile	Tyr	Asn	Val	Gln	Asn	Arg	Lys	Ile	Tyr	Leu	Ser	Asn	Val	Tyr		
	370					375					380						
Thr	Pro	Val	Thr	Ala	Asn	Ile	Leu	Asp	Asp	Asn	Val	Tyr	Asp	Ile	Gln		
	385				390					395					400		
Asn	Gly	Phe	Asn	Ile	Pro	Lys	Ser	Asn	Leu	Asn	Val	Leu	Phe	Met	Gly		
			405						410					415			
Gln	Asn	Leu	Ser	Arg	Asn	Pro	Ala	Leu	Arg	Lys	Val	Asn	Pro	Glu	Asn		
		420						425					430				
Met	Leu	Tyr	Leu	Phe	Thr	Lys	Phe	Cys	Val	Asp	Gly	Ile	Ile	Thr	Ser		
	435					440						445					
Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	Phe	Gly	Gly	Phe	Thr	Gly	Ala		
	450					455					460						
Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn	Gln	Ala	Leu	Ala	Gly	Gly	Gly		
	465				470					475					480		
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu		
			485					490						495			
Gln	Cys	Arg	Glu	Leu	Leu	Val	Lys	Asn	Thr	Asp	Leu	Pro	Phe	Ile	Gly		
			500					505					510				
Asp	Ile	Ser	Asp	Val	Lys	Thr	Asp	Ile	Phe	Leu	Arg	Lys	Asp	Ile	Asn		
		515					520					525					
Glu	Glu	Thr	Glu	Val	Ile	Tyr	Tyr	Pro	Asp	Asn	Val	Ser	Val	Asp	Gln		
	530					535					540						
Val	Ile	Leu	Ser	Lys	Asn	Thr	Ser	Glu	His	Gly	Gln	Leu	Asp	Leu	Leu		
	545				550					555					560		
Tyr	Pro	Ser	Ile	Asp	Ser	Glu	Ser	Glu	Ile	Leu	Pro	Gly	Glu	Asn	Gln		
			565						570					575			
Val	Phe	Tyr	Asp	Asn	Arg	Thr	Gln	Asn	Val	Asp	Tyr	Leu	Asn	Ser	Tyr		
			580					585					590				

-continued

Tyr Tyr Leu Glu Ser Gln Lys Leu Ser Asp Asn Val Glu Asp Phe Thr
 595 600 605
 Phe Thr Arg Ser Ile Glu Glu Ala Leu Asp Asn Ser Ala Lys Val Tyr
 610 615 620
 Thr Tyr Phe Pro Thr Leu Ala Asn Lys Val Asn Ala Gly Val Gln Gly
 625 630 635 640
 Gly Leu Phe Leu Met Trp Ala Asn Asp Val Val Glu Asp Phe Thr Thr
 645 650 655
 Asn Ile Leu Arg Lys Asp Thr Leu Asp Lys Ile Ser Asp Val Ser Ala
 660 665 670
 Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Ser Asn Ser Val Arg
 675 680 685
 Arg Gly Asn Phe Thr Glu Ala Phe Ala Val Thr Gly Val Thr Ile Leu
 690 695 700
 Leu Glu Ala Phe Pro Glu Phe Thr Ile Pro Ala Leu Gly Ala Phe Val
 705 710 715 720
 Ile Tyr Ser Lys Val Gln Glu Arg Asn Glu Ile Ile Lys Thr Ile Asp
 725 730 735
 Asn Cys Leu Glu Gln Arg Ile Lys Arg Trp Lys Asp Ser Tyr Glu Trp
 740 745 750
 Met Met Gly Thr Trp Leu Ser Arg Ile Ile Thr Gln Phe Asn Asn Ile
 755 760 765
 Ser Tyr Gln Met Tyr Asp Ser Leu Asn Tyr Gln Ala Gly Ala Ile Lys
 770 775 780
 Ala Lys Ile Asp Leu Glu Tyr Lys Lys Tyr Ser Gly Ser Asp Lys Glu
 785 790 795 800
 Asn Ile Lys Ser Gln Val Glu Asn Leu Lys Asn Ser Leu Asp Val Lys
 805 810 815
 Ile Ser Glu Ala Met Asn Asn Ile Asn Lys Phe Ile Arg Glu Cys Ser
 820 825 830
 Val Thr Tyr Leu Phe Lys Asn Met Leu Pro Lys Val Ile Asp Glu Leu
 835 840 845
 Asn Glu Phe Asp Arg Asn Thr Lys Ala Lys Leu Ile Asn Leu Ile Asp
 850 855 860
 Ser His Asn Ile Ile Leu Val Gly Glu Val Asp Lys Leu Lys Ala Lys
 865 870 875 880
 Val Asn Asn Ser Phe Gln Asn Thr Ile Pro Phe Asn Ile Phe Ser Tyr
 885 890 895
 Thr Asn Asn Ser Leu Leu Lys Asp Ile Ile Asn Glu Tyr Phe Asn Leu
 900 905 910

Asp

<210> SEQ ID NO 21

<211> LENGTH: 2673

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 21

ggatccatgg agttcggttaa caaacagttc aactataaag acccaggttaa cgggtgttgac 60

attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc 120

-continued

cacaacaaaa tctgggttat cccggaacgt gataccctta ctaaccgga agaaggtgac	180
ctgaaccgcg caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactggt cgagcgattt	300
tactccaccg acctggggcg tatgctgctg actagcatcg ttgcgggtat cccgttctgg	360
ggcggttcta ccatcgatag cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct	480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct cacccgtaac	540
ggctacgggt ccaactcagta catccgttct tctccggact tcaccttcgg ttttgaagaa	600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg	660
gttacacctg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggc cacgacgcta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccatcggtgg taccactgct tctctccagt acatgaagaa cgtttttaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgccgaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggct tggtcgagtt ttacaagctg ctgtgcgctg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatacggg ggtttcatgg cgctagcggg cggtggcggg	1380
agcggcgggt gcggtagcgg cgggtggcgg agcgactag tgctgcagtg tatcaagggt	1440
aacaactggg atttattctt cagcccgagt gaagacaact tcaccaacga cctgaacaaa	1500
ggtaagaaaa tcacctcaga tactaacatc gaagcagccg aagaaaacat ctgcgtggac	1560
ctgatccagc agtactacct gacctttaat ttcgacaacg agccggaaaa catttctatc	1620
gaaaacctga gctctgatat catcgccag ctggaactga tgcggaacat cgaacgtttc	1680
ccaaacggta aaaagtacga gctggacaaa tataccatgt tccactacct gcgcgcgcag	1740
gaatttgaac acggcaaatc cggatcgca ctgactaact ccgttaacga agctctgctc	1800
aaccggtccc gtgtatacac cttcttctct agcgactacg tgaaaaaggt caacaaagcg	1860
actgaagctg caatgttctt gggttgggtt gaacagcttg tttatgattt tacgcacgag	1920
acgtccgaag tatctactac cgacaaaatt gcggatatca ctatcatcat cccgtacatc	1980
ggtcggctc tgaacattgg caacatgctg taaaaagacg acttcgttgg cgcactgac	2040
ttctccgggt cgggtatcct gctggagtct atcccgaaa tcgccatccc ggtactgggc	2100
acctttgctc tggtttctta cattgcaaac aaggtttctg ctgtacaaac catcgacaac	2160
gcgctgagca aacgtaacga aaaatgggat gaagtttaca aatatatcgt gaccaactgg	2220
ctggctaagg ttaatactca gatcgacctc atccgcaaaa aaatgaaaga agcactggaa	2280
aaccaggcgg aagctaccaa ggcaatcatt aactaccagt acaaccagta caccgaggaa	2340
gaaaaaaca acatcaactt caacatcgac gatctgtcct ctaaaactgaa cgaatccatc	2400

-continued

```

aacaaagcta tgatcaacat caacaagttc ctgaaccagt gctctgtaag ctatctgatg 2460
aactccatga tcccgtagcg tgtaaactgt ctggaggact tcgatgcgtc tctgaaagac 2520
gccctgctga aatacattta cgacaacctt ggcaactctga tcggtcaggt tgatcgtctg 2580
aaggacaaaag tgaacaatac cttatcgacc gacatccctt ttcagctcag taaatatgtc 2640
gataaccaac gccttttgtc cactctagac tag 2673

```

```

<210> SEQ ID NO 22
<211> LENGTH: 890
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 22

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10          15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245         250         255
Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
260         265         270
Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
275         280         285
Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser

```

-continued

290	295	300
Ile Val Gly Thr Thr 305	Ala Ser Leu Gln Tyr 310	Met Lys Asn Val Phe Lys 315 320
Glu Lys Tyr Leu 325	Ser Glu Asp Thr 330	Ser Gly Lys Phe Ser Val Asp 335
Lys Leu Lys Phe Asp 340	Lys Leu Tyr 345	Lys Met Leu Thr Glu Ile Tyr Thr 350
Glu Asp Asn Phe Val 355	Lys Phe Phe Lys Val 360	Leu Asn Arg Lys Thr Tyr 365
Leu Asn Phe Asp Lys 370	Ala Val Phe Lys Ile 375	Asn Ile Val Pro Lys Val 380
Asn Tyr Thr Ile Tyr 385	Asp Gly Phe Asn Leu 390	Arg Asn Thr Asn Leu Ala 395 400
Ala Asn Phe Asn Gly 405	Gln Asn Thr Glu Ile 410	Asn Asn Met Asn Phe Thr 415
Lys Leu Lys Asn Phe 420	Thr Gly Leu Phe Glu 425	Phe Tyr Lys Leu Leu Cys 430
Val Asp Gly Ile Ile 435	Thr Ser Lys Thr Lys 440	Ser Leu Ile Glu Gly Arg 445
Tyr Gly Gly Phe Met 450	Ala Leu Ala Gly Gly 455	Gly Gly Ser Gly Gly Gly 460
Gly Ser Gly Gly Gly 465	Gly Ser Ala Leu Val 470	Leu Gln Cys Ile Lys Val 475 480
Asn Asn Trp Asp Leu 485	Phe Phe Ser Pro Ser 490	Glu Asp Asn Phe Thr Asn 495
Asp Leu Asn Lys Gly 500	Glu Glu Ile Thr Ser 505	Asp Thr Asn Ile Glu Ala 510
Ala Glu Glu Asn Ile 515	Ser Leu Asp Leu Ile 520	Gln Gln Tyr Tyr Leu Thr 525
Phe Asn Phe Asp Asn 530	Glu Pro Glu Asn Ile 535	Ser Ile Glu Asn Leu Ser 540
Ser Asp Ile Ile Gly 545	Gln Leu Glu Leu Met 550	Pro Asn Ile Glu Arg Phe 555 560
Pro Asn Gly Lys Lys 565	Tyr Glu Leu Asp Lys 570	Tyr Thr Met Phe His Tyr 575
Leu Arg Ala Gln Glu 580	Phe Glu His Gly Lys 585	Ser Arg Ile Ala Leu Thr 590
Asn Ser Val Asn Glu 595	Ala Leu Leu Asn Pro 600	Ser Arg Val Tyr Thr Phe 605
Phe Ser Ser Asp Tyr 610	Val Lys Lys Val Asn 615	Lys Ala Thr Glu Ala Ala 620
Met Phe Leu Gly Trp 625	Val Glu Gln Leu Val 630	Tyr Asp Phe Thr Asp Glu 635 640
Thr Ser Glu Val Ser 645	Thr Thr Asp Lys Ile 650	Ala Asp Ile Thr Ile Ile 655
Ile Pro Tyr Ile Gly 660	Pro Ala Leu Asn Ile 665	Gly Asn Met Leu Tyr Lys 670
Asp Asp Phe Val Gly 675	Ala Leu Ile Phe Ser 680	Gly Ala Val Ile Leu Leu 685
Glu Phe Ile Pro Glu 690	Ile Ala Ile Pro Val 695	Leu Gly Thr Phe Ala Leu 700

-continued

Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn
705					710					715					720
Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile
			725						730					735	
Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg
			740					745					750		
Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala
		755				760						765			
Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn
	770					775					780				
Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile
785					790					795					800
Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val
			805						810					815	
Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu	Glu
		820						825					830		
Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp
	835						840					845			
Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val
	850					855					860				
Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr	Val
865					870					875					880
Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp						
			885						890						

<210> SEQ ID NO 23

<211> LENGTH: 2751

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 23

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac	60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc	120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaaggtgac	180
ctgaacccgc caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactggt cgagcgtatt	300
tactccaccg acctgggccc tatgctgctg actagcatcg ttccgggtat cccgttctgg	360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggagc gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct	480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct caccgtaac	540
ggctacgggt cactcagta catccgttcc tctccggact tcacottcgg ttttgaagaa	600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgct	660
gttacctctg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggc cagcagceta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900

-continued

```

aaagcgaaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaaa 960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc 1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagtctctt 1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc 1140
gtgccgaaa gtaactacac tatctacgat gggttcaacc tgcgtaacac caacctggct 1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac 1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa 1320
actaaatctc tgatcgaagg tcgttacggt gggttcatga cctctgaaaa atctcagacc 1380
ccgctgggta cctgttcaa aaacgctatc atcaaaaacg cttacaaaaa aggtgaagcg 1440
ctagcgggtg gtggtgggtc tgggtgggtg gggtctggtg gtggtgggtc tgcactagtg 1500
ctgcagtgtg tcaagggtta caactgggat ttattcttca gcccgagtga agacaacttc 1560
accaacgacc tgaacaaagg tgaagaaatc acctcagata ctaacatcga agcagccgaa 1620
gaaaacatct cgctggacct gatccagcag tactacctga cctttaattt cgacaacgag 1680
ccgaaaaaca tttctatcga aaacctgagc tctgatatca tcggccagct ggaactgatg 1740
ccgaacatcg aacgtttccc aaacggtaaa aagtacgagc tggacaaata taccatgttc 1800
cactacctgc gcgcgcagga atttgaacac ggcaaattcc gtatcgact gactaactcc 1860
gttaacgaag ctctgctcaa cccgtcccg gtataacct tcttctctag cgactacgtg 1920
aaaaaggtca acaaagcgac tgaagctgca atgttcttgg gttgggttga acagcttggt 1980
tatgatttta ccgacgagac gtccgaagta tctactaccg acaaaattgc ggatatcact 2040
atcatcatcc cgtacatcgg tccggtctg aacattggca acatgctgta caaagacgac 2100
ttcgttggcg cactgatctt ctccggtcgc gtgatcctgc tggagtccat cccggaatc 2160
gccatcccg tactgggac ctttgcctg gtttcttaca ttgcaaaaaa ggttctgact 2220
gtacaaaacca tcgacaacgc gctgagcaaa cgtaacgaaa aatgggatga agtttacaaa 2280
tatatcgtga ccaactggct ggctaaggtt aatactcaga tcgacctcat ccgcaaaaaa 2340
atgaaagaag cactgaaaa ccaggcggaa gctaccaagg caatcattaa ctaccgtac 2400
aaccagtaca ccgaggaaga aaaaaacaac atcaacttca acatcgacga tctgtcctct 2460
aaactgaacg aatccatcaa caaagctatg atcaacatca acaagtctct gaaccagtgc 2520
tctgtaagct atctgatgaa ctccatgatc ccgtacggtg ttaaacgtct ggaggacttc 2580
gatgctctc tgaaagacgc cctgctgaaa tacatttacg acaaccgtgg cactctgatc 2640
ggtcagggtg atcgtctgaa ggacaaaagt aacaatacct tatcgaccga catccctttt 2700
cagctcagta aatatgtcga taaccaacgc cttttgtcca ctctagacta g 2751

```

<210> SEQ ID NO 24

<211> LENGTH: 916

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 24

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1 5 10 15

-continued

Asn	Gly	Val	Asp	Ile	Ala	Tyr	Ile	Lys	Ile	Pro	Asn	Ala	Gly	Gln	Met
			20					25					30		
Gln	Pro	Val	Lys	Ala	Phe	Lys	Ile	His	Asn	Lys	Ile	Trp	Val	Ile	Pro
		35					40					45			
Glu	Arg	Asp	Thr	Phe	Thr	Asn	Pro	Glu	Glu	Gly	Asp	Leu	Asn	Pro	Pro
	50					55					60				
Pro	Glu	Ala	Lys	Gln	Val	Pro	Val	Ser	Tyr	Tyr	Asp	Ser	Thr	Tyr	Leu
	65				70					75					80
Ser	Thr	Asp	Asn	Glu	Lys	Asp	Asn	Tyr	Leu	Lys	Gly	Val	Thr	Lys	Leu
			85						90					95	
Phe	Glu	Arg	Ile	Tyr	Ser	Thr	Asp	Leu	Gly	Arg	Met	Leu	Leu	Thr	Ser
			100					105						110	
Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu
		115					120					125			
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly
	130					135					140				
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala
	145				150					155					160
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn
				165					170					175	
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro
		180						185					190		
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro
		195					200					205			
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala
	210					215				220					
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn
	225				230					235					240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser
			245					250						255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260					265						270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
		275				280						285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290				295						300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
	305				310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325					330						335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
			340				345						350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
		355				360						365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
	385				390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405					410						415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys

-continued

420						425						430					
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg		
435						440						445					
Tyr	Gly	Gly	Phe	Met	Thr	Ser	Glu	Lys	Ser	Gln	Thr	Pro	Leu	Val	Thr		
450						455						460					
Leu	Phe	Lys	Asn	Ala	Ile	Ile	Lys	Asn	Ala	Tyr	Lys	Lys	Gly	Glu	Ala		
465						470						475					
Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly		
485						490						495					
Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe		
500						505						510					
Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu		
515						520						525					
Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser		
530						535						540					
Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu		
545						550						555					
Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln		
565						570						575					
Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr		
580						585						590					
Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe		
595						600						605					
Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala		
610						615						620					
Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val		
625						630						635					
Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val		
645						650						655					
Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr		
660						665						670					
Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro		
675						680						685					
Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala		
690						695						700					
Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile		
705						710						715					
Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn		
725						730						735					
Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn		
740						745						750					
Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala		
755						760						765					
Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala		
770						775						780					
Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr		
785						790						795					
Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp		
805						810						815					
Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn		
820						825						830					

-continued

Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser
835 840 845

Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu
850 855 860

Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile
865 870 875 880

Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr
885 890 895

Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu
900 905 910

Ser Thr Leu Asp
915

<210> SEQ ID NO 25

<211> LENGTH: 2709

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 25

```

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaaggtgac    180
ctgaacccgc caccggaagc gaaacaggtg ccggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt    300
tactccaccg acctggggccg tatgtgctg actagcatcg ttcgcggtat cccgttctgg    360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcagagt taagagcttt ggtcacgaag ttctgaacct caccgtaac    540
ggctacgggt ccactcagta catccgttcc tctccggact tcaccttcgg ttttgaagaa    600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg    660
gttaccctgg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttgccggt cagcagccta aattcatoga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaaat ccatcgtggg taccactgct tctctccagt acatgaagaa cgttttttaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaatc   1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaaacac caacctggct   1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tgttcagatt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320
actaaatctc tgatagaagg tagatttggc gggttcacgg gcgcacgcaa atcagcgcg   1380
aaacgtaaga accaggcgct agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt   1440

```

-continued

```

ggcggtagcg cactagtgtc gcagtgtatc aagggttaaca actgggattt attcttcagc 1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctcagatact 1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc 1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc 1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg 1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatcccgt 1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccgtgt atacaccttc 1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttgggt 1920
tgggttgaac agcttgttta tgattttacc gacgagacgt ccgaagtatc tactaccgac 1980
aaaaattgagg atatcactat catcatcccg tacatcggtc cggctctgaa cattggcaac 2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct cgggtgcggt gatcctgctg 2100
gagttcatcc cgaaaatcgc catcccggtg ctgggcacct ttgctctggt ttcttacatt 2160
gcaacaacgg ttctgactgt acaaacctac gacaacgcgc tgagcaaacg taacgaaaaa 2220
tgggatgaag ttacaaata tatcgtgacc aactggctgg ctaagggtta tactcagatc 2280
gacctcatcc gaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgac tgctcttaa actgaacgaa tccatcaaca aagctatgat caacatcaac 2460
aagttcctga accagtgtc tgtaagctat ctgatgaact ccatgatccc gtacggtgtt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgctgaaata catctacgac 2580
aaccgtggca ctctgatcgg tcaggttgat cgtctgaagg acaaaagtga caatacctta 2640
tcgaccgaca tcccttttca gtcagtaaa tatgtcgata accaacgcct tttgtccact 2700
ctagactag 2709

```

<210> SEQ ID NO 26

<211> LENGTH: 902

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 26

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15

```

```

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30

```

```

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45

```

```

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60

```

```

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80

```

```

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95

```

```

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110

```

-continued

Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu
	115						120					125			
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly
	130					135					140				
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala
145					150					155					160
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn
				165					170					175	
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro
		180						185					190		
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro
		195					200					205			
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala
	210					215				220					
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn
225					230					235					240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser
				245					250					255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260						265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
		275					280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290					295					300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325						330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
			340					345					350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
		355					360					365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405						410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
		435					440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Arg	Lys	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465				470						475					480
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp
			485						490					495	
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys
			500					505					510		
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn

-continued

515					520					525					
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp
530					535					540					
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile
545					550					555				560	
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys
				565					570					575	
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln
				580				585					590		
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn
				595			600					605			
Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp
				610			615				620				
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly
625					630					635					640
Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val
				645					650					655	
Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile
				660				665					670		
Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val
				675			680					685			
Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro
				690			695				700				
Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile
705					710					715					720
Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys
				725					730					735	
Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp
				740				745					750		
Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys
				755			760					765			
Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr
				770			775				780				
Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn
785					790					795					800
Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met
				805					810					815	
Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met
				820				825					830		
Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala
				835			840					845			
Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr
				850			855				860				
Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu
865					870					875					880
Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg
				885					890					895	
Leu	Leu	Ser	Thr	Leu	Asp										
				900											

-continued

```

<211> LENGTH: 2691
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 27

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgtgtgtgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaagggtgac    180
ctgaacccgc caccggaagc gaaacagggt cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaagggtgta ctaaaactgt cgagcgtatt    300
tactccaccg acctggggcg tatgtgctg actagcatcg ttcgcggtat cccgttcttg    360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcgagtg taagagcttt ggtcagcaag ttctgaacct cacccgtaac    540
ggctacgggt ccaactcagta catccgttct tctccggact tcaccttcgg ttttgaagaa    600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg    660
gttacccttg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaagt taaccaaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttgccggt cagcagccta aattcatcga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct   1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320
actaaatctc tgatagaagg tagatttggc gggttcacgg gcgcacgcaa atcagcggcg   1380
ctagcggggc gtggcggtag cggcgggtgg ggtagcggcg gtggcggtag cgcactagtg   1440
ctgcagtgta tcaagggtta caactgggat ttattcttca gcccagtgta agacaacttc   1500
accaacgacc tgaacaaagg tgaagaaatc acctcagata ctaacatcga agcagccgaa   1560
gaaaacatct cgctggacct gatccagcag tactacctga cctttaattt cgacaacgag   1620
ccggaacaaa tttctatcga aaacctgagc tctgatatca tcggccagct ggaactgatg   1680
ccgaacatcg aacgtttccc aaacggtaaa aagtacgagc tggacaaata taccatgttc   1740
cactacctgc gcgcgcagga atttgaacac ggcaaatccc gtatcgact gactaactcc   1800
gttaacgaag ctctgctcaa cccgtcccgt gtatacacct tcttctctag cgactacgtg   1860
aaaaaggtea acaaaagcag tgaagctgca atgttcttgg gttgggttga acagcttggt   1920
tatgatttta ccgacgagac gtccgaagta tctactaccg acaaaattgc ggatatact   1980
atcatcatcc cgtacatcgg tccggctctg aacattggca acatgctgta caaagacgac   2040

```

-continued

```

ttcgttggcg cactgatctt ctccggtgcg gtgacctgc tggagttcat cccggaaatc 2100
gccatcccggt tactgggcac ctttgcctcg gtttcttaca ttgcaaaca ggttctgact 2160
gtacaaacca tcgacaacgc gctgagcaaa cgtaacgaaa aatgggatga agtttacaaa 2220
tatatcgtga ccaactggct ggctaagggt aatactcaga tcgacctcat ccgcaaaaaa 2280
atgaaagaag cactggaaaa ccaggcggaa gctaccaagg caatcattaa ctaccagtac 2340
aaccagtaca ccgaggaaga aaaaaacaac atcaacttca acatcgacga tctgtcctct 2400
aaactgaacg aatccatcaa caaagctatg atcaacatca acaagttcct gaaccagtgc 2460
tctgtaagct atctgatgaa ctccatgatc ccgtacggtg ttaaactgtc ggaggacttc 2520
gatgcgtctc tgaaagacgc cctgtgaaa tacatttacg acaaccgtgg cactctgatc 2580
ggtcagggtg atcgtctgaa ggacaaagtg aacaatacct tatcgaccga catccctttt 2640
cagctcagta aatatgtcga taaccaacgc cttttgtcca ctctagacta g 2691

```

<210> SEQ ID NO 28

<211> LENGTH: 896

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 28

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240

```

-continued

Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	245	250	255
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp	260	265	270
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	275	280	285
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	290	295	300
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	305	310	315
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	325	330	335
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	340	345	350
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	355	360	365
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	435	440	445
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Ala	Leu	Ala	Gly	Gly	450	455	460
Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	465	470	475
Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	485	490	495
Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	500	505	510
Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	515	520	525
Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	530	535	540
Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	545	550	555
Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	565	570	575
Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	580	585	590
Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	595	600	605
Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	610	615	620
Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	625	630	635

-continued

Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr Asp Lys Ile
 645 650 655
 Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile
 660 665 670
 Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser
 675 680 685
 Gly Ala Val Ile Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val
 690 695 700
 Leu Gly Thr Phe Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr
 705 710 715 720
 Val Gln Thr Ile Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp
 725 730 735
 Glu Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr
 740 745 750
 Gln Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln
 755 760 765
 Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr
 770 775 780
 Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser
 785 790 795 800
 Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe
 805 810 815
 Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr
 820 825 830
 Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu
 835 840 845
 Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp
 850 855 860
 Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe
 865 870 875 880
 Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp
 885 890 895

<210> SEQ ID NO 29

<211> LENGTH: 2691

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 29

```

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctggggttat cccggaacgt gataccttta ctaacccgga agaaggtgac    180
ctgaacccgc caccggaagc gaaacagggtg cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactgtt cgagcgtatt    300
tactccaccg acctggggccg tatgtgctg actagcatcg ttccggttat cccgttctgg    360
ggcggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcagatg taagagcttt ggtcacgaag ttctgaacct caccgtaac    540

```

-continued

ggctacggtt ccactcagta catccgttcc tctccggact tcaccttcgg ttttgaagaa	600
tccttggaag tagacacgaa cccactgtgt ggcgtggta aattcgcaac tgatcctgcg	660
gttacccctgg ctacagaaat gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggc cacgacgcta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccactgtggg taccactgct tctctccagt acatgaagaa cgtttttaaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatttgcc ggtttcacgg gcgcacgcaa atatgcggcg	1380
ctagcgggcg gtggcggtag cggcgggtggc ggtagcggcg gtggcggtag cgcactagtg	1440
ctgcagtgtg tcaagggtta caactgggat ttattcttca gcccgagtga agacaacttc	1500
accaacgacc tgaacaaagg tgaagaaatc acctcagata ctaacatcga agcagccgaa	1560
gaaaacatct cgctggacct gatccagcag tactacctga cctttaattt cgacaacgag	1620
ccggaaaaca tttctatcga aaacctgagc tctgatatca tcggccagct ggaactgatg	1680
ccgaacatcg aacgtttccc aaacggtaaa aagtacgagc tggacaata taccatgttc	1740
cactacctgc gcgcgcagga atttgaacac ggcaaatccc gtatcgcaat gactaactcc	1800
gttaacgaag ctctgctcaa cccgtcccgt gtataacct tcttctctag cgactacgtg	1860
aaaaaggtca acaaagcgac tgaagctgca atgttcttgg gttgggttga acagcttggt	1920
tatgatttta ccgacgagac gtccgaagta tctactaccg acaaaattgc ggatatcaat	1980
atcatcatcc cgtacatcgg tccggtctg aacattggca acatgctgta caaagacgac	2040
ttcgttggcg cactgatctt ctccggtgcg gtgatcctgc tggagttcat cccggaaatc	2100
gccatcccgg tactgggac ctttgcctcg gtttcttaca ttgcaaaaa ggttctgact	2160
gtacaaaacca tcgacaacgc gctgagcaaa cgtaacgaaa aatgggatga agtttacaaa	2220
tatatcgtga ccaactggct ggctaaggtt aatactcaga tcgacctcat ccgcaaaaaa	2280
atgaaagaag cactggaaaa ccaggcggaa gctaccaagg caatcattaa ctaccagtac	2340
aaccagtaca ccgaggaaga aaaaaacaac atcaacttca acatcgacga tctgtcctct	2400
aaactgaacg aatccatcaa caaagctatg atcaacatca acaagttcct gaaccagtgc	2460
tctgtaagct atctgatgaa ctccatgac ccgtacggtg ttaaacgtct ggaggacttc	2520
gatgctctc tgaaagacgc cctgctgaaa tacatttacg acaaccgtgg cactctgac	2580
ggtcagggtg atcgtctgaa ggacaaagtg aacaatacct tatcgaccga catccctttt	2640
cagctcagta aatatgtcga taaccaacgc cttttgtcca ctctagacta g	2691

<210> SEQ ID NO 30

<211> LENGTH: 896

<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 30

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
 1 5 10 15
 Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
 20 25 30
 Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
 35 40 45
 Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
 50 55 60
 Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
 65 70 75 80
 Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
 85 90 95
 Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
 100 105 110
 Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
 115 120 125
 Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
 130 135 140
 Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
 145 150 155 160
 Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
 165 170 175
 Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
 180 185 190
 Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
 195 200 205
 Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
 210 215 220
 His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
 225 230 235 240
 Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
 245 250 255
 Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
 260 265 270
 Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
 275 280 285
 Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser
 290 295 300
 Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys
 305 310 315 320
 Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp
 325 330 335
 Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr
 340 345 350
 Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr
 355 360 365
 Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val

-continued

370					375					380					
Asn 385	Tyr	Thr	Ile	Tyr	Asp 390	Gly	Phe	Asn	Leu	Arg 395	Asn	Thr	Asn	Leu	Ala 400
Ala	Asn	Phe	Asn	Gly 405	Gln	Asn	Thr	Glu	Ile 410	Asn	Asn	Met	Asn	Phe	Thr 415
Lys	Leu	Lys	Asn	Phe 420	Thr	Gly	Leu	Phe 425	Glu	Phe	Tyr	Lys	Leu 430	Leu	Cys
Val	Asp	Gly 435	Ile	Ile	Thr	Ser	Lys 440	Thr	Lys	Ser	Leu	Ile 445	Glu	Gly	Arg
Phe	Gly 450	Gly	Phe	Thr	Gly	Ala 455	Arg	Lys	Tyr	Ala	Ala 460	Leu	Ala	Gly	Gly
Gly 465	Gly	Ser	Gly	Gly	Gly 470	Gly	Ser	Gly	Gly	Gly 475	Gly	Ser	Ala	Leu	Val 480
Leu	Gln	Cys	Ile 485	Lys	Val	Asn	Asn	Trp	Asp 490	Leu	Phe	Phe	Ser	Pro 495	Ser
Glu	Asp	Asn	Phe 500	Thr	Asn	Asp	Leu	Asn 505	Lys	Gly	Glu	Glu	Ile 510	Thr	Ser
Asp	Thr	Asn	Ile 515	Glu	Ala	Ala	Glu 520	Glu	Asn	Ile	Ser	Leu 525	Asp	Leu	Ile
Gln	Gln	Tyr	Tyr 530	Leu	Thr	Phe 535	Asn	Phe	Asp	Asn	Glu 540	Pro	Glu	Asn	Ile
Ser 545	Ile	Glu	Asn	Leu	Ser 550	Ser	Asp	Ile	Ile	Gly 555	Gln	Leu	Glu	Leu	Met 560
Pro	Asn	Ile	Glu 565	Arg	Phe	Pro	Asn	Gly	Lys 570	Lys	Tyr	Glu	Leu	Asp 575	Lys
Tyr	Thr	Met	Phe 580	His	Tyr	Leu	Arg	Ala 585	Gln	Glu	Phe	Glu	His 590	Gly	Lys
Ser	Arg	Ile 595	Ala	Leu	Thr	Asn	Ser 600	Val	Asn	Glu	Ala	Leu 605	Leu	Asn	Pro
Ser 610	Arg	Val	Tyr	Thr	Phe 615	Phe	Ser	Ser	Asp	Tyr	Val 620	Lys	Lys	Val	Asn
Lys 625	Ala	Thr	Glu	Ala	Ala 630	Met	Phe	Leu	Gly	Trp 635	Val	Glu	Gln	Leu	Val 640
Tyr	Asp	Phe	Thr 645	Asp	Glu	Thr	Ser	Glu	Val 650	Ser	Thr	Thr	Asp	Lys 655	Ile
Ala	Asp	Ile 660	Thr	Ile	Ile	Ile	Pro	Tyr 665	Ile	Gly	Pro	Ala	Leu 670	Asn	Ile
Gly	Asn	Met 675	Leu	Tyr	Lys	Asp	Asp 680	Phe	Val	Gly	Ala	Leu 685	Ile	Phe	Ser
Gly 690	Ala	Val	Ile	Leu	Leu	Glu 695	Phe	Ile	Pro	Glu	Ile 700	Ala	Ile	Pro	Val
Leu 705	Gly	Thr	Phe	Ala	Leu 710	Val	Ser	Tyr	Ile 715	Ala	Asn	Lys	Val	Leu	Thr 720
Val	Gln	Thr	Ile 725	Asp	Asn	Ala	Leu	Ser	Lys 730	Arg	Asn	Glu	Lys	Trp 735	Asp
Glu	Val	Tyr	Lys 740	Tyr	Ile	Val	Thr	Asn 745	Trp	Leu	Ala	Lys	Val 750	Asn	Thr
Gln	Ile 755	Asp	Leu	Ile	Arg	Lys	Lys 760	Met	Lys	Glu	Ala	Leu 765	Glu	Asn	Gln
Ala	Glu 770	Ala	Thr	Lys	Ala	Ile 775	Ile	Asn	Tyr	Gln	Tyr 780	Asn	Gln	Tyr	Thr

-continued

Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser
785 790 795 800

Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe
805 810 815

Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr
820 825 830

Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu
835 840 845

Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp
850 855 860

Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe
865 870 875 880

Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp
885 890 895

<210> SEQ ID NO 31

<211> LENGTH: 2691

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 31

```

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaaggtgac    180
ctgaacccgc caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt    300
tactccaccg acctggggccg tatgtgctg actagcatcg ttccgggtat cccgttctgg    360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcagagt taagagcttt ggtcacgaag ttctgaacct caccgtaac    540
ggctacgggt ccactcagta catccgttcc tctccggact tcacctcggg ttttgaagaa    600
tccttggaag tagacacgaa cccactgctg ggcgtggta aattcgcaac tgatcctgcg    660
gttaccctgg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttgccggt cagcagccta aattcatoga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaagt ccatcgtggg taccactgct tctctccagt acatgaagaa cgttttttaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat gggttcaacc tgcgtaacac caacctggct   1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tgttcagatt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320

```

-continued

```

actaaatctc tgatagaagg tagatttggc ggtttcacgg gcgcacgcaa atcatatgcg 1380
ctagcgggcg gtggcggttag cggcggtggc ggtagcggcg gtggcggttag cgcactagtg 1440
ctgcagtgtg tcaaggttaa caactgggat ttattcttca gcccgagtga agacaacttc 1500
accaacgacc tgaacaaagg tgaagaaatc acctcagata ctaacatcga agcagccgaa 1560
gaaaacatct cgctggacct gatccagcag tactacctga cctttaattt cgacaacgag 1620
ccggaaaaca tttctatcga aaacctgagc tctgatatca tcggccagct ggaactgatg 1680
ccgaacatcg aacgtttccc aaacggtaaa aagtacgagc tggacaaata taccatgttc 1740
cactacctgc gcgcgcagga atttgaacac ggcaaatccc gtatcgact gactaactcc 1800
gttaacgaag ctctgctcaa ccggtcccg gtatacacct tcttctctag cgactacgtg 1860
aaaaagggtc acaaagcgac tgaagctgca atgttcttgg gttgggttga acagcttgtt 1920
tatgatttta ccgacgagac gtccgaagta tctactaccg acaaaattgc ggatatcact 1980
atcatcatcc cgtacatcgg tccggctctg aacattggca acatgctgta caaagacgac 2040
ttcgttggcg cactgatctt ctccggtgcy gtgatcctgc tggagtcat ccggaaatc 2100
gccatcccg tactgggcac ctttgcctg gtttcttaca ttgcaaaca ggttctgact 2160
gtacaaacca tcgacaacgc gctgagcaaa cgtaacgaaa aatgggatga agtttcaaaa 2220
tatatcgtga ccaactggct ggctaagggt aatactcaga tcgacctcat ccgcaaaaaa 2280
atgaaagaag cactggaaaa ccaggcggaa gctaccaagg caatcattaa ctaccagtac 2340
aaccagtaca ccgaggaaga aaaaaacaac atcaacttca acatcgacga tctgtcctct 2400
aaactgaacg aatccatcaa caaagctatg atcaacatca acaagttcct gaaccagtgc 2460
tctgtaagct atctgatgaa ctccatgatc ccgtacggtg ttaaacgtct ggaggacttc 2520
gatgctctc tgaaagacgc cctgtgaaa tacatttacg acaaccgtgg cactctgatc 2580
ggtcagggtg atcgtctgaa ggacaaagtg aacaatacct tatcgaccga catccctttt 2640
cagctcagta aatatgtcga taaccaacgc cttttgtcca ctctagacta g 2691

```

<210> SEQ ID NO 32

<211> LENGTH: 896

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 32

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15

```

```

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20           25           30

```

```

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35           40           45

```

```

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50           55           60

```

```

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65           70           75           80

```

```

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85           90           95

```

```

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100          105          110

```

-continued

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu	115	120	125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly	130	135	140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala	145	150	155
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn	165	170	175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro	180	185	190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro	195	200	205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala	210	215	220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn	225	230	235
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser	245	250	255
Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp	260	265	270
Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr	275	280	285
Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser	290	295	300
Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys	305	310	315
Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp	325	330	335
Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr	340	345	350
Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr	355	360	365
Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val	370	375	380
Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala	385	390	395
Ala Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr	405	410	415
Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys	420	425	430
Val Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg	435	440	445
Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Tyr Ala Leu Ala Gly Gly	450	455	460
Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Ser Ala Leu Val	465	470	475
Leu Gln Cys Ile Lys Val Asn Asn Trp Asp Leu Phe Phe Ser Pro Ser	485	490	495
Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser	500	505	510

-continued

Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile
		515					520					525			
Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile
		530				535					540				
Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met
545				550						555					560
Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys
				565					570					575	
Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys
			580					585					590		
Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro
		595					600					605			
Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn
	610				615						620				
Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val
625					630					635					640
Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile
				645					650					655	
Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile
			660					665					670		
Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser
		675					680					685			
Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val
		690					695					700			
Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr
705					710					715					720
Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp
				725					730					735	
Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr
			740					745					750		
Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln
		755					760					765			
Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr
		770				775						780			
Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser
785					790					795					800
Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe
				805					810					815	
Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr
			820					825					830		
Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu
		835					840					845			
Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp
		850				855						860			
Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe
865					870					875					880
Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp
				885					890						895

<210> SEQ ID NO 33

<211> LENGTH: 2709

<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 33

```

ggatccatgg agttcggttaa caaacagttc aactataaag acccagttaa cgggtgttgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccggga agaaggtgac    180
ctgaacccgc caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt    300
tactccaccg acctggggccg tatgtctctg actagcatcg ttccgggtat cccgttcttg    360
ggcggttcta ccatcgatgc cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcgatg taagagcttt ggtcacgaag ttctgaacct cacccgtaac    540
ggctacgggt ccaactcagta catccgtttc tctccggact tcaccttcgg ttttgaagaa    600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg    660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccctg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttgccggt cacgacgcta aattcatcga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagtctcga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct   1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320
actaaatctc tgatagaagg tagatttggc ggtttcacgg gcgcacgcaa atcagcgcgt   1380
aaatatgcta accagggcgt agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt   1440
ggcggtagcg cactagtgtc gcagtgtatc aagggttaaca actgggattt attcttcagc   1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctacagatac   1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc   1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc   1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg   1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatcccg   1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccggtg atacaccttc   1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttgagg   1920
tgggttgaa agcttgttta tgattttacc gacgagacgt ccgaagtatc tactaccgac   1980
aaaaattcgg atatcactat catcatcccg tacatcggtc cggctctgaa cattggcaac   2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct ccggtgcggg gatcctgctg   2100

```

-continued

```

gagttcatcc cggaaatcgc catcccggtg ctgggcacct ttgctctggt ttcttacatt 2160
gcaaacaagg ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa 2220
tgggatgaag ttacaaata tatcgtgacc aactggctgg ctaagggttaa tactcagatc 2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgatc tgcctcctaa actgaacgaa tccatcaaca aagctatgat caacatcaac 2460
aagttcctga accagtgtct tgtaagctat ctgatgaact ccatgatccc gtacggtggt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgcgtgaaata catttacgac 2580
aaccgtggca ctctgatcgg tcaggttgat cgtctgaagg acaaaagtga caatacctta 2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct ttgtccact 2700
ctagactag 2709

```

```

<210> SEQ ID NO 34
<211> LENGTH: 902
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 34

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240

```

-continued

Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	245	250	255
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp	260	265	270
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	275	280	285
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	290	295	300
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	305	310	315
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	325	330	335
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	340	345	350
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	355	360	365
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	435	440	445
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Tyr	Ala	Asn	450	455	460
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	465	470	475
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	485	490	495
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	500	505	510
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	515	520	525
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	530	535	540
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	545	550	555
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	565	570	575
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	580	585	590
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	595	600	605
Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	610	615	620
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	625	630	635
																		640

-continued

Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	
				645					650					655		
Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	
			660					665					670			
Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	
		675					680					685				
Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	
		690				695					700					
Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	
	705				710					715					720	
Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	
			725					730						735		
Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	
		740						745					750			
Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	
		755					760				765					
Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	
	770				775						780					
Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	
	785				790					795					800	
Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	
			805						810					815		
Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	
			820					825					830			
Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	
		835					840					845				
Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	
	850					855					860					
Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	
	865				870				875						880	
Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	
			885						890					895		
Leu	Leu	Ser	Thr	Leu	Asp											
				900												

<210> SEQ ID NO 35

<211> LENGTH: 2697

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 35

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac	60
attgcttaca tcaaaatccc gaacgtggc cagatgcagc cggtaaaggc attcaaaatc	120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac	180
ctgaacccgc caccggaagc gaaacagggtg ccggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactgtt cgagcgtatt	300
tactccaccg acctggggccg tatgtgtctg actagcatcg ttgcggtat cccgttctgg	360
ggcggttcta ccatcgatgc cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct	480

-continued

gatatcatcc agttcagatg taagagcttt ggtcacgaag ttctgaacct cacccgtaac	540
ggctacgggt ccaactcagta catccgttcc tctccggact tcaccttcgg ttttgaagaa	600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg	660
gttacacctg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
cgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggc cacgacgcta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaac ctatctgaac ttgcacaagg cagtattcaa aatcaacatc	1140
gtgccgaag ttaactacac tatctacgat ggtttcaacc tgcgtaaac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatttggc ggtttcacgg gcgcacgcaa atcagcgcgt	1380
aaagcgctag cgggcggtg cggtagcggc ggtggcggtg gcggcggtg cggtagcgca	1440
ctagtgtgc agtgtatcaa ggttaacaac tgggatttat tcttcagccc gagtgaagac	1500
aacttcacca acgacctgaa caaagggtgaa gaaatcacct cagatactaa catcgaagca	1560
gccgaagaaa acatctcgct ggacctgatc cagcagtact acctgacctt taatttcgac	1620
aacgagccgg aaaacatttc tatcgaaaac ctgagctctg atatcatcgg ccagctggaa	1680
ctgatgccga acatcgaacg tttcccaaac ggtaaaaagt acgagctgga caaataatac	1740
atgttcact acctgcgcgc gcaggaattt gaacacggca aatcccgat cgactgact	1800
aactccgta acgaagctct gctcaaccg tcccggtgat acaccttctt ctctagcgac	1860
tacgtgaaaa aggtcaacaa agcgactgaa gctgcaatgt tcttgggttg ggttgaacag	1920
cttgtttatg attttaccga cgagacgtcc gaagtatcta ctaccgacaa aattgcggat	1980
atcactatca tcatcccgta catcgggtccg gctctgaaca ttggcaacat gctgtacaaa	2040
gacgacttcg ttggcgact gatcttctcc ggtgcggtga tcctgctgga gttcatcccg	2100
gaaatcgcca tcccggtact gggcaccttt gctctggttt cttacattgc aaacaagggt	2160
ctgactgtac aaaccatcga caacgcgtg agcaaacgta acgaaaaatg ggatgaagtt	2220
tacaaatata tcgtgaccaa ctggctggct aagggttaata ctcatcga cctcatccgc	2280
aaaaaaatga aagaagcact ggaaaaccag gcggaagcta ccaaggcaat cattaactac	2340
cagtacaacc agtacaccga ggaagaaaaa aacaacatca acttcaacat cgacgatctg	2400
tcctctaacc tgaacgaatc catcaacaaa gctatgatca acatcaacaa gttcctgaac	2460
cagtgtctg taagctatct gatgaactcc atgatccgt acggtgttaa acgtctggag	2520
gacttcgatg cgtctctgaa agacgcctg ctgaaataca ttacgacaa ccgtggcact	2580
ctgatcggtc aggttgatcg tctgaaggac aaagtgaaca atacctatc gaccgacatc	2640
ccttttcagc tcagtaaata tgcgataac caacgccttt tgcactctc agactag	2697

-continued

```

<210> SEQ ID NO 36
<211> LENGTH: 898
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 36

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1          5          10          15

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125

Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140

Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160

Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175

Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190

Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205

Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220

His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240

Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245         250         255

Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
260         265         270

Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
275         280         285

Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser
290         295         300

Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys
305         310         315         320

Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp
325         330         335

Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr
340         345         350

Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr

```

-continued

355					360					365					
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
370						375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
				405					410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
			420					425					430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
		435					440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Ala	Leu	Ala
	450					455					460				
Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Gly	Ser	Ala
465					470				475						480
Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser
				485					490					495	
Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile
			500					505					510		
Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp
		515					520					525			
Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu
	530					535					540				
Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu
545					550					555					560
Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu
			565						570					575	
Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His
			580					585					590		
Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu
		595					600					605			
Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys
	610					615					620				
Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln
625					630					635					640
Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp
			645						650					655	
Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu
		660					665						670		
Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile
		675					680					685			
Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile
	690					695					700				
Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val
705					710					715					720
Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys
			725						730					735	
Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val
		740					745					750			
Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu
	755					760						765			

-continued

Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln
 770 775 780
 Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu
 785 790 795 800
 Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn
 805 810 815
 Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile
 820 825 830
 Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp
 835 840 845
 Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln
 850 855 860
 Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile
 865 870 875 880
 Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr
 885 890 895

Leu Asp

<210> SEQ ID NO 37
 <211> LENGTH: 51
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 37

tttggcgggtt tcacggggcgc acgcaaatca gcgcgtaaat tagctaacca g

51

<210> SEQ ID NO 38
 <211> LENGTH: 17
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 38

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Leu Ala Asn
 1 5 10 15

Gln

<210> SEQ ID NO 39
 <211> LENGTH: 33
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 39

tttggcgggtt tcacggggcgc acgcaaatca gcg

33

<210> SEQ ID NO 40
 <211> LENGTH: 11
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 40

-continued

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala
1 5 10

<210> SEQ ID NO 41
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 41

tttggcgggtt tcacgggcgc acgcaaatat gcg 33

<210> SEQ ID NO 42
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 42

Phe Gly Gly Phe Thr Gly Ala Arg Lys Tyr Ala
1 5 10

<210> SEQ ID NO 43
<211> LENGTH: 33
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 43

tttggcgggtt tcacgggcgc acgcaaatca tat 33

<210> SEQ ID NO 44
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 44

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Tyr
1 5 10

<210> SEQ ID NO 45
<211> LENGTH: 51
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 45

tttggcgggtt tcacgggcgc acgcaaatca gcgcgtaaat atgctaacca g 51

<210> SEQ ID NO 46
<211> LENGTH: 17
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 46

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Tyr Ala Asn

-continued

1	5	10	15	
---	---	----	----	--

Gln

<210> SEQ ID NO 47
 <211> LENGTH: 39
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 47

tttggcgggtt tcacggggcgc acgcaaatca ggcggtaaa	39
--	----

<210> SEQ ID NO 48
 <211> LENGTH: 13
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 48

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys	
1 5 10	

<210> SEQ ID NO 49
 <211> LENGTH: 51
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 49

tttggcgggtt tcacggggcgc acgcaaatca ggcggtaaac gcaaaaacca g	51
--	----

<210> SEQ ID NO 50
 <211> LENGTH: 17
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 50

Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Arg Lys Asn	
1 5 10 15	

Gln

<210> SEQ ID NO 51
 <211> LENGTH: 2736
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 51

ctcgggattg agggtcgttt tggcggtttc acggggcgac gcaaatcagc gcgtaaatta	60
gctaaccaga ctagtggcgg tgggggtagt ggcggtggcg gttcgggcgg ggggtgggagc	120
cctaggggat ccatggagtt cgtaaataaa cagttcaact ataaagaccc agttaacggt	180
gttgacattg cttacatcaa aatcccgaac gctggccaga tgcagccggt aaaggcattc	240
aaaaatccaca acaaaatctg gggttatccc gaacgtgata cctttactaa cccggaagaa	300

-continued

ggtagacctga acccgccacc ggaagcgaaa caggtgccgg tatottacta tgactccacc	360
tacctgtcta ccgataacga aaaggacaac tacctgaaag gtgttactaa actgttcgag	420
cgtatttact ccaccgacct gggccgtatg ctgctgacta gcctcggtcg cggtatcccg	480
ttctggggcg gttctacat cgataccgaa ctgaaagtaa tcgacactaa ctgcatcaac	540
gttattcagc cggacgggtc ctatcggttc gaagaactga acctggtgat catcggcccg	600
tctgctgata tcatccagtt cgagtgtgag agctttggtc acgaagttct gaacctcacc	660
cgtaacggct acggttccac tcagtacatc cgtttctctc cggacttcac ctteggtttt	720
gaagaatccc tggaagtaga cacgaacca ctgctggcg ctggtaaatt cgcaactgat	780
cctgcgggta ccctggctca cgaactgatt catgcaggcc accgcctgta cggtatcgcc	840
atcaatccga accgtgtctt caaagttaac accaacgcgt attacgagat gtccgggtctg	900
gaagttagct tcgaagaact gcgtactttt ggcggtcacg acgctaaatt catcgactct	960
ctgcaagaaa acgagttccg tctgtactac tataacaagt tcaagatat cgcattccacc	1020
ctgaacaaag cgaaatccat cgtgggtacc actgcttctc tccagtacat gaagaacgtt	1080
tttaagaaa aatacctgct cagcgaagac acctccgca aattctctgt agacaagttg	1140
aaattcgata aactttacaa aatgctgact gaaatttaca ccgaagacaa ctteggtaag	1200
ttctttaaag ttctgaacgg caaaacctat ctgaacttcg acaaggcagt attcaaaatc	1260
aacatcgtgc cgaaagttaa ctacactatc tacgatgggt tcaacctgcg taacaccaac	1320
ctggctgcta attttaacgg ccagaacacg gaaatcaaca acatgaactt cacaaaactg	1380
aaaaacttca ctggtctgtt cgagttttac aagctgctgt gcgtcgacgg catcattacc	1440
tccaaaacta aatctctgat agaaggtaga aacaaagcgc tgaacctgca gtgtatcaag	1500
gttaacaact gggatttatt cttcagcccg agtgaagaca acttcaccaa cgacctgaac	1560
aaaggtgaag aaatcacctc agataactac atcgaagcag ccgaagaaaa catctcgtctg	1620
gacctgatcc agcagtaacta cctgacctt aatttcgaca acgagccgga aaacatttct	1680
atcgaaaacc tgagctctga tatcatcggc cagctggaac tgatgccgaa catcgaaactg	1740
ttcccaaacg gtaaaaagta cgagctggac aaatatacca tgttccacta cctgcgcgcg	1800
caggaatttg aacacggcaa atcccgatc gcaactgact actccgttaa cgaagctctg	1860
ctcaaccctg cccgtgtata caccttcttc tctagcgact acgtgaaaaa ggtcaacaaa	1920
gcgactgaag ctgcaatgtt cttgggttgg gttgaacagc ttgtttatga tttaccgac	1980
gagacgtccg aagtatctac taccgacaaa attgcggata tcactatcat catcccgta	2040
atcggtcccg ctctgaacat tggcaacatg ctgtacaaag acgacttcgt tggcgcaactg	2100
atcttctccg gtgcgggtgat cctgctggag ttcattcccg aaatcgccat cccggtactg	2160
ggcacctttg ctctggtttc ttacattgca aacaagggtc tgactgtaca aaccatcgac	2220
aacgcgctga gcaaacgtaa cgaaaaatgg gatgaagttt acaaatatat cgtgaccaac	2280
tggttggtga aggttaatac tcagatcgac ctcatccgca aaaaaatgaa agaagcactg	2340
gaaaaccagg cggaagctac caaggcaatc attaactacc agtacaacca gtacaccgag	2400
gaagaaaaaa acaacatcaa cttcaacatc gacgatctgt cctctaaact gaacgaatcc	2460
atcaacaaag ctatgatcaa catcaacaag ttctgaacc agtgctctgt aagctatctg	2520
atgaactcca tgatcccgta cgggtgttaa cgtctggagg acttcgatgc gtctctgaaa	2580

-continued

```

gacgcctgc tgaatacat ttacgacaac cgtggcactc tgatecgtca ggttgatcgt 2640
ctgaaggaca aagtgaacaa taccttatcg accgacatcc cttttcagct cagtaaatat 2700
gtcgataaacc aacgcctttt gtccactcta gactag 2736

```

```

<210> SEQ ID NO 52
<211> LENGTH: 911
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 52

```

```

Leu Gly Ile Glu Gly Arg Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser
1      5      10      15
Ala Arg Lys Leu Ala Asn Gln Thr Ser Gly Gly Gly Gly Ser Gly Gly
20     25     30
Gly Gly Ser Gly Gly Gly Gly Ser Pro Arg Gly Ser Met Glu Phe Val
35     40     45
Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val Asn Gly Val Asp Ile Ala
50     55     60
Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met Gln Pro Val Lys Ala Phe
65     70     75     80
Lys Ile His Asn Lys Ile Trp Val Ile Pro Glu Arg Asp Thr Phe Thr
85     90     95
Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro Glu Ala Lys Gln Val
100    105    110
Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu Ser Thr Asp Asn Glu Lys
115    120    125
Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu Phe Glu Arg Ile Tyr Ser
130    135    140
Thr Asp Leu Gly Arg Met Leu Leu Thr Ser Ile Val Arg Gly Ile Pro
145    150    155    160
Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu Leu Lys Val Ile Asp Thr
165    170    175
Asn Cys Ile Asn Val Ile Gln Pro Asp Gly Ser Tyr Arg Ser Glu Glu
180    185    190
Leu Asn Leu Val Ile Ile Gly Pro Ser Ala Asp Ile Ile Gln Phe Glu
195    200    205
Cys Lys Ser Phe Gly His Glu Val Leu Asn Leu Thr Arg Asn Gly Tyr
210    215    220
Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro Asp Phe Thr Phe Gly Phe
225    230    235    240
Glu Glu Ser Leu Glu Val Asp Thr Asn Pro Leu Leu Gly Ala Gly Lys
245    250    255
Phe Ala Thr Asp Pro Ala Val Thr Leu Ala His Glu Leu Ile His Ala
260    265    270
Gly His Arg Leu Tyr Gly Ile Ala Ile Asn Pro Asn Arg Val Phe Lys
275    280    285
Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser Gly Leu Glu Val Ser Phe
290    295    300
Glu Glu Leu Arg Thr Phe Gly Gly His Asp Ala Lys Phe Ile Asp Ser
305    310    315    320

```

-continued

Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	Tyr	Asn	Lys	Phe	Lys	Asp
			325						330					335	
Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	Ile	Val	Gly	Thr	Thr	Ala
			340					345					350		
Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	Glu	Lys	Tyr	Leu	Leu	Ser
		355					360					365			
Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	Lys	Leu	Lys	Phe	Asp	Lys
	370					375					380				
Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	Glu	Asp	Asn	Phe	Val	Lys
385					390					395					400
Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	Leu	Asn	Phe	Asp	Lys	Ala
			405					410						415	
Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	Asn	Tyr	Thr	Ile	Tyr	Asp
			420					425					430		
Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	Ala	Asn	Phe	Asn	Gly	Gln
		435				440						445			
Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	Lys	Leu	Lys	Asn	Phe	Thr
	450					455					460				
Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	Val	Asp	Gly	Ile	Ile	Thr
465				470					475					480	
Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	Asn	Lys	Ala	Leu	Asn	Leu
			485					490					495		
Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu
		500					505						510		
Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp
	515					520						525			
Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln
	530					535					540				
Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser
545					550				555					560	
Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro
			565					570					575		
Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr
		580					585						590		
Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser
	595					600						605			
Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser
	610				615						620				
Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys
625				630					635					640	
Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr
			645					650					655		
Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala
		660					665					670			
Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Gly
	675					680						685			
Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser	Gly
	690				695						700				
Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val	Leu
705				710					715					720	
Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val

725										730					735				
Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu				
740										745					750				
Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr	Gln				
755										760					765				
Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln	Ala				
770										775					780				
Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr	Glu				
785										790					795				
Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser	Lys				
805										810					815				
Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe	Leu				
820										825					830				
Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr	Gly				
835										840					845				
Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu	Leu				
850										855					860				
Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp	Arg				
865										870					875				
Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Gln				
885										890					895				
Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp					
900										905					910				
<210> SEQ ID NO 53 <211> LENGTH: 177 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic <400> SEQUENCE: 53																			
ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga 60 tttggcgggtt tcacggggcgc acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg 120 ggtggtggtg gttctgcact agtgcctgcac acgcacggtc tagaatgata aaagctt 177																			
<210> SEQ ID NO 54 <211> LENGTH: 192 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic <400> SEQUENCE: 54																			
ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga 60 tttggcgggtt tcacggggcgc acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg 120 ggtggtggtg gttctggtgg tgggtggttct gcactagtgc tgcagacgca cggtctagaa 180 tgataaaagc tt 192																			
<210> SEQ ID NO 55 <211> LENGTH: 222 <212> TYPE: DNA <213> ORGANISM: Artificial Sequence <220> FEATURE: <223> OTHER INFORMATION: Synthetic																			

-continued

<400> SEQUENCE: 55

```

ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga      60
tttggcgggtt tcacgggcgc acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg    120
ggtggtggtg gttctggtgg tgggtggttct ggtggtggtg gttctggtgg tgggtggttct  180
gcactagtgc tgcagacgca cgggtctagaa tgataaaagc tt                        222

```

<210> SEQ ID NO 56

<211> LENGTH: 237

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 56

```

ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga      60
tttggcgggtt tcacgggcgc acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg    120
ggtggtggtg gttctggtgg tgggtggttct ggtggtggtg gttctggtgg tgggtggttct  180
ggtggtggtg gttctgcact agtgctgcag acgcacggtc tagaatgata aaagctt       237

```

<210> SEQ ID NO 57

<211> LENGTH: 228

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 57

```

ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga      60
tttggcgggtt tcacgggcgc acgcaaatca gcgcgtaaat tagctaacca ggcgctagcg    120
gctgaagctg ctgctaaga agctgctgct aaagaagctg ctgctaagc tggtggcggt      180
ggttccgcac tagtgctgca gacgcacggt ctagaatgat aaaagctt                  228

```

<210> SEQ ID NO 58

<211> LENGTH: 2694

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 58

```

ggatccatgg agttcggtta caaacagttc aactataaag acccagttaa cgggtgttgac      60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacacaacaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaaggtagc    180
ctgaaccgcg caccggaagc gaaacagggt cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgt cgagcgtatt    300
tactccaccg acctggggcg tatgctgctg actagcatcg ttccgggtat cccgttctgg    360
ggcgggttcta ccatcgatgc cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccgggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct cacccgtaac    540
ggctacgggt ccactcagta catccgtttc tctccggact tcaccttcgg ttttgaagaa    600

```

-continued

tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg	660
gttacccttg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg ttttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggg cagcagccta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaaatctt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa caccgaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatttggc ggtttcacgg gcgcacgcaa atcagcgcgt	1380
aaattagcta accagggcgt agcgggtggg ggtggttctg gtggtggtgg ttctgacta	1440
gtgctgcagt gtatcaaggt taacaactgg gatttattct tcagcccgag tgaagacaac	1500
ttcaccaacg acctgaacaa aggtgaagaa atcacctcag atactaacat cgaagcagcc	1560
gaagaaaaca tctcgttgga cctgatccag cagtactacc tgacctttaa ttctgacaac	1620
gagccggaac acatttctat cgaaaacctg agctctgata tcatcggccca gctggaactg	1680
atgccgaaca tcgaacgttt cccaaacggt aaaaagtacg agctggacaa atataccatg	1740
ttccactacc tgcgcgcgca ggaatttgaa caccgcaaat cccgtatcgc actgactaac	1800
tccgttaacg aagctctgct caacccgctc cgtgtataca ccttcttctc tagcgactac	1860
gtgaaaaagg tcaacaaagc gactgaagct gcaatgttct tgggttgggt tgaacagctt	1920
gtttatgatt ttaccgacga gacgtccgaa gtatctacta ccgacaaaat tgcggatata	1980
actatcatca tcccgtacat cggtcoggtc ctgaacattg gcaacatgct gtacaaagac	2040
gacttcgttg gcgcactgat cttctccggt gcggtgatcc tgcgtgagtt catcccgga	2100
atcgccatcc cgggtactgg cacccttgct ctggtttctt acattgcaaa caaggttctg	2160
actgtacaaa ccatcgacaa cgcgctgagc aaacgtaacg aaaaatggga tgaagtttac	2220
aaatatatcg tgaccaactg gctggctaag gtttaatactc agatcgacct catccgcaaa	2280
aaaatgaaag aagcactgga aaaccaggcg gaagctacca aggcaatcat taactaccag	2340
tacaaccagt acaccgagga agaaaaaac aacatcaact tcaacatcga cgatctgtcc	2400
tctaaactga acgaatccat caacaaagct atgatcaaca tcaacaagtt cctgaaccag	2460
tgctctgtaa gctatctgat gaactccatg atcccgtacg gtgttaaacg tctggaggac	2520
ttcgatgcgt ctctgaaaga cgccttgctg aaatacatct acgacaaccg tggcactctg	2580
atcggtcagg ttgategtct gaaggacaaa gtgaacaata ccttatcgac cgacatccct	2640
tttcagctca gtaaatatgt cgataaccaa cgccttttgt ccactctaga ctag	2694

<210> SEQ ID NO 59

<211> LENGTH: 897

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

-continued

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 59

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
 1 5 10 15
 Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
 20 25 30
 Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
 35 40 45
 Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
 50 55 60
 Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
 65 70 75 80
 Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
 85 90 95
 Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
 100 105 110
 Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
 115 120 125
 Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
 130 135 140
 Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
 145 150 155 160
 Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
 165 170 175
 Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
 180 185 190
 Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
 195 200 205
 Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
 210 215 220
 His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
 225 230 235 240
 Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
 245 250 255
 Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
 260 265 270
 Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
 275 280 285
 Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser
 290 295 300
 Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys
 305 310 315 320
 Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp
 325 330 335
 Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr
 340 345 350
 Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr
 355 360 365
 Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val
 370 375 380

-continued

Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405					410						415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
	435					440						445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu
465				470					475						480
Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro
			485						490					495	
Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr
		500					505						510		
Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu
	515					520						525			
Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn
	530					535					540				
Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu
545				550					555						560
Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp
			565						570					575	
Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly
		580					585						590		
Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn
	595					600					605				
Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val
	610					615					620				
Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu
625				630						635					640
Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys
			645					650						655	
Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn
		660						665					670		
Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe
	675						680				685				
Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro
	690					695					700				
Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu
705				710						715					720
Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp
			725						730					735	
Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn
		740						745					750		
Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn
	755						760					765			
Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr
	770					775						780			

-continued

Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser
785					790					795				800	
Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys
			805						810					815	
Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro
			820					825					830		
Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala
		835					840					845			
Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val
	850					855					860				
Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro
865					870					875				880	
Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu
				885					890					895	

Asp

<210> SEQ ID NO 60
 <211> LENGTH: 2724
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 60

```

ggatccatgg agttcggttaa caaacagttc aactataaag acccagttaa cgggtgttgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac    180
ctgaaccgcg caccggaagc gaaacaggtg ccggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt    300
tactccaccg acctggggcg tatgtgctg actagcatcg ttccgggtat cccgttctgg    360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct cacccgtaac    540
ggctacgggt ccaactcagta catccgttct tctccggact tcaccttcgg tttgaagaa    600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg    660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttgccggt cagcagccta aattcatcga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgttttttaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaac ctatctgaac ttgcacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat gggttcaacc tgcgtaacac caacctggct   1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tggtcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320

```

-continued

```

actaaatctc tgatagaagg tagatttggc ggtttcacgg gcgcacgcaa atcagcgcgt 1380
aaattagcta accagggcgt agcgggtggg ggtggttctg gtggtggtgg ttctggtggt 1440
ggtggttctg gtggtggtgg ttctgacta gtgctgcagt gtatcaaggt taacaactgg 1500
gatttattct tcagcccgag tgaagacaac ttcaccaacg acctgaacaa aggtgaagaa 1560
atcacctcag atactaacat cgaagcagcc gaagaaaaca tctcgctgga cctgatccag 1620
cagtactacc tgacctttaa ttctgacaac gagccggaaa acatttctat cgaaaacctg 1680
agctctgata tcatcggcc gctggaactg atgccgaaca tcgaacgttt cccaaacggt 1740
aaaaagtacg agctggacaa atataccatg ttccactacc tgcgcgcgca ggaatttgaa 1800
cacggcaaat cccgtatcgc actgactaac tccgttaacg aagctctgct caaccctcc 1860
cgtgtataca cttcttctc tagcgactac gtgaaaaagg tcaacaaagc gactgaagct 1920
gcaatgttct tgggttgggt tgaacagctt gtttatgatt ttaccgacga gacgtccgaa 1980
gtatctacta ccgacaaaat tgcggatata actatcatca tcccgtaacat cggtcgggct 2040
ctgaacattg gcaacatgct gtacaaagac gacttcgttg gcgcactgat cttctccggt 2100
gcggtgatcc tgctggagtt catcccgga atcgccatcc cggctactggg cacctttgct 2160
ctggtttctt acattgcaaa caaggttctg actgtacaaa ccatcgacaa cgcgctgagc 2220
aaacgtaacg aaaaatggga tgaagtttac aaatatatcg tgaccaactg gctggctaag 2280
gttaatactc agatcgacct catccgcaaa aaaatgaaag aagcactgga aaaccaggcg 2340
gaagctacca aggcaatcat taactaccag tacaaccagt acaccgagga agaaaaaac 2400
aacatcaact tcaacatcga cgatctgtcc tctaaactga acgaatccat caacaaagct 2460
atgatcaaca tcaacaagtt cctgaaccag tgctctgtaa gctatctgat gaactccatg 2520
atcccgtaac gtgttaaacg tctggaggac ttcgatgcgt ctctgaaaga cgccctgctg 2580
aaatacattt acgacaacgg tggcactctg atcggtcagg ttgatcgtct gaaggacaaa 2640
gtgaacaata cttatcgac cgacatccct tttcagctca gtaaatatgt cgataaccaa 2700
cgccctttgt ccactctaga ctag 2724

```

```

<210> SEQ ID NO 61
<211> LENGTH: 907
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 61

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95

```

-continued

Phe	Glu	Arg	Ile	Tyr	Ser	Thr	Asp	Leu	Gly	Arg	Met	Leu	Leu	Thr	Ser
			100					105					110		
Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu
		115					120					125			
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly
	130					135					140				
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala
145				150						155					160
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn
			165						170					175	
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro
		180						185					190		
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro
	195						200					205			
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala
	210					215					220				
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn
225				230						235					240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser
			245						250					255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260						265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
	275						280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290				295						300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325						330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
		340					345						350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
	355					360						365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405						410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
		435					440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465					470					475					480
Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys
			485						490					495	

-continued

Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr
			500					505					510		
Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu
		515					520					525			
Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu
		530				535					540				
Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu
545					550				555						560
Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg
			565						570					575	
Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His
			580					585					590		
Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu
		595					600					605			
Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr
						615				620					
Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala
625					630					635					640
Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp
				645					650					655	
Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile
			660					665					670		
Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr
		675					680					685			
Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu
	690					695					700				
Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala
705					710				715						720
Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp
				725					730					735	
Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr
			740					745					750		
Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile
		755					760					765			
Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys
	770					775					780				
Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn
785					790				795						800
Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser
				805					810					815	
Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser
			820					825					830		
Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu
		835					840					845			
Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr
	850					855					860				
Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys
865					870					875					880
Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr
				885					890						895
Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp					

-continued

900	905	
<210> SEQ ID NO 62		
<211> LENGTH: 207		
<212> TYPE: DNA		
<213> ORGANISM: Artificial Sequence		
<220> FEATURE:		
<223> OTHER INFORMATION: Synthetic		
<400> SEQUENCE: 62		
ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctgacga tgacgataaa		60
tttggcgggtt tcacgggcgc acgcaaatca gcgcgtaaac gtaagaacca ggcgctagcg		120
ggcgggtggcg gtacggcgcg tggcggtagc ggcgggtggcg gtacgcgcact agtgctgcag		180
acgcacgggtc tagaatgata aaagctt		207
<210> SEQ ID NO 63		
<211> LENGTH: 2709		
<212> TYPE: DNA		
<213> ORGANISM: Artificial Sequence		
<220> FEATURE:		
<223> OTHER INFORMATION: Synthetic		
<400> SEQUENCE: 63		
ggatccatgg agttcggttaa caaacagttc aactataaag acccagttaa cgggtgttgac		60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc		120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac		180
ctgaaccgcg caccggaagc gaaacaggtg ccggtatctt actatgactc cacctacctg		240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt		300
tactccaccg acctgggcgg tatgtgctg actagcatcg ttcgcggtat cccgttctgg		360
ggcggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt		420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct		480
gatatcatcc agttcgatg taagagcttt ggtcacgaag ttctgaacct cacccgtaac		540
ggctacgggt ccaactcagta catccgttcc tctccggact tcaccttcgg tttgaagaa		600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg		660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat		720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt		780
agcttcgaag aactgcgtac ttttggcggc cagcagccta aattcatcga ctctctgcaa		840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac		900
aaagcgaaat ccatcgtggg taccactgct tctctccagt acatgaagaa cgtttttaaa		960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc		1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt		1080
aaagttctga accgcaaac ctatctgaac ttgcacaagg cagtattcaa aatcaacatc		1140
gtgccgaaag ttaactacac tatctacgat gggttcaacc tgcgtaacac caacctggct		1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac		1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa		1320
actaaatctg acgatgacga taaatttggc gggtttcacgg gcgcacgcaa atcagcgcgt		1380

-continued

```

aaacgtaaga accagggcgt agcggggcgt ggcggtagcg gcggtggcgg tagcggcgggt 1440
ggcggtagcg cactagtgtc gcagtgtatc aaggttaaca actgggattt attcttcagc 1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctacagatact 1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc 1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc 1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg 1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatcccg 1800
atcgcaactga ctaactccgt taacgaagct ctgctcaacc cgtcccgtgt atacaccttc 1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttgggt 1920
tgggttgaac agcttgttta tgattttacc gacgagacgt ccgaagtatc tactaccgac 1980
aaaattgcgg atactactat catcatcccg tacatcggtc cggtcttgaa cattggcaac 2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct ccggtgcgggt gatcctgctg 2100
gagttcatcc cggaatcgc catcccggtg ctgggcacct ttgctctggt ttcttacatt 2160
gcaacaacgg ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa 2220
tgggatgaag ttacaaata tatcgtgacc aactggctgg ctaagggtta tactcagatc 2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgacg tgctctctaa actgaacgaa tccatcaaca aagctatgat caacatcaac 2460
aagttcttga accagtgtct tgtaagctat ctgatgaact ccatgatccc gtacgggtgtt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgctgaaata catttacgac 2580
aaccgtaggca ctctgatcgg tcaggttgat cgtctgaagg acaaaagtga caatacctta 2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct ttgttccact 2700
ctagactag 2709

```

```

<210> SEQ ID NO 64
<211> LENGTH: 902
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 64

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110

```

-continued

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu	115	120	125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly	130	135	140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala	145	150	155
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn	165	170	175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro	180	185	190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro	195	200	205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala	210	215	220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn	225	230	235
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser	245	250	255
Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp	260	265	270
Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr	275	280	285
Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser	290	295	300
Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys	305	310	315
Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp	325	330	335
Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr	340	345	350
Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr	355	360	365
Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val	370	375	380
Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala	385	390	395
Ala Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr	405	410	415
Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys	420	425	430
Val Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Asp Asp Asp Lys	435	440	445
Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Arg Lys Asn	450	455	460
Gln Ala Leu Ala Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly	465	470	475
Gly Gly Ser Ala Leu Val Leu Gln Cys Ile Lys Val Asn Asn Trp Asp	485	490	495
Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp Leu Asn Lys	500	505	510

-continued

<210> SEQ ID NO 65
<211> LENGTH: 207
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 65

```
ggatccacgc acgtcgacgg catcattacc tccaaaacta aatctctgat agaaggtaga    60
tttggcggtt tcacgggcgc acgcaaatca gcgcgtaaac gtaagaacca ggcgctagcg    120
ggcgggtggcg gtacggcgcg tggcggtagc ggcgggtggcg gtacgcgact agtgctgcag    180
acgcacgggtc tagaatgata aaagctt                                     207
```

<210> SEQ ID NO 66
<211> LENGTH: 2742
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 66

```
ggatccgaat tcatgccgat caccatcaac aacttcaact acagcgatcc ggtggataac    60
aaaaacatcc tgtacctgga taccatctg aataccctgg cgaacgaacc ggaaaaagcg    120
tttcgtatca cgggcaacat ttgggttatt cgggatcggt ttagccgtaa cagcaaccgc    180
aatctgaata aaccgcccgc tgttaccagc ccgaaaagcg gttattacga tccgaactat    240
ctgagcaccg atagcgataa agataccttc ctgaaagaaa tcatcaact gttcaaacgc    300
atcaacagcc gtgaaattgg cgaagaactg atctatcgcc tgagcaccga tattccgttt    360
ccgggcaaca acaacacccc gatcaacacc tttgatttcg atgtggattt caacagcggt    420
gatgttaaaa cccgccaggg taacaattgg gtgaaaaccg gcagcattaa cccgagcgtg    480
attattaccg gtccgcgcga aaacattatt gatccggaaa ccagcacctt taaactgacc    540
aacaacacct ttgcggcgca ggaagggttt ggcgcgctga gcattattag cattagcccg    600
cgctttatgc tgacctatag caacgcgacc aacgatgttg gtgaaggccg ttccagcaaa    660
agcgaatttt gcatggaccc gatcctgatc ctgatgcacg aactgaacca tgcgatgcac    720
aacctgtatg gcatcgcgat tccgaacgat cagaccatta gcagcgtgac cagcaacatc    780
ttttacagcc agtacaacgt gaaactggaa tatgcggaaa tctatgcggt tggcggtccg    840
accattgatc tgattccgaa aagcgcgcgc aaatacttcg aagaaaaagc gctggattac    900
tategcagca ttgcgaaacg tctgaacagc attaccaccg cgaatccgag cagcttcaac    960
aaatatatcg gcgaatataa acagaaaactg atccgcaaat atcgctttgt ggtggaaagc   1020
agcggcgaag ttaccgttaa ccgcaataaa ttcgtggaac tgtacaacga actgacccag   1080
atcttcaccg aatttaacta tgcgaaaatc tataacgtgc agaaccgtaa aatctacctg   1140
agcaacgtgt ataccccggt gaccgcgaat attctggatg ataacgtgta cgatatccag   1200
aacggcttta acatcccgaa aagcaacctg aacgttctgt ttatgggcca gaacctgagc   1260
cgtaatccgg cgctgcgtaa agtgaacccg gaaaacatgc tgtacctgtt caccaaattt   1320
tgcgtcgacg gcatcattac ctccaaaact aaatctctga tagaaggtag atttggcggt   1380
ttcacgggcg cagcgaatc agcgcgtaaa cgtaagaacc aggcgctagc gggcggtggc   1440
ggtagcggcg gtggcggtag cggcggtggc ggtagcgcac tagtgctgca gtgtcgtgaa   1500
```

-continued

```

ctgctggtga aaaacaccga tctgccgttt attggcgata tcagcgatgt gaaaaccgat 1560
atcttctctgc gcaaagatat caacgaagaa accgaagtga tctactaccc ggataacgtg 1620
agcgttgatc aggtgatcct gagcaaaaac accagcgaac atggtcagct ggatctgctg 1680
tatccgagca ttgatagcga aagcgaaatt ctgccgggcg aaaaccaggt gttttacgat 1740
aaccgtaccc agaactgga ttacctgaac agctattact acctggaaag ccagaaactg 1800
agcgataacg tggaagattt tacctttacc cgcagcattg aagaagcgct ggataacagc 1860
gcgaaagttt acacctatct tccgaccctg gcgaacaaag ttaatgcggg tggtcagggc 1920
ggtctgtttc tgatgtgggc gaacgatgtg gtggaagatt tcaccaccaa catcctgcgt 1980
aaagataccc tggataaaat cagcgatgtt agcgcgatta ttccgtatat tggtcggcg 2040
ctgaacatta gcaatagcgt gcgtcgtggc aattttaccg aagcgtttgc ggttaccggt 2100
gtgaccattc tgctggaagc gtttcgggaa ttaccattc cggcgctggg tgcgtttgtg 2160
atctatagca aagtgcagga acgcaacgaa atcatcaaaa ccatcgataa ctgcctggaa 2220
cagcgatatta aacgctggaa agatagctat gaatggatga tgggcacctg gctgagccgt 2280
attatcacc agttcaacaa catcagctac cagatgtacg atagcctgaa ctatcaggcg 2340
ggtgcgatta aagcgaaaat cgatctggaa taaaaaaat acagcggcag cgataaagaa 2400
aacatcaaaa gccaggttga aaacctgaaa aacagcctgg atgtgaaaat tagcgaaagc 2460
atgaataaca tcaacaaatt catccgcgaa tgcagcgtga cctacctgtt caaaaacatg 2520
ctgccgaaag tgatcgatga actgaacgaa tttgatcgca acaccaaagc gaaactgatc 2580
aacctgatcg atagccacaa cattattctg gtggcgcaag tggataaact gaaagcgaaa 2640
gttaacaaca gttccagaa caccatcccg ttaacatct tcagctatac caacaacagc 2700
ctgctgaaag atatcatcaa cgaatacttc aatctagact ag 2742

```

```

<210> SEQ ID NO 67
<211> LENGTH: 913
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 67

```

```

Gly Ser Glu Phe Met Pro Ile Thr Ile Asn Asn Phe Asn Tyr Ser Asp
1           5           10           15
Pro Val Asp Asn Lys Asn Ile Leu Tyr Leu Asp Thr His Leu Asn Thr
20          25          30
Leu Ala Asn Glu Pro Glu Lys Ala Phe Arg Ile Thr Gly Asn Ile Trp
35          40          45
Val Ile Pro Asp Arg Phe Ser Arg Asn Ser Asn Pro Asn Leu Asn Lys
50          55          60
Pro Pro Arg Val Thr Ser Pro Lys Ser Gly Tyr Tyr Asp Pro Asn Tyr
65          70          75          80
Leu Ser Thr Asp Ser Asp Lys Asp Thr Phe Leu Lys Glu Ile Ile Lys
85          90          95
Leu Phe Lys Arg Ile Asn Ser Arg Glu Ile Gly Glu Glu Leu Ile Tyr
100         105         110
Arg Leu Ser Thr Asp Ile Pro Phe Pro Gly Asn Asn Asn Thr Pro Ile
115         120         125

```

-continued

Asn Thr Phe Asp Phe Asp Val Asp Phe Asn Ser Val Asp Val Lys Thr	
130	135 140
Arg Gln Gly Asn Asn Trp Val Lys Thr Gly Ser Ile Asn Pro Ser Val	
145	150 155 160
Ile Ile Thr Gly Pro Arg Glu Asn Ile Ile Asp Pro Glu Thr Ser Thr	
	165 170 175
Phe Lys Leu Thr Asn Asn Thr Phe Ala Ala Gln Glu Gly Phe Gly Ala	
	180 185 190
Leu Ser Ile Ile Ser Ile Ser Pro Arg Phe Met Leu Thr Tyr Ser Asn	
	195 200 205
Ala Thr Asn Asp Val Gly Glu Gly Arg Phe Ser Lys Ser Glu Phe Cys	
	210 215 220
Met Asp Pro Ile Leu Ile Leu Met His Glu Leu Asn His Ala Met His	
	225 230 235 240
Asn Leu Tyr Gly Ile Ala Ile Pro Asn Asp Gln Thr Ile Ser Ser Val	
	245 250 255
Thr Ser Asn Ile Phe Tyr Ser Gln Tyr Asn Val Lys Leu Glu Tyr Ala	
	260 265 270
Glu Ile Tyr Ala Phe Gly Gly Pro Thr Ile Asp Leu Ile Pro Lys Ser	
	275 280 285
Ala Arg Lys Tyr Phe Glu Glu Lys Ala Leu Asp Tyr Tyr Arg Ser Ile	
	290 295 300
Ala Lys Arg Leu Asn Ser Ile Thr Thr Ala Asn Pro Ser Ser Phe Asn	
	305 310 315 320
Lys Tyr Ile Gly Glu Tyr Lys Gln Lys Leu Ile Arg Lys Tyr Arg Phe	
	325 330 335
Val Val Glu Ser Ser Gly Glu Val Thr Val Asn Arg Asn Lys Phe Val	
	340 345 350
Glu Leu Tyr Asn Glu Leu Thr Gln Ile Phe Thr Glu Phe Asn Tyr Ala	
	355 360 365
Lys Ile Tyr Asn Val Gln Asn Arg Lys Ile Tyr Leu Ser Asn Val Tyr	
	370 375 380
Thr Pro Val Thr Ala Asn Ile Leu Asp Asp Asn Val Tyr Asp Ile Gln	
	385 390 395 400
Asn Gly Phe Asn Ile Pro Lys Ser Asn Leu Asn Val Leu Phe Met Gly	
	405 410 415
Gln Asn Leu Ser Arg Asn Pro Ala Leu Arg Lys Val Asn Pro Glu Asn	
	420 425 430
Met Leu Tyr Leu Phe Thr Lys Phe Cys Val Asp Gly Ile Ile Thr Ser	
	435 440 445
Lys Thr Lys Ser Leu Ile Glu Gly Arg Phe Gly Gly Phe Thr Gly Ala	
	450 455 460
Arg Lys Ser Ala Arg Lys Arg Lys Asn Gln Ala Leu Ala Gly Gly Gly	
	465 470 475 480
Gly Ser Gly Gly Gly Ser Gly Gly Gly Gly Ser Ala Leu Val Leu	
	485 490 495
Gln Cys Arg Glu Leu Leu Val Lys Asn Thr Asp Leu Pro Phe Ile Gly	
	500 505 510
Asp Ile Ser Asp Val Lys Thr Asp Ile Phe Leu Arg Lys Asp Ile Asn	
	515 520 525

-continued

Glu	Glu	Thr	Glu	Val	Ile	Tyr	Tyr	Pro	Asp	Asn	Val	Ser	Val	Asp	Gln
530						535					540				
Val	Ile	Leu	Ser	Lys	Asn	Thr	Ser	Glu	His	Gly	Gln	Leu	Asp	Leu	Leu
545					550					555					560
Tyr	Pro	Ser	Ile	Asp	Ser	Glu	Ser	Glu	Ile	Leu	Pro	Gly	Glu	Asn	Gln
				565					570					575	
Val	Phe	Tyr	Asp	Asn	Arg	Thr	Gln	Asn	Val	Asp	Tyr	Leu	Asn	Ser	Tyr
			580					585					590		
Tyr	Tyr	Leu	Glu	Ser	Gln	Lys	Leu	Ser	Asp	Asn	Val	Glu	Asp	Phe	Thr
		595					600					605			
Phe	Thr	Arg	Ser	Ile	Glu	Glu	Ala	Leu	Asp	Asn	Ser	Ala	Lys	Val	Tyr
	610					615					620				
Thr	Tyr	Phe	Pro	Thr	Leu	Ala	Asn	Lys	Val	Asn	Ala	Gly	Val	Gln	Gly
625					630					635					640
Gly	Leu	Phe	Leu	Met	Trp	Ala	Asn	Asp	Val	Val	Glu	Asp	Phe	Thr	Thr
				645					650					655	
Asn	Ile	Leu	Arg	Lys	Asp	Thr	Leu	Asp	Lys	Ile	Ser	Asp	Val	Ser	Ala
			660					665					670		
Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Ser	Asn	Ser	Val	Arg
		675					680					685			
Arg	Gly	Asn	Phe	Thr	Glu	Ala	Phe	Ala	Val	Thr	Gly	Val	Thr	Ile	Leu
	690					695					700				
Leu	Glu	Ala	Phe	Pro	Glu	Phe	Thr	Ile	Pro	Ala	Leu	Gly	Ala	Phe	Val
705					710					715					720
Ile	Tyr	Ser	Lys	Val	Gln	Glu	Arg	Asn	Glu	Ile	Ile	Lys	Thr	Ile	Asp
				725					730					735	
Asn	Cys	Leu	Glu	Gln	Arg	Ile	Lys	Arg	Trp	Lys	Asp	Ser	Tyr	Glu	Trp
			740					745					750		
Met	Met	Gly	Thr	Trp	Leu	Ser	Arg	Ile	Ile	Thr	Gln	Phe	Asn	Asn	Ile
		755					760					765			
Ser	Tyr	Gln	Met	Tyr	Asp	Ser	Leu	Asn	Tyr	Gln	Ala	Gly	Ala	Ile	Lys
		770				775					780				
Ala	Lys	Ile	Asp	Leu	Glu	Tyr	Lys	Lys	Tyr	Ser	Gly	Ser	Asp	Lys	Glu
785					790					795					800
Asn	Ile	Lys	Ser	Gln	Val	Glu	Asn	Leu	Lys	Asn	Ser	Leu	Asp	Val	Lys
				805					810					815	
Ile	Ser	Glu	Ala	Met	Asn	Asn	Ile	Asn	Lys	Phe	Ile	Arg	Glu	Cys	Ser
			820					825					830		
Val	Thr	Tyr	Leu	Phe	Lys	Asn	Met	Leu	Pro	Lys	Val	Ile	Asp	Glu	Leu
			835				840					845			
Asn	Glu	Phe	Asp	Arg	Asn	Thr	Lys	Ala	Lys	Leu	Ile	Asn	Leu	Ile	Asp
					855						860				
Ser	His	Asn	Ile	Ile	Leu	Val	Gly	Glu	Val	Asp	Lys	Leu	Lys	Ala	Lys
865					870					875					880
Val	Asn	Asn	Ser	Phe	Gln	Asn	Thr	Ile	Pro	Phe	Asn	Ile	Phe	Ser	Tyr
				885					890					895	
Thr	Asn	Asn	Ser	Leu	Leu	Lys	Asp	Ile	Ile	Asn	Glu	Tyr	Phe	Asn	Leu
			900					905					910		

Asp

-continued

<211> LENGTH: 2673

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 68

```
ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgtgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaagggtgac    180
ctgaacccgc caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaagggtgta ctaaaactgtt cgagcgtatt    300
tactccaccg acctggggccg tatgtctctg actagcatcg ttccgggtat cccgttcttg    360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct    480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct caccgtaac    540
ggctacgggt ccaactcagta catccgttct tctccggact tcaccttcgg ttttgaagaa    600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgct    660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaagt taaccaaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttggcggc cagcagccta aattcatcga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaat ccatcgtggg taccactgct tctctccagt acatgaagaa cgtttttaa    960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaacttt aaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc   1140
gtgccgaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct   1200
gctaatttta acggccagaa caccgaaatc aacaacatga acttcacaaa actgaaaaac   1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa   1320
actaaatctc tgatagaagg tagatacggg gggttctctg cgctagcggg cgggtggcgg   1380
agcggcggtg gcggtagcgg cgggtggcgg agcgactag tgcgtagtg tatcaagggt   1440
aacaactggg atttattctt cagcccgagt gaagacaact tcaccaacga cctgaacaaa   1500
ggtgaagaaa tcacctcaga tactaacatc gaagcagccg aagaaaacat ctgcctggac   1560
ctgatccagc agtactacct gacctttaat ttcgacaacg agccggaaaa catttctatc   1620
gaaaacctga gctctgatat catcggccag ctggaactga tgccgaacat cgaacgtttc   1680
ccaaacggta aaaagtagca gctggacaaa tataccatgt tccactacct gcgcgcgcag   1740
gaatttgaac acggcaaatc ccgtatcgca ctgactaact ccgttaacga agctctgctc   1800
aaccgcgtcc gtgtatacac cttcttctct agcgactacg tgaaaaaggt caacaaagcg   1860
actgaagctg caatgttctt ggggtgggtt gaacagcttg tttatgattt taccgacgag   1920
acgtccgaag tatctactac cgacaaaatt gcggatatca ctatcatcat cccgtacatc   1980
ggtcgggctc tgaacattgg caacatgctg tacaagacg acttcgttgg cgcactgatc   2040
```

-continued

```

ttctccggtg cgggtgatcct gctggagttc atcccggaaa tcgccatccc ggtactgggc 2100
acctttgctc tggtttctta cattgcaaac aaggttctga ctgtacaaac catcgacaac 2160
gcgctgagca aacgtaacga aaaatgggat gaagtttaca aatatatcgt gaccaactgg 2220
ctggctaagg ttaatactca gatcgacctc atccgcaaaa aaatgaaaga agcactggaa 2280
aaccaggcgg aagctaccaa ggcaatcatt aactaccagt acaaccagta caccgaggaa 2340
gaaaaaaaca acatcaactt caacatcgac gatctgtcct ctaaaactgaa cgaatccatc 2400
aacaaagcta tgatcaacat caacaagttc ctgaaccagt gctctgtaag ctatctgatg 2460
aactccatga tcccgtacgg tgtaaactg ctggaggact tcgatgcgtc tctgaaagac 2520
gccctgctga aatacattta cgacaaccgt ggcactctga tcggtcaggt tgatcgtctg 2580
aaggacaaaag tgaacaatac cttatcgacc gacatccctt ttcagctcag taaatatgtc 2640
gataaccaac gccttttgtc cactctagac tag 2673

```

<210> SEQ ID NO 69

<211> LENGTH: 890

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 69

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15

```

```

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20           25           30

```

```

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35           40           45

```

```

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50           55           60

```

```

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65           70           75           80

```

```

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85           90           95

```

```

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100          105          110

```

```

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115          120          125

```

```

Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130          135          140

```

```

Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145          150          155          160

```

```

Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165          170          175

```

```

Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180          185          190

```

```

Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195          200          205

```

```

Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210          215          220

```

```

His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225          230          235          240

```

-continued

Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	245	250	255
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp	260	265	270
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	275	280	285
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	290	295	300
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	305	310	315
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	325	330	335
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	340	345	350
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	355	360	365
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	435	440	445
Tyr	Gly	Gly	Phe	Leu	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	450	455	460
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	465	470	475
Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	485	490	495
Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	500	505	510
Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	515	520	525
Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	530	535	540
Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	545	550	555
Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	565	570	575
Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	580	585	590
Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	595	600	605
Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	610	615	620
Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	625	630	635

-continued

Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile
 645 650 655
 Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys
 660 665 670
 Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu
 675 680 685
 Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu
 690 695 700
 Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn
 705 710 715 720
 Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile
 725 730 735
 Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg
 740 745 750
 Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala
 755 760 765
 Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn
 770 775 780
 Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile
 785 790 795 800
 Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val
 805 810 815
 Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu
 820 825 830
 Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp
 835 840 845
 Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val
 850 855 860
 Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val
 865 870 875 880
 Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp
 885 890

<210> SEQ ID NO 70

<211> LENGTH: 2709

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 70

```

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac      60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc      120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaaggtgac      180
ctgaacccgc caccggaagc gaaacagggtg cgggtatett actatgactc cacctacctg      240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactgtt cgagcgtatt      300
tactccaccg acctggggccg tatgtgctg actagcatcg ttgcggtat cccgttctgg      360
ggcggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt      420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct      480
gatatcatcc agttcagatg taagagcttt ggtcacgaag ttctgaacct caccgtaac      540

```

-continued

ggctacggtt ccactcagta catccgttcc tctccggact tcaccttcgg ttttgaagaa	600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg	660
gttacacctg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggc cacgacgcta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccactgtggg taccactgct tctctccagt acatgaagaa cgtttttaaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatatggc ggtttcacgg gcgcacgcaa atcagcgcgt	1380
aaattagcta accagcgctg agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt	1440
ggcggtagcg cactagtgtc gcagtgtatc aaggttaaca actgggattt attcttcagc	1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctacagatact	1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc	1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc	1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg	1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatcccgt	1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccgtgt atacaccttc	1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttgggg	1920
tgggttgaa acgtttgtta tgattttacc gacgagacgt ccgaagtatc tactaccgac	1980
aaaaattcgg atatcactat catcatcccg tacatcggtc cggctctgaa cattggcaac	2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct ccggtgcggg gatcctgctg	2100
gagttcatcc cggaaatcgc catcccggtg ctgggacact ttgctctggg ttcttacatt	2160
gcaaacaagg ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa	2220
tgggatgaag tttaacaata tatcgtgacc aactggctgg ctaagggtta tactcagatc	2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca	2340
atcattaact accagtagaa ccagtagacc gaggaagaaa aaaacaacat caacttcaac	2400
atcgacgac tgctctctaa actgaacgaa tccatcaaca aagctatgat caacatcaac	2460
aagttcctga accagtgtct tgtaagctat ctgatgaact ccatgatccc gtacgggtgt	2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgctgaaata catttacgac	2580
aaccgtggca ctctgatcgg tcagggtgat cgtctgaagg acaaaagtga caatacctta	2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct tttgtccact	2700
ctagactag	2709

-continued

```

<211> LENGTH: 902
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 71

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1          5          10          15

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125

Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140

Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160

Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175

Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190

Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205

Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220

His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240

Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245         250         255

Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
260         265         270

Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
275         280         285

Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser
290         295         300

Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys
305         310         315         320

Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp
325         330         335

Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr
340         345         350

Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr
355         360         365

```

-continued

Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380	
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395	400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430	
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	435	440	445	
Tyr	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Leu	Ala	Asn	450	455	460	
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	465	470	475	480
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	485	490	495	
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	500	505	510	
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	515	520	525	
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	530	535	540	
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	545	550	555	560
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	565	570	575	
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	580	585	590	
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	595	600	605	
Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	610	615	620	
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	625	630	635	640
Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	645	650	655	
Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	660	665	670	
Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	675	680	685	
Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	690	695	700	
Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	705	710	715	720
Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	725	730	735	
Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	740	745	750	
Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	755	760	765	

-continued

Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys Ala Ile Ile Asn Tyr
 770 775 780
 Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys Asn Asn Ile Asn Phe Asn
 785 790 795 800
 Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu Ser Ile Asn Lys Ala Met
 805 810 815
 Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys Ser Val Ser Tyr Leu Met
 820 825 830
 Asn Ser Met Ile Pro Tyr Gly Val Lys Arg Leu Glu Asp Phe Asp Ala
 835 840 845
 Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr
 850 855 860
 Leu Ile Gly Gln Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu
 865 870 875 880
 Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg
 885 890 895
 Leu Leu Ser Thr Leu Asp
 900

<210> SEQ ID NO 72
 <211> LENGTH: 2709
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 72

```

ggatccatgg agttcgtaa caaacagttc aactataaag acccagtaa cgggtgtgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc    120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac    180
ctgaacccgc caccggaagc gaaacagggt cgggtatctt actatgactc cacctacctg    240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactggt cgagcgtatt    300
tactccaccg acctgggccg tatgctgctg actagcatcg ttcgcggtat cccgttctgg    360
ggcggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt    420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgateatcgg cccgtctgct    480
gatatcatcc agttcgagtg taagagcttt ggtcacgaag ttctgaacct caccgtaac    540
ggctacgggt cactcagta catccgtttc tctccggact tcaccttcgg ttttgaagaa    600
tcccgtgaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgctg    660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat    720
ccgaaccgtg tcttcaaaat taacaccaac gcgtattacg agatgtccgg tctggaagtt    780
agcttcgaag aactgcgtac ttttggcggt cagcagccta aattcatcga ctctctgcaa    840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac    900
aaagcgaaat ccatcggtgg taccactgct tctctccagt acatgaagaa cgttttttaa    960
gaaaataacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc   1020
gataaaactt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt   1080
aaagttctga accgcaaac ctatctgaac ttgcacaagg cagtattcaa aatcaacatc   1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct   1200
  
```

-continued

```

gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac 1260
ttcactgggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa 1320
actaaatctc tgatagaagg tagatatggc gggtttcacgg gcgcacgcaa atcagcgcgt 1380
aaacgtaaga accagggcgt agcggggcgt ggcggtagcg gcggtggcgg tagcggcgggt 1440
ggcggtagcg cactagtgtc gcagtgtatc aaggtttaaca actgggattt attcttcagc 1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctacagatact 1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc 1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcacc 1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg 1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatcccg 1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccggtg atacaccttc 1860
ttctctagcg actacgtgaa aaagggtcaac aaagcgactg aagctgcaat gttcttggtt 1920
tgggttgaa agcttgttta tgattttacc gacgagacgt ccgaagtatc tactaccgac 1980
aaaaattgcg atatcactat catcatcccg tacatcggtc cggctctgaa cattggcaac 2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct cgggtgcggg gatectgctg 2100
gagttcatcc cggaaatcgc catcccggtc ctgggcacct ttgctctggt ttcttacatt 2160
gcaacaacgg ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa 2220
tgggatgaag ttacaaaata tatcgtgacc aactggctgg ctaagggtta tactcagatc 2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgac tgctctctaa actgaacgaa tocatcaaca aagctatgat caacatcaac 2460
aagttctctga accagtgtc tgtaagctat ctgatgaact ccatgatccc gtacgggtgtt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgctgaaata catttacgac 2580
aaccggtggca ctctgatcgg tcaggttgat cgtctgaagg acaaagtga caatacetta 2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct tttgtccact 2700
ctagactag 2709

```

<210> SEQ ID NO 73

<211> LENGTH: 902

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 73

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1           5           10           15

```

```

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20           25           30

```

```

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35           40           45

```

```

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50           55           60

```

```

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu

```

-continued

65	70	75	80
Ser Thr Asp Asn Glu 85	Lys Asp Asn Tyr Leu 90	Lys Gly Val Thr Lys Leu 95	
Phe Glu Arg Ile Tyr 100	Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser 110		
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu 115	120	125	
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly 130	135	140	
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala 145	150	155	160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn 165	170	175	
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro 180	185	190	
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro 195	200	205	
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala 210	215	220	
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn 225	230	235	240
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser 245	250	255	
Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp 260	265	270	
Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr 275	280	285	
Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser 290	295	300	
Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys 305	310	315	320
Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp 325	330	335	
Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr 340	345	350	
Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr 355	360	365	
Leu Asn Phe Asp Lys Ala Val Phe Lys Ile Asn Ile Val Pro Lys Val 370	375	380	
Asn Tyr Thr Ile Tyr Asp Gly Phe Asn Leu Arg Asn Thr Asn Leu Ala 385	390	395	400
Ala Asn Phe Asn Gly Gln Asn Thr Glu Ile Asn Asn Met Asn Phe Thr 405	410	415	
Lys Leu Lys Asn Phe Thr Gly Leu Phe Glu Phe Tyr Lys Leu Leu Cys 420	425	430	
Val Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly Arg 435	440	445	
Tyr Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Arg Lys Asn 450	455	460	
Gln Ala Leu Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly 465	470	475	480

Gly 485	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp
Leu 500	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys
Gly 515	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn
Ile 530	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp
Asn 545	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile
Gly 565	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys
Lys 580	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln
Glu 595	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn
Glu 610	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp
Tyr 625	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly
Trp 645	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val
Ser 660	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile
Gly 675	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val
Gly 690	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro
Glu 705	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile
Ala 725	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys
Arg 740	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp
Leu 755	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys
Glu 770	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr
Gln 785	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn
Ile 805	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met
Ile 820	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met
Asn 835	Ser	Met	Ile	Pro	Tyr	Gly	Val	Arg	Leu	Glu	Asp	Phe	Asp	Ala	
Ser 850	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr
Leu 865	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu

-continued

Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg
885 890 895

Leu Leu Ser Thr Leu Asp
900

<210> SEQ ID NO 74

<211> LENGTH: 2889

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 74

```

ggatccttgg tacgagatga cgttgactat caaatTTTcc gcgactttgc ggaaaataaa      60
ggtaagtttt tcgtcgccgc cacagacctg tccgtcaaaa ataagagagg ccagaacatc     120
ggtaacgcac tgagcaacgt ccctatgatt gatttttagtg tagcggacgt taataaacgg     180
attgcaaccg tcgttgatcc gcagtatgct gtcagcgtca aacatgctaa agcgggaagtt     240
catacgttct attacgggca atataacggc cataacgatg tggctgataa agaaaatgaa     300
tatcgcggtgg tcgagcagaa caattacgaa ccgcacaaag cgtggggcgc gagtaattta     360
ggccgccttg aggactataa catggcccgt ttcaataaat tcgtgaccga ggtagcaccg     420
atcgccccc cagatgctgg tgggggcctg gataacctaca aagataaaaa ccgcttctct     480
agcttcgtgc gcattggcgc cggtcgtcag ctcgtgtacg agaaggggtg ctatcaccag     540
gaaggtaatg aaaaggggta cgacctccgt gatttgtccc aggcgtatcg ctacgctatt     600
gccggaaccc cgtataaaga tattaatata gatcaaacca tgaataccga aggcctaatt     660
ggtttcggga atcataataa gcaatatagc gcagaagagc taaagcaggc cctcagccaa     720
gatgcgttaa ccaattacgg agtggttaggc gatagcggca gtccgctgtt tgccttcgat     780
aaacagaaaa atcaatgggt gtttctgggc acttatgatt attgggccgg atatggtaaa     840
aagagctggc aggaatggaa tatttataaa aaggaattcg cagacaaaat caagcagcat     900
gacaacgcag gtacggtgaa ggggaacggc gaacatcact ggaagacgac cggcacgaat     960
agtcataatc gatcgacggc cgttcgcctg gcgaacaatg agggcgatgc aaacaatggg    1020
caaaacgtga cttttgagga caacgggtacc ctggtcctta accagaacat aaatcagggc    1080
gcgggaggct tgttctttaa aggcgactat actgttaagg gagcaaaaa tgacatcacc    1140
tggttagggg ccggtattga cgttgcggat ggaaaaaagg tggtttgga ggttaaaaac    1200
cctaacgggg accggtggc aaaaatcggc aaaggacat tggaaattaa tggtaaccgg    1260
gtgaatcagg gtcagctgaa agtgggagat gggaccgtga ttctgaacca gaaagcagac    1320
gctgacaaaa aggtgcaagc ctttagccaa gtaggaattg ttagtggtcg tggcacactc    1380
gtcttgaaat caagcaacca aataaatccg gataacctgt actttggatt tcgtggcgga    1440
cgctgggatg ctaacgggaa tgatctgacc tttgaacata tccgtaacgt tgacgagggt    1500
gcgcgcatag ttaatcataa tactgaccat gcataacta tcaccttgac cgggaaaagt    1560
ctgattacaa acccaaaact tctgtcagta cattccatcc agaataatta tgatgaagac    1620
gattactcat actattaccg gccgcgtaga ccaattccac aaggtaaaga tctttattac    1680
aaaaattacc gttattacgc attaaaatcc ggaggcggc tgaatgcacc tatgccgga    1740
aatggcgtgg ccgaaaaaaa tgactggatt tttatgggtt atactcaaga agaggctcgc    1800

```

-continued

aaaaatgcaa tgaaccataa aaataaccga aggatcggtg atttcggcgg atttttcgat	1860
gaggaaaatg gtaaagggtca caatgggtgcg ctgaatctaa attttaacgg caaaagtgcc	1920
cagaaaacgtt tccttctgac tgggtggcgt aatctgaatg gtaaaatcag tgtgacgcag	1980
ggtaacgtgc tgctttctgg ccggccaact ccgcatgcac gtgattttgt aaataaatcg	2040
agcgctcgta aagatgcgca tttttctaaa aataacgagg tcgtgtttga agatgactgg	2100
ataaatcgca cctttaaacg gccagaaaac gcggttaac agagtgcgag cttttcatcg	2160
ggtaggaatg tatctgatat tacagcaaac attacagcca ctgataatgc gaaggccaac	2220
ctgggttata aaaacggtga tgaagtttgt gttcgatcgg attacacggg ctatgttacc	2280
tgcaacactg gcaatctgtc tgataaacg cttactctt ttgacgccac gcgcattaac	2340
gggaatgtga acctgaacca aaacgctgcc ttggtacttg gtaaggccgc gttgtgggg	2400
aaaattcagg gccagggcaa ctcccggtg tctctgaacc agcactcgaa gtggcacctg	2460
acgggggact cgcaggtgca caacttgcc ctggccgata gccatattca ccttaacaat	2520
gcgtccgatg cccagtcagc taataaatat catacgatca aaatcaatca cctctctggc	2580
aacggtcact ttcactactt aacggattta gcaaaaaact taggggataa agtctctgta	2640
aaagaatcag cgagcggaca ttatcagtta catgtacaga acaaacagg cgagccaaat	2700
caggaaggcc ttgacttatt tgatgcttca tcggtacaag atcgttccag actgttcgtt	2760
tcactcgca atcactactg tgatctgggt gcgctgcgt atactataaa gacggaaaat	2820
ggcataacac gcctctataa tccctatgcc ggtaacggcc gtcgggtgaa acctgctccc	2880
tgcgtcgac	2889

<210> SEQ ID NO 75
 <211> LENGTH: 4296
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 75

ggatccttgg tacgagatga cgttgactat caaattttcc gcgactttgc ggaaaataaa	60
ggtaagtttt tcgtcggcgc cacagacctg tccgtcaaaa ataagagagg ccagaacatc	120
ggtaacgcac tgagcaacgt ccctatgatt gatttttagt tagcggacgt taataaacgg	180
attgcaacgc tcgttgatcc gcagtatgct gtcagcgtea aacatgctaa agcgggaagt	240
catacgttct attacgggca atataacggc cataacgatg tggctgataa agaaaatgaa	300
tatcgcgtgg tcgagcagaa caattacgaa ccgcacaaag cgtggggcgc gagtaattta	360
ggccgcctgg aggactataa catggcccgt ttcaataaat tcgtgaccga ggtagcaccg	420
atcgccccc cagatgctgg tgggggcctg gatacctaca aagataaaaa ccgcttctct	480
agcttcgtgc gcattggcgc cggctcgtcag ctctgttacg agaaggggtgt ctatcaccag	540
gaaggtaatg aaaaggggta cgacctccgt gatttgtccc aggcgtatcg ctacgctatt	600
gccggaaccc cgtataaaga tattaatatc gatcaaacca tgaataccga aggcctaatt	660
ggtttcggga atcataataa gcaatatagc gcagaagagc taaagcaggc cctcagccaa	720
gatgcgttaa ccaattacgg agtggttagc gatagcggca gtcgctgtt tgccttcgat	780
aaacagaaaa atcaatgggt gtttctgggc acctatgatt attgggccc ataggtaaa	840

-continued

aagagctggc	aggaatggaa	tatttataaa	aaggaattcg	cagacaaaat	caagcagcat	900
gacaacgcag	gtacggtgaa	ggggaacggc	gaacatcact	ggaagacgac	cggcacgaat	960
agtcatatcg	gatcgacggc	cgttcgccctg	gcgaacaatg	aggcgcatgc	aaacaatggg	1020
caaaacgtga	cctttgagga	caacgggtacc	ctggtcctta	accagaacat	aatcagggc	1080
gcgggaggct	tgttctttaa	aggcgactat	actgttaagg	gagcaaaaa	tgacatcacc	1140
tggttagggg	ccggtattga	cgttgcggat	ggaaaaaagg	tggtttgcca	ggttaaaaa	1200
cctaacgggg	accggctggc	aaaaatcgcc	aaaggacat	tggaaattaa	tggtagcggg	1260
gtgaatcagg	gtcagctgaa	agtgggagat	gggaccgtga	ttctgaacca	gaaagcagac	1320
gctgacaaaa	aggtgcaagc	ctttagccaa	gtaggaattg	ttagtggtcg	tggcacactc	1380
gtcttgaaat	caagcaacca	aataaatccg	gataacctgt	actttggatt	tctgtggcga	1440
cgctgggatg	ctaacgggaa	tgatctgacc	tttgaacata	tccgtaacgt	tgacgagggg	1500
gcgcgcatag	ttaatcataa	tactgaccat	gcatacaata	tcaccttgac	cgggaaaagt	1560
ctgattacaa	acccaaactc	tctgtcagta	cattccatcc	agaatgatta	tgatgaagac	1620
gattactcat	actattaccg	gccgcgtaga	ccaattccac	aaggtaaaga	tctttattac	1680
aaaaattacc	gttattacgc	attaaaatcc	ggaggggcgc	tgaatgcacc	tatgccggaa	1740
aatggcgtgg	ccgaaaaaaa	tgactggatt	tttatgggtt	atactcaaga	agaggctcgc	1800
aaaaatgcaa	tgaaccataa	aaataaccga	aggatcggtg	atttcggcgg	atttttcgat	1860
gaggaaaatg	gtaaagggtc	caatgggtcg	ctgaatctaa	attttaacgg	caaaagtggc	1920
cagaaacgtt	tccttctgac	tgggtggcgt	aatctgaatg	gtaaaatcag	tgtgacgcag	1980
ggtaacgtgc	tgttttctgg	ccggccaact	ccgcatgcac	gtgattttgt	aaataaatcg	2040
agcgctcgta	aagatgcgca	tttttctaaa	aataacgagg	tcgtgtttga	agatgactgg	2100
ataaatcgca	cctttaaagc	ggcagaaatc	gcggttaatc	agagtgcgag	cttttcatcg	2160
ggtaggaatg	tatctgatat	tacagcaaac	attacagcca	ctgataatgc	gaagggtcaac	2220
ctgggttata	aaaacgggtg	tgaagtttgt	gttcgatcgg	attacacggg	ctatgttacc	2280
tgcaacactg	gcaatctgtc	tgataaagcg	cttaactctt	ttgacgccac	gcgcattaac	2340
gggaatgtga	acctgaacca	aaacgctgcc	ttgggtactg	gtaaggccgc	gttggtgggt	2400
aaaattcagg	gccagggcaa	ctcccggtgt	tctctgaacc	agcactcgaa	gtggcacctg	2460
acgggggact	cgcaggtgca	caacttgtcc	ctggccgata	gccatattca	ccttaacaat	2520
gcgtccgatg	cccagtcagc	taataaatat	catacgatca	aaatcaatca	cctctctggc	2580
aacggtcact	ttcactactt	aacggattta	gcaaaaaact	taggggataa	agtccctggta	2640
aaagaatcag	cgagcggaca	ttatcagtta	catgtacaga	acaaaacagg	cgagccaaat	2700
caggaaggcc	ttgacttatt	tgatgcttca	tcggtacaag	atcgttccag	actgttcggt	2760
tcaactcgca	atcactacgt	tgatctgggt	gcgctgcgct	atactataaa	gacggaaaat	2820
ggcataaac	gcctctataa	tccctatgcc	ggtaacggcc	gtccgggtgaa	acctgtccc	2880
tgcgctgacg	gcatcattac	ctccaaaact	aatctctga	tagaaggtag	atttgccggg	2940
ttcacgggcg	cacgcaaatc	agcgcgtaaa	cgtaagaacc	aggcgctagc	gggagggtggc	3000
ggtagcggcg	gtggcggtag	cggcggtggc	ggtagcgcac	tagtgctgca	gtgtatcaag	3060
gttaacaact	gggatttatt	cttcagcccc	agtgaagaca	acttcaccaa	cgacctgaac	3120

-continued

```

aaaggtgaag aaatcacctc agatactaac atcgaagcag ccgaagaaaa catctcgctg 3180
gacctgatcc agcagtacta cctgaccttt aatttcgaca acgagccgga aaacatttct 3240
atcgaaaacc tgagctctga tatcatcggc cagctggaac tgatgccgaa catcgaaact 3300
ttcccaaacg gtaaaaagta cgagctggac aaatatacca tgttccacta cctgcgcgcg 3360
caggaatttg aacacggcaa atcccgatc gcactgacta actccgttaa cgaagctctg 3420
ctcaacccgt cccgtgtata caccttcttc tctagcgact acgtgaaaaa ggtcaacaaa 3480
gcgactgaag ctgcaatggt cttgggttgg gttgaacagc ttgtttatga ttttaccgac 3540
gagacgtccg aagtatctac taccgacaaa attgcggata tcactatcat catcccgta 3600
atcggtccg ctctgaacat tggcaacatg ctgtacaaag acgacttcgt tggcgactg 3660
atcttctccg gtgcggtgat cctgctggag ttcattcccg aaatcgccat cccggtactg 3720
ggcacctttg ctctggtttc ttacattgca aacaagggtc tgactgtaca aaccatcgac 3780
aacgcgctga gcaaacgtaa cgaataatgg gatgaagttt acaaatatat cgtgaccaac 3840
tggttggtga aggttaatac tcagatcgac ctcatccgca aaaaaatgaa agaagcactg 3900
gaaaaccagg cggaagctac caaggcaatc attaactacc agtacaacca gtacaccgag 3960
gaagaaaaaa acaacatcaa cttcaacatc gacgatctgt cctctaaact gaacgaatcc 4020
atcaacaag ctatgatcaa catcaacaag ttctgaacc agtgctctgt aagctatctg 4080
atgaactcca tgatcccgta cgggtgttaa cgtctggagg acttcgatgc gtctctgaaa 4140
gacgcctgc tgaatacat ttacgacaac cgtggcactc tgatcggtca ggttgatcgt 4200
ctgaaggaca aagtgaacaa taccttatcg accgacatcc cttttcagct cagtaaatat 4260
gtcgataacc aacgcctttt gtccactcta gactag 4296

```

<210> SEQ ID NO 76

<211> LENGTH: 1431

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 76

```

Gly Ser Leu Val Arg Asp Asp Val Asp Tyr Gln Ile Phe Arg Asp Phe
1           5           10           15

```

```

Ala Glu Asn Lys Gly Lys Phe Phe Val Gly Ala Thr Asp Leu Ser Val
20           25           30

```

```

Lys Asn Lys Arg Gly Gln Asn Ile Gly Asn Ala Leu Ser Asn Val Pro
35           40           45

```

```

Met Ile Asp Phe Ser Val Ala Asp Val Asn Lys Arg Ile Ala Thr Val
50           55           60

```

```

Val Asp Pro Gln Tyr Ala Val Ser Val Lys His Ala Lys Ala Glu Val
65           70           75           80

```

```

His Thr Phe Tyr Tyr Gly Gln Tyr Asn Gly His Asn Asp Val Ala Asp
85           90           95

```

```

Lys Glu Asn Glu Tyr Arg Val Val Glu Gln Asn Asn Tyr Glu Pro His
100          105          110

```

```

Lys Ala Trp Gly Ala Ser Asn Leu Gly Arg Leu Glu Asp Tyr Asn Met
115          120          125

```

```

Ala Arg Phe Asn Lys Phe Val Thr Glu Val Ala Pro Ile Ala Pro Thr
130          135          140

```

-continued

Asp	Ala	Gly	Gly	Gly	Leu	Asp	Thr	Tyr	Lys	Asp	Lys	Asn	Arg	Phe	Ser	145	150	155	160
Ser	Phe	Val	Arg	Ile	Gly	Ala	Gly	Arg	Gln	Leu	Val	Tyr	Glu	Lys	Gly	165	170	175	
Val	Tyr	His	Gln	Glu	Gly	Asn	Glu	Lys	Gly	Tyr	Asp	Leu	Arg	Asp	Leu	180	185	190	
Ser	Gln	Ala	Tyr	Arg	Tyr	Ala	Ile	Ala	Gly	Thr	Pro	Tyr	Lys	Asp	Ile	195	200	205	
Asn	Ile	Asp	Gln	Thr	Met	Asn	Thr	Glu	Gly	Leu	Ile	Gly	Phe	Gly	Asn	210	215	220	
His	Asn	Lys	Gln	Tyr	Ser	Ala	Glu	Glu	Leu	Lys	Gln	Ala	Leu	Ser	Gln	225	230	235	240
Asp	Ala	Leu	Thr	Asn	Tyr	Gly	Val	Leu	Gly	Asp	Ser	Gly	Ser	Pro	Leu	245	250	255	
Phe	Ala	Phe	Asp	Lys	Gln	Lys	Asn	Gln	Trp	Val	Phe	Leu	Gly	Thr	Tyr	260	265	270	
Asp	Tyr	Trp	Ala	Gly	Tyr	Gly	Lys	Lys	Ser	Trp	Gln	Glu	Trp	Asn	Ile	275	280	285	
Tyr	Lys	Lys	Glu	Phe	Ala	Asp	Lys	Ile	Lys	Gln	His	Asp	Asn	Ala	Gly	290	295	300	
Thr	Val	Lys	Gly	Asn	Gly	Glu	His	His	Trp	Lys	Thr	Thr	Gly	Thr	Asn	305	310	315	320
Ser	His	Ile	Gly	Ser	Thr	Ala	Val	Arg	Leu	Ala	Asn	Asn	Glu	Gly	Asp	325	330	335	
Ala	Asn	Asn	Gly	Gln	Asn	Val	Thr	Phe	Glu	Asp	Asn	Gly	Thr	Leu	Val	340	345	350	
Leu	Asn	Gln	Asn	Ile	Asn	Gln	Gly	Ala	Gly	Gly	Leu	Phe	Phe	Lys	Gly	355	360	365	
Asp	Tyr	Thr	Val	Lys	Gly	Ala	Asn	Asn	Asp	Ile	Thr	Trp	Leu	Gly	Ala	370	375	380	
Gly	Ile	Asp	Val	Ala	Asp	Gly	Lys	Lys	Val	Val	Trp	Gln	Val	Lys	Asn	385	390	395	400
Pro	Asn	Gly	Asp	Arg	Leu	Ala	Lys	Ile	Gly	Lys	Gly	Thr	Leu	Glu	Ile	405	410	415	
Asn	Gly	Thr	Gly	Val	Asn	Gln	Gly	Gln	Leu	Lys	Val	Gly	Asp	Gly	Thr	420	425	430	
Val	Ile	Leu	Asn	Gln	Lys	Ala	Asp	Ala	Asp	Lys	Lys	Val	Gln	Ala	Phe	435	440	445	
Ser	Gln	Val	Gly	Ile	Val	Ser	Gly	Arg	Gly	Thr	Leu	Val	Leu	Asn	Ser	450	455	460	
Ser	Asn	Gln	Ile	Asn	Pro	Asp	Asn	Leu	Tyr	Phe	Gly	Phe	Arg	Gly	Gly	465	470	475	480
Arg	Leu	Asp	Ala	Asn	Gly	Asn	Asp	Leu	Thr	Phe	Glu	His	Ile	Arg	Asn	485	490	495	
Val	Asp	Glu	Gly	Ala	Arg	Ile	Val	Asn	His	Asn	Thr	Asp	His	Ala	Ser	500	505	510	
Thr	Ile	Thr	Leu	Thr	Gly	Lys	Ser	Leu	Ile	Thr	Asn	Pro	Asn	Ser	Leu	515	520	525	
Ser	Val	His	Ser	Ile	Gln	Asn	Asp	Tyr	Asp	Glu	Asp	Asp	Tyr	Ser	Tyr	530	535	540	

-continued

Tyr 545	Tyr	Arg	Pro	Arg 550	Arg	Pro	Ile	Pro	Gln	Gly 555	Lys	Asp	Leu	Tyr	Tyr 560
Lys	Asn	Tyr	Arg	Tyr 565	Tyr	Ala	Leu	Lys	Ser 570	Gly	Gly	Arg	Leu	Asn	Ala 575
Pro	Met	Pro	Glu 580	Asn	Gly	Val	Ala	Glu 585	Asn	Asn	Asp	Trp	Ile	Phe	Met 590
Gly	Tyr	Thr 595	Gln	Glu	Glu	Ala	Arg 600	Lys	Asn	Ala	Met	Asn 605	His	Lys	Asn
Asn	Arg 610	Arg	Ile	Gly	Asp	Phe 615	Gly	Gly	Phe	Phe	Asp 620	Glu	Glu	Asn	Gly
Lys	Gly 625	His	Asn	Gly	Ala 630	Leu	Asn	Leu	Asn	Phe 635	Asn	Gly	Lys	Ser	Ala 640
Gln	Lys	Arg	Phe 645	Leu	Leu	Thr	Gly	Gly	Ala 650	Asn	Leu	Asn	Gly	Lys	Ile 655
Ser	Val	Thr 660	Gln	Gly	Asn	Val	Leu	Leu	Ser 665	Gly	Arg	Pro	Thr	Pro	His 670
Ala	Arg 675	Asp	Phe	Val	Asn	Lys	Ser 680	Ser	Ala	Arg	Lys	Asp 685	Ala	His	Phe
Ser	Lys 690	Asn	Asn	Glu	Val	Val	Phe 695	Glu	Asp	Asp	Trp 700	Ile	Asn	Arg	Thr
Phe	Lys 705	Ala	Ala	Glu	Ile 710	Ala	Val	Asn	Gln	Ser 715	Ala	Ser	Phe	Ser	Ser 720
Gly	Arg	Asn 725	Val	Ser	Asp	Ile	Thr	Ala	Asn 730	Ile	Thr	Ala	Thr	Asp	Asn 735
Ala	Lys	Val 740	Asn	Leu	Gly	Tyr	Lys 745	Asn	Gly	Asp	Glu	Val	Cys	Val	Arg 750
Ser	Asp 755	Tyr	Thr	Gly	Tyr	Val	Thr 760	Cys	Asn	Thr	Gly	Asn 765	Leu	Ser	Asp
Lys	Ala 770	Leu	Asn	Ser	Phe	Asp 775	Ala	Thr	Arg	Ile	Asn 780	Gly	Asn	Val	Asn
Leu	Asn 785	Gln	Asn	Ala	Ala 790	Leu	Val	Leu	Gly	Lys 795	Ala	Ala	Leu	Trp	Gly 800
Lys	Ile	Gln	Gly	Gln 805	Gly	Asn	Ser	Arg	Val 810	Ser	Leu	Asn	Gln	His	Ser 815
Lys	Trp	His 820	Leu	Thr	Gly	Asp	Ser	Gln	Val 825	His	Asn	Leu	Ser	Leu	Ala 830
Asp	Ser 835	His	Ile	His	Leu	Asn	Asn 840	Ala	Ser	Asp	Ala	Gln 845	Ser	Ala	Asn
Lys	Tyr 850	His	Thr	Ile	Lys	Ile 855	Asn	His	Leu	Ser	Gly 860	Asn	Gly	His	Phe
His	Tyr 865	Leu	Thr	Asp	Leu 870	Ala	Lys	Asn	Leu	Gly 875	Asp	Lys	Val	Leu	Val 880
Lys	Glu	Ser	Ala	Ser 885	Gly	His	Tyr	Gln	Leu 890	His	Val	Gln	Asn	Lys	Thr 895
Gly	Glu	Pro 900	Asn	Gln	Glu	Gly	Leu 905	Asp	Leu	Phe	Asp	Ala	Ser	Ser	Val 910
Gln	Asp 915	Arg	Ser	Arg	Leu	Phe	Val 920	Ser	Leu	Ala	Asn	His 925	Tyr	Val	Asp
Leu	Gly 930	Ala	Leu	Arg	Tyr	Thr 935	Ile	Lys	Thr	Glu	Asn 940	Gly	Ile	Thr	Arg
Leu	Tyr	Asn	Pro	Tyr	Ala	Gly	Asn	Gly	Arg	Pro	Val	Lys	Pro	Ala	Pro

-continued

945	950					955					960				
Cys Val Asp Gly Ile Ile Thr Ser Lys Thr Lys Ser Leu Ile Glu Gly	965					970					975				
Arg Phe Gly Gly Phe Thr Gly Ala Arg Lys Ser Ala Arg Lys Arg Lys	980					985					990				
Asn Gln Ala Leu Ala Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly	995					1000					1005				
Gly Gly Gly Ser Ala Leu Val Leu Gln Cys Ile Lys Val Asn Asn	1010					1015					1020				
Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp Asn Phe Thr Asn Asp	1025					1030					1035				
Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr Asn Ile Glu Ala	1040					1045					1050				
Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln Gln Tyr Tyr Leu	1055					1060					1065				
Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile Ser Ile Glu Asn	1070					1075					1080				
Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro Asn Ile	1085					1090					1095				
Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr Thr	1100					1105					1110				
Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser	1115					1120					1125				
Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala Leu Leu Asn Pro	1130					1135					1140				
Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr Val Lys Lys Val	1145					1150					1155				
Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly Trp Val Glu Gln	1160					1165					1170				
Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu Val Ser Thr Thr	1175					1180					1185				
Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro Tyr Ile Gly Pro	1190					1195					1200				
Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp Asp Phe Val Gly	1205					1210					1215				
Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu Glu Phe Ile Pro	1220					1225					1230				
Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala Leu Val Ser Tyr	1235					1240					1245				
Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala Leu	1250					1255					1260				
Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile Val	1265					1270					1275				
Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile Arg	1280					1285					1290				
Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr Lys	1295					1300					1305				
Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys	1310					1315					1320				
Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn	1325					1330					1335				

-continued

Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn	
1340 1345 1350	
Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly	
1355 1360 1365	
Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu	
1370 1375 1380	
Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val	
1385 1390 1395	
Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile	
1400 1405 1410	
Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser	
1415 1420 1425	
Thr Leu Asp	
1430	

<210> SEQ ID NO 77

<211> LENGTH: 1357

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 77

gctagcgggc ggtggcggta gcggcgggtg cggtagcggc ggtggcggta ggcactagt	60
gctgcagtgt atcaaggta acaactggga tttattcttc agcccgagtg aagacaactt	120
caccaacgac ctgaacaaag gtgaagaaat cacctcagat actaacatcg aagcagccga	180
agaaaacatc tcgctggacc tgatccagca gtactacctg acctttaatt tcgacaacga	240
gccggaaaac atttctatcg aaaacctgag ctctgatatc atcggccagc tggaaactgat	300
gccgaacatc gaacgtttcc caaacggtaa aaagtacgag ctggacaaat ataccatgtt	360
ccactacctg cgcgcgcagg aatttgaaca cggcaaatcc cgtatcgac tgactaaactc	420
cgttaacgaa gctctgctca acccgctccg tgtatacacc ttcttctcta gcgactacgt	480
gaaaaaggtc aacaaagcga ctgaagctgc aatgttcttg ggttgggttg aacagcttgt	540
ttatgatattt accgacgaga cgtccgaagt atctactacc gacaaaattg cggatatcac	600
tatcatcatc ccgtacatcg gtccggctct gaacattggc aacatgctgt acaaagacga	660
cttcgttggc gcactgatct tctccgtg cgtgatctct ctggagtcca tcccggaat	720
cgccatcccc gtactgggca cctttgctct gggttcttac attgcaaaca aggttctgac	780
tgtacaaacc atcgacaacg cgctgagcaa acgtaacgaa aaatgggatg aagtttacia	840
atatatcggt accaactggc tggctaaggt taatactcag atcgacctca tccgcaaaaa	900
aatgaaagaa gcactggaaa accaggcgga agctaccaag gcaatcatta actaccagta	960
caaccagtac accgaggaag aaaaaacaa catcaacttc aacatcgacg atctgtcttc	1020
taaaactgaac gaatccatca acaaagctat gatcaacatc aacaagttcc tgaaccagt	1080
ctctgtaagc tatctgatga actccatgat cccgtacggg gttaaactgc tggaggactt	1140
cgatgcgtct ctgaaagacg cctgtctgaa atacatttac gacaaccgtg gcaactctgat	1200
cggtcagggt gatcgtctga aggacaaagt gaacaatacc ttatcgaccg acatcccttt	1260
tcagctcagt aaatatgtcg ataaccaacg ccttttgtcc actctagaaa tagaaggtag	1320

-continued

aagtgggcac catcaccatc accattaatg aaagctt 1357

<210> SEQ ID NO 78

<211> LENGTH: 2745

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 78

```

ggatccatgg agttcggttaa caaacagttc aactataaag acccagttaa cgtgtgtgac    60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc   120
cacacaacaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac   180
ctgaaccgcg caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg   240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt   300
tactccaccg acctggggcg tatgtgctg actagcatcg ttccgggtat cccgttctgg   360
ggcggttcta ccatcgatgc cgaactgaaa gtaatcgaca ctaactgcat caacgttatt   420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct   480
gatatcatcc agttcgatg taagagcttt ggtcacgaag ttctgaacct cacccgtaac   540
ggctacgggt ccaactcagta catccgttct tctccggact tcaccttcgg ttttgaagaa   600
tccttggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgct   660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat   720
ccgaaccctg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt   780
agcttcgaag aactgcgtac ttttgccggt cacgacgcta aattcatcga ctctctgcaa   840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac   900
aaagcgaaat ccatcgtggg taccactgct tctctccagt acatgaagaa cgtttttaaa   960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc  1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt  1080
aaagttctga accgcaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc  1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaaacac caacctggct  1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac  1260
ttcactggtc tgttcgatgt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa  1320
actaaatctc tgatagaagg tagatttggt gggttcacgg gcgcacgcaa atcagcgctg  1380
aaacgtaaga accagggcgt agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt  1440
ggcggtagcg cactagtgtc gcagtgtatc aaggttaaca actgggattt attcttcagc  1500
ccgagtgaag acaacttcac caacgacctg aacaaagggtg aagaaatcac ctacgatact  1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc  1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc  1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg  1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaaat ttgaacacgg caaatcccgt  1800
atcgactga ctaactccgt taacgaagct ctgctcaacc cgtcccgtgt atacaccttc  1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttggtt  1920

```

-continued

```

tggggtgaac agcttggttta tgattttacc gacgagacgt ccgaagtatc tactaccgac 1980
aaaattgctg atactactat catcatcccg tacatcggtc cggctctgaa cattggcaac 2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct cgggtgctgt gatcctgctg 2100
gagttcatcc cggaaatcgc catcccggtg ctgggcacct ttgctctggt ttcttacatt 2160
gcaaacaagg ttctgactgt acaaaccatc gacaacgcgc tgagcaaacg taacgaaaaa 2220
tgggatgaag ttacaaata tatcgtgacc aactggctgg ctaagggtta tactcagatc 2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca 2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac 2400
atcgacgacg tgctctctaa actgaacgaa tccatcaaca aagctatgat caacatcaac 2460
aagttctctg accagtgtct tgtaagctat ctgatgaact ccatgatccc gtacggtgtt 2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgetgaaata catttacgac 2580
aaccgtggca ctctgatcgg tcaggttgat cgtctgaagg acaaaagtga caatacctta 2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct ttgtccact 2700
ctagaaatag aaggtagaag tgggcaccat caccatcacc attaa 2745

```

```

<210> SEQ ID NO 79
<211> LENGTH: 914
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 79

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1             5             10            15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20            25            30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35            40            45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50            55            60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65            70            75            80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85            90            95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100           105           110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115           120           125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130           135           140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145           150           155           160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165           170           175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180           185           190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195           200           205

```

-continued

Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala
210						215					220				
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn
225					230					235					240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser
			245						250					255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260						265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
		275					280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290					295					300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325						330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
		340						345					350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
	355						360					365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405					410						415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg
	435						440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Arg	Lys	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465				470					475						480
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp
			485						490					495	
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys
		500						505					510		
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn
	515						520					525			
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp
	530					535					540				
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile
545					550					555					560
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys
			565						570					575	
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln
		580						585					590		
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn
	595						600					605			

-continued

Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp
610						615					620				
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly
625					630					635					640
Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val
			645						650					655	
Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile
			660					665					670		
Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val
		675					680					685			
Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro
	690					695					700				
Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile
705					710					715					720
Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys
			725						730					735	
Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp
		740						745					750		
Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys
		755					760				765				
Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr
	770				775						780				
Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn
785					790					795					800
Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met
			805						810					815	
Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met
			820					825					830		
Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala
		835					840					845			
Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr
	850					855					860				
Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu
865					870				875						880
Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg
			885						890					895	
Leu	Leu	Ser	Thr	Leu	Glu	Ile	Glu	Gly	Arg	Ser	Gly	His	His	His	His
		900					905					910			

His His

<210> SEQ ID NO 80

<211> LENGTH: 619

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 80

gctagcgggc ggtggcggtg gggcggtgg cggtagcggc ggtggcggtg ggcactagt	60
gctgcagtgt atcaatctgg attgggacgt aatccgtgat aagacaaaa caaaaatcga	120
gtctttgaaa gaacacggcc cgatcaaaaa taagatgtct gaatcaccca ataaaactgt	180
ttcgaggaaa aaagcgaaac agtatattgga agagtttcat caaacgcgc ttgaacatcc	240

-continued

ggagctcagt gaactgaaaa cagtgacggg aacgaatcct gtttttgcag gcgcaaaacta	300
tgcggcttgg gccgtgaatg ttgcccaagt aattgatagt gagaccgcag acaacctgga	360
aaagacgacc gcagcgtaa gcattttacc ggggattggt tccgtgatgg gtatagcgga	420
tggagcggtc caccataaca ctgaggaaat tgcgcccag tcaatcgctc tgagttccct	480
gatggttgca caggctatcc cactcgtggg ggaactggtt gacatagggt tcgccgccta	540
caacttcgta gaaagcatta ttaatctttt tcagggtgtg cataacagct acaaccgccc	600
tctagaatga taaaagctt	619

<210> SEQ ID NO 81

<211> LENGTH: 1971

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 81

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac	60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc	120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaacccgga agaagggtgac	180
ctgaacccgc caccggaagc gaaacaggtg ccggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaactggt cgagcgtatt	300
tactccaccg acctgggccg tatgtgctg actagcatcg ttcgcggtat cccgttcttg	360
ggcggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgatcatcgg cccgtctgct	480
gatatcatcc agttcgagt taagagcttt ggtcacgaag ttctgaacct cacccgtaac	540
ggctacgggt ccaactagta catccgtttc tctccggact tcaccttcgg ttttgaagaa	600
tccctggaag tagacacgaa cccactgctg ggcgctgta aattcgcaac tgatcctgag	660
gttaccctgg ctacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccgtg tcttcaaaat taacaccaac gcgtattacg agatgtccgg tctggaagtt	780
agcttcgaag aactgcgtac ttttggcggt cagcagccta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccatcgtagg taccactgct tctctccagt acatgaagaa cgtttttaaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgcgaaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctggct	1200
gctaatttta acggccagaa cacggaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcaactggtc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctc tgatagaagg tagatacggg ggtttcctgg cgctagcggg cggtagcggt	1380
agcggcgggt gcggtagcgg cggtagcggt agcgcactag tgctgcagtg tatcaatctg	1440
gattgggacg taatccgtga taagaacaaa acaaaaatcg agtctttgaa agaacacggc	1500
ccgatcaaaa ataagatgtc tgaatcacc aataaaactg tttcggagga aaaagcgaaa	1560

-continued

```

cagtatttgg aagagtttca tcaaaccgcg cttgaacatc cggagctcag tgaactgaaa 1620
acagtgcacgg gaacgaatcc tgtttttgca ggcgcaaact atgcggcttg ggcctgtaat 1680
gttgcccaag taattgatag tgagaccgca gacaacctgg aaaagacgac cgcagcgtaa 1740
agcattttac cggggattgg ttccgtgatg ggtatagcgg atggagcggg ccaccataac 1800
actgaggaaa ttgtcgccca gtcaatcgct ctgagttccc tgatggttgc acaggctatc 1860
ccactcgtgg gggaactggg tgacataggt ttcgccgcct acaacttcgt agaaagcatt 1920
attaatcttt ttcaggtggg gcataacagc tacaaccgcc ctctagaatg a 1971

```

```

<210> SEQ ID NO 82
<211> LENGTH: 656
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 82

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1      5      10      15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20     25     30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35     40     45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50     55     60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65     70     75     80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85     90     95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100    105    110
Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115    120    125
Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130    135    140
Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145    150    155    160
Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165    170    175
Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180    185    190
Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195    200    205
Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210    215    220
His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225    230    235    240
Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245    250    255
Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
260    265    270

```

-continued

Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	275	280	285
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	290	295	300
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	305	310	315
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	325	330	335
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	340	345	350
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	355	360	365
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	435	440	445
Tyr	Gly	Gly	Phe	Leu	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	450	455	460
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Asn	Leu	465	470	475
Asp	Trp	Asp	Val	Ile	Arg	Asp	Lys	Thr	Lys	Thr	Lys	Ile	Glu	Ser	Leu	485	490	495
Lys	Glu	His	Gly	Pro	Ile	Lys	Asn	Lys	Met	Ser	Glu	Ser	Pro	Asn	Lys	500	505	510
Thr	Val	Ser	Glu	Glu	Lys	Ala	Lys	Gln	Tyr	Leu	Glu	Glu	Phe	His	Gln	515	520	525
Thr	Ala	Leu	Glu	His	Pro	Glu	Leu	Ser	Glu	Leu	Lys	Thr	Val	Thr	Gly	530	535	540
Thr	Asn	Pro	Val	Phe	Ala	Gly	Ala	Asn	Tyr	Ala	Ala	Trp	Ala	Val	Asn	545	550	555
Val	Ala	Gln	Val	Ile	Asp	Ser	Glu	Thr	Ala	Asp	Asn	Leu	Glu	Lys	Thr	565	570	575
Thr	Ala	Ala	Leu	Ser	Ile	Leu	Pro	Gly	Ile	Gly	Ser	Val	Met	Gly	Ile	580	585	590
Ala	Asp	Gly	Ala	Val	His	His	Asn	Thr	Glu	Glu	Ile	Val	Ala	Gln	Ser	595	600	605
Ile	Ala	Leu	Ser	Ser	Leu	Met	Val	Ala	Gln	Ala	Ile	Pro	Leu	Val	Gly	610	615	620
Glu	Leu	Val	Asp	Ile	Gly	Phe	Ala	Ala	Tyr	Asn	Phe	Val	Glu	Ser	Ile	625	630	635
Ile	Asn	Leu	Phe	Gln	Val	Val	His	Asn	Ser	Tyr	Asn	Arg	Pro	Leu	Glu	645	650	655

<210> SEQ ID NO 83

<211> LENGTH: 1329

<212> TYPE: DNA

-continued

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 83

```

ggatccatgc ctattactat taacaatttt cgttatagcg atcccgtaa caatgacacc      60
attatcatga tggaaaccgc atattgcaaa ggactggaca ttactataa agccttcaag    120
attactgacc gcatttggat tgttccagag cgttacgagt tcgggacgaa accagaagat    180
ttaaaccgcg cttcatcgct gatcgaagga gcatcagagt attacgatcc gaactatctg    240
cgtacggaca gcgataaaga ccgcttctta cagaccatgg tcaaaacttt taaccgtatt    300
aagaacaatg tggccggaga agcactcttg gataagatta tcaacgcgat tccatacctg    360
ggcaattctt acagcctgct ggataaattt gacacaaata gtaattcagt cagctttaac    420
ctgttagaac aagatccgag tggcgcaacc acgaagtctg ccatgctgac aaatctgac     480
attttgggtc caggtcctgt actgaataaa aatgaagtac gcggcatcgt tctccgcgtg    540
gacaataaga actacttccc atgccgtgac ggcttcggtt cgatcatgca gatggctttc    600
tgtccggagt acgttccgac gtttgataat gttattgaga atatcacgag tttaacaatc    660
ggtaagtcaa aatattttta agatccggcc cttctcctta tgcatgaact gattcacgtg    720
ctgcacggct tatatggtat gcaagtgtcc tcgcatgaaa tcattccgct caaacaggaa    780
atztatatgc agcataccta cccgatttca gctgaagagt tgtttacggt ttgtggccag    840
gacgcgaatt tgatctccat cgacatcaaa aacgatctgt atgagaaaac attaaatgac    900
tataaagcga ttgcgaacaa actgtctcag gtgactagct gcaacgatcc taacattgat    960
attgattcct acaaaacaaat ttatcaacag aaataccagt tcgataaaga cagcaatggt   1020
cagtatatcg taaacgaaga taaatttcag atcctgtata acagcattat gtatggcttt   1080
accgaaattg agttggggaa gaaatttaac attaaaaccc gtctgtctta ttttagtatg   1140
aaccatgacg cggtgaaaat ccccaatctg cttgatgata ccatttataa tgataccgaa   1200
gggttcaaca ttgaatctaa ggatctgaaa tccgaataca aaggccaaaa tatgcgtgtt   1260
aatactaacg ctttccgtaa tgttgatggt agtggactcg tctcgaaact gattgggttg   1320
tgtgtcgac                                     1329

```

<210> SEQ ID NO 84

<211> LENGTH: 2736

<212> TYPE: DNA

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 84

```

ggatccatgc ctattactat taacaatttt cgttatagcg atcccgtaa caatgacacc      60
attatcatga tggaaaccgc atattgcaaa ggactggaca ttactataa agccttcaag    120
attactgacc gcatttggat tgttccagag cgttacgagt tcgggacgaa accagaagat    180
ttaaaccgcg cttcatcgct gatcgaagga gcatcagagt attacgatcc gaactatctg    240
cgtacggaca gcgataaaga ccgcttctta cagaccatgg tcaaaacttt taaccgtatt    300
aagaacaatg tggccggaga agcactcttg gataagatta tcaacgcgat tccatacctg    360
ggcaattctt acagcctgct ggataaattt gacacaaata gtaattcagt cagctttaac    420

```

-continued

ctgttagaac aagatccgag tggcgcaacc acgaagtctg ccatgctgac aaatctgate	480
atTTTTggtc caggctcctgt actgaataaa aatgaagtac gcggcatcgt tctccgcgtg	540
gacaataaga actacttccc atgccgtgac ggcttcggtt cgatcatgca gatggctttc	600
tgcccgagtg acgttccgac gtttgataat gttattgaga atatcacgag ttaacaatc	660
ggtaagtcaa aatattttca agatccggcc cttctcctta tgcatagaact gattcacgtg	720
ctgcacggct tatatgggat gcaagtgtcc tcgcatgaaa tcattccgtc caaacaggaa	780
atTTatatgc agcatacctc cccgatttca gctgaagagt tgTTtacgtt tggTggccag	840
gacgcgaatt tgatctccat cgacatcaaa aacgatctgt atgagaaaac attaaatgac	900
tataaagcga ttgcgaacaa actgtctcag gtgactagct gcaacgatcc taacattgat	960
attgattcct acaaacaaat ttatcaacag aaataccagt tcgataaaga cagcaatggt	1020
cagtatatcg taaacgaaga taaatttcag atcctgtata acagcattat gtatggcttt	1080
accgaaattg agttggggaa gaaatttaac attaaaacc gtctgtctta ttttagtatg	1140
aaccatgac cggtgaaat ccccaatctg cttgatgata ccatttataa tgataccgaa	1200
gggttcaaca ttgaatctaa ggatctgaaa tccgaataca aaggccaaaa tatgcgtgtt	1260
aatactaacg ctttccgtaa tgTTgatggt agTggactcg tctcgaaact gattgggttg	1320
tgTgtcgacg gcatcattac ctccaaaact aaatctctga tagaaggtag attTggcggt	1380
ttcacggcg caccgaaatc agcgcgtaaa cgtaagaacc aggcgctagc gggcggtggc	1440
ggtagcggcg gtggcggtag cggcggtggc ggtagcgac tagTgctgca gtgtatcaag	1500
gttaacaact gggatttatt cttcagcccg agtgaagaca acttcaccaa cgacctgaac	1560
aaaggTgaag aaatcacctc agatactaac atcgaagcag ccgaagaaaa catctcgctg	1620
gacctgatcc agcagtacta cctgacctt aatttcgaca acgagccgga aaacatttct	1680
atcgaaaaac tgagctctga tatcatcggc cagctggaac tgatgccgaa catcgaaagt	1740
ttcccaaacg gtaaaaagta cgagctggac aaatatacca tgttccacta cctgcgcgcg	1800
caggaaattg aacacggcaa atcccgatc gactgacta actccgttaa cgaagctctg	1860
ctcaaccctg cccgtgtata caccttcttc tctagcgact acgtgaaaaa ggtcaacaaa	1920
gcgactgaag ctgcaatgtt cttgggttg gttgaacagc ttgtttatga tttaccgac	1980
gagacgtccg aagtatctac taccgacaaa attgcgata tcactatcat catccgtac	2040
atcggTccgg ctctgaacat tggcaacatg ctgtacaaag acgacttctg tggcgactg	2100
atcttctccg gtgcggtgat cctgctggag ttcattcccg aaatcgccat cccggtactg	2160
ggcacctttg ctctggtttc ttacattgca aacaaggTtc tgactgtaca aaccatcgac	2220
aacgcgctga gcaaacgtaa cgaaaaatgg gatgaagTtt acaatatat cgtgaccaac	2280
tggctggcta aggttaatac tcagatcgac ctcatccgca aaaaaatgaa agaagcactg	2340
gaaaaccagg cggaagctac caaggcaatc attaaactacc agtacaacca gtacaccgag	2400
gaagaaaaaa acaacatcaa cttcaacatc gacgatctgt cctctaaact gaacgaatcc	2460
atcaacaaag ctatgatcaa catcaacaag ttctgaacc agtgctctgt aagctatctg	2520
atgaactcca tgatcccgta cgggtgttaa cgtctggagg acttcgatgc gtctctgaaa	2580
gacgcctgc tgaaatacat ttacgacaac cgtggcactc tgatcggtca ggttgatcgt	2640
ctgaaggaca aagtgaacaa taccttatcg accgacatcc cttttcagct cagtaaatat	2700

-continued

gtcgataacc aacgcctttt gtccactcta gactag

2736

<210> SEQ ID NO 85

<211> LENGTH: 911

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 85

Gly Ser Met Pro Ile Thr Ile Asn Asn Phe Arg Tyr Ser Asp Pro Val
 1 5 10 15
 Asn Asn Asp Thr Ile Ile Met Met Glu Pro Pro Tyr Cys Lys Gly Leu
 20 25 30
 Asp Ile Tyr Tyr Lys Ala Phe Lys Ile Thr Asp Arg Ile Trp Ile Val
 35 40 45
 Pro Glu Arg Tyr Glu Phe Gly Thr Lys Pro Glu Asp Phe Asn Pro Pro
 50 55 60
 Ser Ser Leu Ile Glu Gly Ala Ser Glu Tyr Tyr Asp Pro Asn Tyr Leu
 65 70 75 80
 Arg Thr Asp Ser Asp Lys Asp Arg Phe Leu Gln Thr Met Val Lys Leu
 85 90 95
 Phe Asn Arg Ile Lys Asn Asn Val Ala Gly Glu Ala Leu Leu Asp Lys
 100 105 110
 Ile Ile Asn Ala Ile Pro Tyr Leu Gly Asn Ser Tyr Ser Leu Leu Asp
 115 120 125
 Lys Phe Asp Thr Asn Ser Asn Ser Val Ser Phe Asn Leu Leu Glu Gln
 130 135 140
 Asp Pro Ser Gly Ala Thr Thr Lys Ser Ala Met Leu Thr Asn Leu Ile
 145 150 155 160
 Ile Phe Gly Pro Gly Pro Val Leu Asn Lys Asn Glu Val Arg Gly Ile
 165 170 175
 Val Leu Arg Val Asp Asn Lys Asn Tyr Phe Pro Cys Arg Asp Gly Phe
 180 185 190
 Gly Ser Ile Met Gln Met Ala Phe Cys Pro Glu Tyr Val Pro Thr Phe
 195 200 205
 Asp Asn Val Ile Glu Asn Ile Thr Ser Leu Thr Ile Gly Lys Ser Lys
 210 215 220
 Tyr Phe Gln Asp Pro Ala Leu Leu Leu Met His Glu Leu Ile His Val
 225 230 235 240
 Leu His Gly Leu Tyr Gly Met Gln Val Ser Ser His Glu Ile Ile Pro
 245 250 255
 Ser Lys Gln Glu Ile Tyr Met Gln His Thr Tyr Pro Ile Ser Ala Glu
 260 265 270
 Glu Leu Phe Thr Phe Gly Gly Gln Asp Ala Asn Leu Ile Ser Ile Asp
 275 280 285
 Ile Lys Asn Asp Leu Tyr Glu Lys Thr Leu Asn Asp Tyr Lys Ala Ile
 290 295 300
 Ala Asn Lys Leu Ser Gln Val Thr Ser Cys Asn Asp Pro Asn Ile Asp
 305 310 315 320
 Ile Asp Ser Tyr Lys Gln Ile Tyr Gln Gln Lys Tyr Gln Phe Asp Lys
 325 330 335
 Asp Ser Asn Gly Gln Tyr Ile Val Asn Glu Asp Lys Phe Gln Ile Leu

-continued

340					345					350					
Tyr	Asn	Ser	Ile	Met	Tyr	Gly	Phe	Thr	Glu	Ile	Glu	Leu	Gly	Lys	Lys
		355					360					365			
Phe	Asn	Ile	Lys	Thr	Arg	Leu	Ser	Tyr	Phe	Ser	Met	Asn	His	Asp	Pro
	370					375					380				
Val	Lys	Ile	Pro	Asn	Leu	Leu	Asp	Asp	Thr	Ile	Tyr	Asn	Asp	Thr	Glu
385					390					395					400
Gly	Phe	Asn	Ile	Glu	Ser	Lys	Asp	Leu	Lys	Ser	Glu	Tyr	Lys	Gly	Gln
				405					410					415	
Asn	Met	Arg	Val	Asn	Thr	Asn	Ala	Phe	Arg	Asn	Val	Asp	Gly	Ser	Gly
			420					425					430		
Leu	Val	Ser	Lys	Leu	Ile	Gly	Leu	Cys	Val	Asp	Gly	Ile	Ile	Thr	Ser
	435						440					445			
Lys	Thr	Lys	Ser	Leu	Ile	Glu	Gly	Arg	Phe	Gly	Gly	Phe	Thr	Gly	Ala
	450					455					460				
Arg	Lys	Ser	Ala	Arg	Lys	Arg	Lys	Asn	Gln	Ala	Leu	Ala	Gly	Gly	Gly
465					470					475					480
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu
				485					490					495	
Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu
			500					505					510		
Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp
	515						520					525			
Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln
	530					535					540				
Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser
545					550					555					560
Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro
				565					570					575	
Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr
			580					585					590		
Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser
	595						600					605			
Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser
	610					615					620				
Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys
625					630					635					640
Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr
				645					650					655	
Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala
			660					665					670		
Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Gly
	675						680					685			
Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser	Gly
	690					695					700				
Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val	Leu
705					710					715					720
Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val
				725					730					735	
Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu
			740					745					750		

-continued

Val Tyr Lys Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln
755 760 765

Ile Asp Leu Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala
770 775 780

Glu Ala Thr Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu
785 790 795 800

Glu Glu Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys
805 810 815

Leu Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu
820 825 830

Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly
835 840 845

Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu
850 855 860

Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg
865 870 875 880

Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln
885 890 895

Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp
900 905 910

<210> SEQ ID NO 86
<211> LENGTH: 180
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 86

```

ggatccacgc acgtcgacgc gattgatggt cgttttggtg gtttcacggg cgcacgcaaa    60
tcagcgcgta aacgtaagaa ccaggcgcta gcgggcggtg gcggtagcgg cggtggcggt    120
agcggcggtg gcggtagcgc actagtgtctg cagacgcacg gtctagaatg ataaaagctt    180

```

<210> SEQ ID NO 87
<211> LENGTH: 2715
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 87

```

ggatccgaat tcatgccgat caccatcaac aacttcaact acagcgatcc ggtggataac    60
aaaaacatcc tgtacctgga taccatctg aataccctgg cgaacgaacc ggaaaaagcg    120
tttcgtatca ccggcaacat ttgggttatt ccggatcggt ttagccgtaa cagcaaccgg    180
aatctgaata aaccgccgcg tgttaccagc ccgaaaagcg gttattacga tccgaactat    240
ctgagcaccg atagcgataa agataccttc ctgaaagaaa tcatcaaact gttcaaacgc    300
atcaacagcc gtgaaattgg cgaagaactg atctatcgcc tgagcacoga tattccggtt    360
ccgggcaaca acaacacccc gatcaacacc tttgatttcg atgtggattt caacagcggt    420
gatgttaaaa cccgccaggg taacaattgg gtgaaaaccg gcagcattaa cccgagcggt    480
attattaccg gtccgcgcga aaacattatt gatccggaaa ccagcacott taaactgacc    540
aacaacacct ttgcggcgca ggaagggttt ggccgcgtga gcattattag cattagcccg    600

```

-continued

cgctttatgc tgacctatag caacgcgacc aacgatgttg gtgaaggccg tttcagcaaa	660
agcgaatttt gcatggaccc gatcctgacg ctgatgcatg aactgaacca tgcgatgcat	720
aacctgtatg gcatcgcatg tccgaacgat cagaccatta gcagcgtgac cagcaacatc	780
ttttacagcc agtacaacgt gaaactggaa tatgcggaaa tctatgcgtt tggcggtcgcg	840
accattgacg tgattccgaa aagcgcgcgc aaatacttcg aagaaaaagc gctggattac	900
tatcgcagca ttgcgaaacg tctgaacagc attaccaccg cgaatccgag cagcttcaac	960
aaatatatcg gcgaatataa acagaaactg atccgcaaat atcgctttgt ggtggaaagc	1020
agcggcgaag ttaccgttaa ccgcaataaa ttcgtggaac tgtacaacga actgaccag	1080
atcttcaccg aatttaacta tgcgaaaatc tataacgtgc agaaccgtaa aatctacctg	1140
agcaacgtgt ataccccggt gaccgcgaat attctggatg ataacgtgta cgatatccag	1200
aacggcttta acatcccgaa aagcaacctg aacgttctgt ttatgggcca gaacctgagc	1260
cgtaatccgg cgctgcgtaa agtgaaccgc gaaaacatgc tgtacctgtt caccaaattt	1320
tgcgtcgacg cgattgatgg tcgttttggc ggtttcacgg gcgcacgcaa atcagcgcg	1380
aaacgtaaga accaggcgct agcggggcgt ggcggtagcg gcggtggcgg tagcggcggt	1440
ggcggtagcg cactagtgtc gcagtgtcgt gaactgctgg tgaaaaacac cgatctgccg	1500
tttattggcg atatcagcga tgtgaaaacc gatattctcc tgcgcaaaga tatcaacgaa	1560
gaaaccgaag tgatctacta cccggataac gtgagcgttg atcaggtgat cctgagcaaa	1620
aacaccagcg aacatggtca gctggatctg ctgtatccga gcattgatag cgaaagcgaa	1680
attctgccgg gcgaaaacca ggtgttttac gataaccgta ccagaaacgt ggattacctg	1740
aacagctatt actacctgga aagccagaaa ctgagcgata acgtggaaga ttttaccttt	1800
accgcgagca ttgaagaagc gctggataac agcgcgaaag tttacaccta tttccgacc	1860
ctggcgaaac aagttaatgc ggggtgttcag ggcggtctgt ttctgatgtg gcggaacgat	1920
gtggtggaag atttcaccac caacatcctg cgtaaagata ccctggataa aatcagcgat	1980
gttagcgcga ttattccgta tattggtccg gcgctgaaca ttagcaatag cgtgcgtcgt	2040
ggcaatttta ccgaagcgtt tgcggttacc ggtgtgacca ttctgctgga agcgtttccg	2100
gaatttacca ttccggcgct ggggtcgctt gtgatctata gcaaagtgca ggaacgcaac	2160
gaaatcatca aaaccatcga taactgcctg gaacagcgta ttaaacgctg gaaagatagc	2220
tatgaatgga tgatgggac ctggctgagc cgtattatca ccagttcaa caacatcagc	2280
taccagatgt acgatagcct gaactatcag gcgggtgcga ttaagcgaa aatcgatctg	2340
gaatacaaaa aatacagcgg cagcgataaa gaaaacatca aaagccaggt tgaaaacctg	2400
aaaaacagcc tggatgtgaa aattagcgaa gcgatgaata acatcaacaa attcatccgc	2460
gaatgcagcg tgacctacct gttcaaaaac atgctgccga aagtgatcga tgaactgaac	2520
gaatttgatc gcaacaccaa agcgaaactg atcaacctga tcgatagcca caacattatt	2580
ctggtgggcg aagtggataa actgaaagcg aaagttaaca acagcttcca gaacaccatc	2640
ccgtttaaca tcttcagcta taccaacaac agcctgctga aagatatcat caacgaatac	2700
ttcaatctag actag	2715

<210> SEQ ID NO 88

<211> LENGTH: 904

-continued

```

<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 88

Gly Ser Glu Phe Met Pro Ile Thr Ile Asn Asn Phe Asn Tyr Ser Asp
1          5          10          15
Pro Val Asp Asn Lys Asn Ile Leu Tyr Leu Asp Thr His Leu Asn Thr
20          25          30
Leu Ala Asn Glu Pro Glu Lys Ala Phe Arg Ile Thr Gly Asn Ile Trp
35          40          45
Val Ile Pro Asp Arg Phe Ser Arg Asn Ser Asn Pro Asn Leu Asn Lys
50          55          60
Pro Pro Arg Val Thr Ser Pro Lys Ser Gly Tyr Tyr Asp Pro Asn Tyr
65          70          75          80
Leu Ser Thr Asp Ser Asp Lys Asp Thr Phe Leu Lys Glu Ile Ile Lys
85          90          95
Leu Phe Lys Arg Ile Asn Ser Arg Glu Ile Gly Glu Glu Leu Ile Tyr
100         105         110
Arg Leu Ser Thr Asp Ile Pro Phe Pro Gly Asn Asn Asn Thr Pro Ile
115         120         125
Asn Thr Phe Asp Phe Asp Val Asp Phe Asn Ser Val Asp Val Lys Thr
130         135         140
Arg Gln Gly Asn Asn Trp Val Lys Thr Gly Ser Ile Asn Pro Ser Val
145         150         155         160
Ile Ile Thr Gly Pro Arg Glu Asn Ile Ile Asp Pro Glu Thr Ser Thr
165         170         175
Phe Lys Leu Thr Asn Asn Thr Phe Ala Ala Gln Glu Gly Phe Gly Ala
180         185         190
Leu Ser Ile Ile Ser Ile Ser Pro Arg Phe Met Leu Thr Tyr Ser Asn
195         200         205
Ala Thr Asn Asp Val Gly Glu Gly Arg Phe Ser Lys Ser Glu Phe Cys
210         215         220
Met Asp Pro Ile Leu Ile Leu Met His Glu Leu Asn His Ala Met His
225         230         235         240
Asn Leu Tyr Gly Ile Ala Ile Pro Asn Asp Gln Thr Ile Ser Ser Val
245         250         255
Thr Ser Asn Ile Phe Tyr Ser Gln Tyr Asn Val Lys Leu Glu Tyr Ala
260         265         270
Glu Ile Tyr Ala Phe Gly Gly Pro Thr Ile Asp Leu Ile Pro Lys Ser
275         280         285
Ala Arg Lys Tyr Phe Glu Glu Lys Ala Leu Asp Tyr Tyr Arg Ser Ile
290         295         300
Ala Lys Arg Leu Asn Ser Ile Thr Thr Ala Asn Pro Ser Ser Phe Asn
305         310         315         320
Lys Tyr Ile Gly Glu Tyr Lys Gln Lys Leu Ile Arg Lys Tyr Arg Phe
325         330         335
Val Val Glu Ser Ser Gly Glu Val Thr Val Asn Arg Asn Lys Phe Val
340         345         350
Glu Leu Tyr Asn Glu Leu Thr Gln Ile Phe Thr Glu Phe Asn Tyr Ala
355         360         365

```

-continued

Lys	Ile	Tyr	Asn	Val	Gln	Asn	Arg	Lys	Ile	Tyr	Leu	Ser	Asn	Val	Tyr
370						375					380				
Thr	Pro	Val	Thr	Ala	Asn	Ile	Leu	Asp	Asp	Asn	Val	Tyr	Asp	Ile	Gln
385					390					395				400	
Asn	Gly	Phe	Asn	Ile	Pro	Lys	Ser	Asn	Leu	Asn	Val	Leu	Phe	Met	Gly
			405						410					415	
Gln	Asn	Leu	Ser	Arg	Asn	Pro	Ala	Leu	Arg	Lys	Val	Asn	Pro	Glu	Asn
			420					425					430		
Met	Leu	Tyr	Leu	Phe	Thr	Lys	Phe	Cys	Val	Asp	Ala	Ile	Asp	Gly	Arg
	435						440					445			
Phe	Gly	Gly	Phe	Thr	Gly	Ala	Arg	Lys	Ser	Ala	Arg	Lys	Arg	Lys	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465					470					475					480
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Arg	Glu	Leu	Leu	Val	Lys	Asn
			485						490					495	
Thr	Asp	Leu	Pro	Phe	Ile	Gly	Asp	Ile	Ser	Asp	Val	Lys	Thr	Asp	Ile
		500						505					510		
Phe	Leu	Arg	Lys	Asp	Ile	Asn	Glu	Glu	Thr	Glu	Val	Ile	Tyr	Tyr	Pro
		515					520					525			
Asp	Asn	Val	Ser	Val	Asp	Gln	Val	Ile	Leu	Ser	Lys	Asn	Thr	Ser	Glu
		530				535					540				
His	Gly	Gln	Leu	Asp	Leu	Leu	Tyr	Pro	Ser	Ile	Asp	Ser	Glu	Ser	Glu
545					550					555					560
Ile	Leu	Pro	Gly	Glu	Asn	Gln	Val	Phe	Tyr	Asp	Asn	Arg	Thr	Gln	Asn
				565					570					575	
Val	Asp	Tyr	Leu	Asn	Ser	Tyr	Tyr	Tyr	Leu	Glu	Ser	Gln	Lys	Leu	Ser
		580						585					590		
Asp	Asn	Val	Glu	Asp	Phe	Thr	Phe	Thr	Arg	Ser	Ile	Glu	Glu	Ala	Leu
		595					600					605			
Asp	Asn	Ser	Ala	Lys	Val	Tyr	Thr	Tyr	Phe	Pro	Thr	Leu	Ala	Asn	Lys
		610				615				620					
Val	Asn	Ala	Gly	Val	Gln	Gly	Gly	Leu	Phe	Leu	Met	Trp	Ala	Asn	Asp
625					630					635					640
Val	Val	Glu	Asp	Phe	Thr	Thr	Asn	Ile	Leu	Arg	Lys	Asp	Thr	Leu	Asp
				645					650					655	
Lys	Ile	Ser	Asp	Val	Ser	Ala	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu
		660						665					670		
Asn	Ile	Ser	Asn	Ser	Val	Arg	Arg	Gly	Asn	Phe	Thr	Glu	Ala	Phe	Ala
		675					680					685			
Val	Thr	Gly	Val	Thr	Ile	Leu	Leu	Glu	Ala	Phe	Pro	Glu	Phe	Thr	Ile
	690					695					700				
Pro	Ala	Leu	Gly	Ala	Phe	Val	Ile	Tyr	Ser	Lys	Val	Gln	Glu	Arg	Asn
705					710					715					720
Glu	Ile	Ile	Lys	Thr	Ile	Asp	Asn	Cys	Leu	Glu	Gln	Arg	Ile	Lys	Arg
				725					730					735	
Trp	Lys	Asp	Ser	Tyr	Glu	Trp	Met	Met	Gly	Thr	Trp	Leu	Ser	Arg	Ile
			740					745					750		
Ile	Thr	Gln	Phe	Asn	Asn	Ile	Ser	Tyr	Gln	Met	Tyr	Asp	Ser	Leu	Asn
		755					760					765			
Tyr	Gln	Ala	Gly	Ala	Ile	Lys	Ala	Lys	Ile	Asp	Leu	Glu	Tyr	Lys	Lys

-continued

770	775	780
Tyr Ser Gly Ser Asp	Lys Glu Asn Ile Lys Ser	Gln Val Glu Asn Leu
785	790	795 800
Lys Asn Ser Leu Asp	Val Lys Ile Ser Glu Ala Met Asn Asn Ile Asn	
	805	810 815
Lys Phe Ile Arg Glu Cys Ser Val Thr Tyr Leu Phe Lys Asn Met Leu		
	820	825 830
Pro Lys Val Ile Asp Glu Leu Asn Glu Phe Asp Arg Asn Thr Lys Ala		
	835	840 845
Lys Leu Ile Asn Leu Ile Asp Ser His Asn Ile Ile Leu Val Gly Glu		
	850	855 860
Val Asp Lys Leu Lys Ala Lys Val Asn Asn Ser Phe Gln Asn Thr Ile		
	865	870 875 880
Pro Phe Asn Ile Phe Ser Tyr Thr Asn Asn Ser Leu Leu Lys Asp Ile		
	885	890 895
Ile Asn Glu Tyr Phe Asn Leu Asp		
	900	

<210> SEQ ID NO 89
 <211> LENGTH: 17
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 89

Tyr Gly Gly Phe Leu Arg Arg Ile Arg Pro Lys Leu Lys Trp Asp Asn
1 5 10 15

Gln

<210> SEQ ID NO 90
 <211> LENGTH: 2709
 <212> TYPE: DNA
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 90

ggatccatgg agttcgttaa caaacagttc aactataaag acccagttaa cgggtgttgac	60
attgcttaca tcaaaatccc gaacgctggc cagatgcagc cggtaaaggc attcaaaatc	120
cacaacaaaa tctgggttat cccggaacgt gataccttta ctaaccgga agaaggtgac	180
ctgaaccgcg caccggaagc gaaacaggtg cgggtatctt actatgactc cacctacctg	240
tctaccgata acgaaaagga caactacctg aaaggtgtta ctaaaactgtt cgagcgtatt	300
tactccaccg acctggggcg tatgtgctg actagcatcg ttccgggtat cccgttctgg	360
ggcgggttcta ccatcgatac cgaactgaaa gtaatcgaca ctaactgcat caacgttatt	420
cagccggacg gttcctatcg ttccgaagaa ctgaacctgg tgaatcggg cccgtctgct	480
gatatcatcc agttcagatg taagagcttt ggtcacgaag ttctgaacct caccgtaac	540
ggctacgggt ccaactcagta catccgttcc tctccggact tcaccttcgg tttgaagaa	600
tccctggaag tagacacgaa cccactgctg ggcgctggta aattcgcaac tgatcctgcg	660
gttacccctg ctcacgaact gattcatgca ggccaccgcc tgtacggtat cgccatcaat	720
ccgaaccctg tcttcaaagt taacaccaac gcgtattacg agatgtccgg tctggaagtt	780

-continued

agcttcgaag aactgcgtac ttttggcggg cagcagccta aattcatcga ctctctgcaa	840
gaaaacgagt tccgtctgta ctactataac aagttcaaag atatcgcatc caccctgaac	900
aaagcgaaat ccatcgtggg taccactgct tctctccagt acatgaagaa cgtttttaaa	960
gaaaaatacc tgctcagcga agacacctcc ggcaaattct ctgtagacaa gttgaaattc	1020
gataaaacttt acaaaatgct gactgaaatt tacaccgaag acaacttcgt taagttcttt	1080
aaagttctga accgcaaac ctatctgaac ttcgacaagg cagtattcaa aatcaacatc	1140
gtgccgaaag ttaactacac tatctacgat ggtttcaacc tgcgtaacac caacctgggt	1200
gctaatttta acggccagaa caccgaaatc aacaacatga acttcacaaa actgaaaaac	1260
ttcactgggc tgttcgagtt ttacaagctg ctgtgcgtcg acggcatcat tacctccaaa	1320
actaaatctg acgatgacga taaatatgga ggttttttga gaaggatagc accaaaatta	1380
aagtgggata atcaagcgct agcggggcgg ggcggtagcg gcggtggcgg tagcggcggg	1440
ggcggtagcg cactagtgtc gcagtgtatc aaggttaaca actgggattt attcttcagc	1500
cggagtgaag acaacttcac caacgacctg aacaaagggt aagaaatcac ctgagatact	1560
aacatcgaag cagccgaaga aaacatctcg ctggacctga tccagcagta ctacctgacc	1620
tttaatttcg acaacgagcc ggaaaacatt tctatcgaaa acctgagctc tgatatcatc	1680
ggccagctgg aactgatgcc gaacatcgaa cgtttcccaa acggtaaaaa gtacgagctg	1740
gacaaatata ccatgttcca ctacctgcgc gcgcaggaat ttgaacacgg caaatccgt	1800
atcgcaactga ctaactccgt taacgaagct ctgctcaacc cgtcccgtgt atacaccttc	1860
ttctctagcg actacgtgaa aaaggtcaac aaagcgactg aagctgcaat gttcttgggt	1920
tgggttgaac agcttgttta tgattttacc gacgagacgt ccgaagtatc tactaccgac	1980
aaaattgctg atatcactat catcatcccg tacatcggtc cggctctgaa catttgcaac	2040
atgctgtaca aagacgactt cgttggcgca ctgatcttct cgggtgcggg gatcctgctg	2100
gagttcatcc cggaatcgc catcccggtg ctgggcacct ttgctctggt ttcttacatt	2160
gcaaacaaag ttctgactgt acaaacctac gacaacgcgc tgagcaaacg taacgaaaaa	2220
tgggatgaag tttaaaaata tatcgtgacc aactggctgg ctaagggtta tactcagatc	2280
gacctcatcc gcaaaaaaat gaaagaagca ctggaaaacc aggcggaagc taccaaggca	2340
atcattaact accagtacaa ccagtacacc gaggaagaaa aaaacaacat caacttcaac	2400
atcgacgatc tgtcctctaa actgaacgaa tccatcaaca aagctatgat caacatcaac	2460
aagttctcga accagtgtct tgtaagctat ctgatgaact ccatgatccc gtaacgtgtt	2520
aaacgtctgg aggacttcga tgcgtctctg aaagacgccc tgctgaaata catttacgac	2580
aacgctggca ctctgatcgg tcaggttgat cgtctgaagg acaaaagtga caatacotta	2640
tcgaccgaca tcccttttca gctcagtaaa tatgtcgata accaacgcct ttgtccact	2700
ctagactag	2709

<210> SEQ ID NO 91

<211> LENGTH: 902

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 91

-continued

Gly	Ser	Met	Glu	Phe	Val	Asn	Lys	Gln	Phe	Asn	Tyr	Lys	Asp	Pro	Val	1	5	10	15
Asn	Gly	Val	Asp	Ile	Ala	Tyr	Ile	Lys	Ile	Pro	Asn	Ala	Gly	Gln	Met	20	25	30	
Gln	Pro	Val	Lys	Ala	Phe	Lys	Ile	His	Asn	Lys	Ile	Trp	Val	Ile	Pro	35	40	45	
Glu	Arg	Asp	Thr	Phe	Thr	Asn	Pro	Glu	Glu	Gly	Asp	Leu	Asn	Pro	Pro	50	55	60	
Pro	Glu	Ala	Lys	Gln	Val	Pro	Val	Ser	Tyr	Tyr	Asp	Ser	Thr	Tyr	Leu	65	70	75	80
Ser	Thr	Asp	Asn	Glu	Lys	Asp	Asn	Tyr	Leu	Lys	Gly	Val	Thr	Lys	Leu	85	90	95	
Phe	Glu	Arg	Ile	Tyr	Ser	Thr	Asp	Leu	Gly	Arg	Met	Leu	Leu	Thr	Ser	100	105	110	
Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu	115	120	125	
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly	130	135	140	
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala	145	150	155	160
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn	165	170	175	
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro	180	185	190	
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro	195	200	205	
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala	210	215	220	
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn	225	230	235	240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	245	250	255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp	260	265	270	
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr	275	280	285	
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser	290	295	300	
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys	305	310	315	320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp	325	330	335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr	340	345	350	
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr	355	360	365	
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380	
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395	400

-continued

Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405						410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
			420					425					430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Asp	Asp	Asp	Asp	Lys
		435					440				445				
Tyr	Gly	Gly	Phe	Leu	Arg	Arg	Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly
465				470					475					480	
Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp
			485					490						495	
Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys
			500					505					510		
Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn
		515					520				525				
Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp
	530					535					540				
Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile
545				550					555					560	
Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys
			565					570						575	
Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln
		580					585						590		
Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn
		595					600					605			
Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp
	610				615						620				
Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly
625				630					635					640	
Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val
			645					650						655	
Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile
		660					665						670		
Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val
		675			680							685			
Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro
	690				695					700					
Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile
705				710					715					720	
Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys
			725					730						735	
Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp
		740					745						750		
Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys
		755				760						765			
Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr
	770				775						780				
Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn
785				790					795					800	
Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met

-continued

805					810					815						
Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	
820					825					830						
Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	
835					840					845						
Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	
850					855					860						
Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	
865					870					875					880	
Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	
885					890					895						
Leu	Leu	Ser	Thr	Leu	Asp											
900																
<210> SEQ ID NO 92																
<211> LENGTH: 892																
<212> TYPE: PRT																
<213> ORGANISM: Artificial Sequence																
<220> FEATURE:																
<223> OTHER INFORMATION: Synthetic																
<400> SEQUENCE: 92																
Gly	Ser	Met	Glu	Phe	Val	Asn	Lys	Gln	Phe	Asn	Tyr	Lys	Asp	Pro	Val	
1				5				10				15				
Asn	Gly	Val	Asp	Ile	Ala	Tyr	Ile	Lys	Ile	Pro	Asn	Ala	Gly	Gln	Met	
			20				25				30					
Gln	Pro	Val	Lys	Ala	Phe	Lys	Ile	His	Asn	Lys	Ile	Trp	Val	Ile	Pro	
			35				40				45					
Glu	Arg	Asp	Thr	Phe	Thr	Asn	Pro	Glu	Glu	Gly	Asp	Leu	Asn	Pro	Pro	
			50				55				60					
Pro	Glu	Ala	Lys	Gln	Val	Pro	Val	Ser	Tyr	Tyr	Asp	Ser	Thr	Tyr	Leu	
			65				70				75				80	
Ser	Thr	Asp	Asn	Glu	Lys	Asp	Asn	Tyr	Leu	Lys	Gly	Val	Thr	Lys	Leu	
			85				90				95					
Phe	Glu	Arg	Ile	Tyr	Ser	Thr	Asp	Leu	Gly	Arg	Met	Leu	Leu	Thr	Ser	
			100				105				110					
Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu	
			115				120				125					
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly	
			130				135				140					
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala	
			145				150				155				160	
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn	
			165				170				175					
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro	
			180				185				190					
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro	
			195				200				205					
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala	
			210				215				220					
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn	
			225				230				235				240	
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser	

-continued

245								250				255			
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
			260					265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
		275					280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290					295					300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
				325					330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
			340					345					350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
		355					360					365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
				405					410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
			420					425					430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Asp	Asp	Asp	Asp	Lys
		435					440					445			
Tyr	Gly	Gly	Phe	Leu	Arg	Arg	Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile
465					470					475					480
Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe
				485					490					495	
Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile
			500					505					510		
Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr
		515					520					525			
Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn
	530					535					540				
Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu
545				550						555					560
Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe
				565					570					575	
His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala
			580					585					590		
Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr
		595					600					605			
Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu
	610					615					620				
Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr
625					630					635					640
Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr
				645					650					655	

-continued

```

Ile Ile Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly Asn Met Leu
    660              665              670
Tyr Lys Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly Ala Val Ile
    675              680              685
Leu Leu Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe
    690              695              700
Ala Leu Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile
    705              710              715
Asp Asn Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys
    725              730              735
Tyr Ile Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu
    740              745              750
Ile Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr
    755              760              765
Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu Lys
    770              775              780
Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu Asn Glu
    785              790              795
Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu Asn Gln Cys
    805              810              815
Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly Val Lys Arg
    820              825              830
Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu Lys Tyr Ile
    835              840              845
Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg Leu Lys Asp
    850              855              860
Lys Val Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln Leu Ser Lys
    865              870              875
Tyr Val Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp
    885              890

```

```

<210> SEQ ID NO 93
<211> LENGTH: 897
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

```

```

<400> SEQUENCE: 93

```

```

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1      5      10      15
Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20     25     30
Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35     40     45
Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50     55     60
Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65     70     75     80
Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85     90     95
Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100    105    110

```

-continued

Ile	Val	Arg	Gly	Ile	Pro	Phe	Trp	Gly	Gly	Ser	Thr	Ile	Asp	Thr	Glu
	115						120					125			
Leu	Lys	Val	Ile	Asp	Thr	Asn	Cys	Ile	Asn	Val	Ile	Gln	Pro	Asp	Gly
	130					135					140				
Ser	Tyr	Arg	Ser	Glu	Glu	Leu	Asn	Leu	Val	Ile	Ile	Gly	Pro	Ser	Ala
145				150						155					160
Asp	Ile	Ile	Gln	Phe	Glu	Cys	Lys	Ser	Phe	Gly	His	Glu	Val	Leu	Asn
			165						170					175	
Leu	Thr	Arg	Asn	Gly	Tyr	Gly	Ser	Thr	Gln	Tyr	Ile	Arg	Phe	Ser	Pro
		180						185					190		
Asp	Phe	Thr	Phe	Gly	Phe	Glu	Glu	Ser	Leu	Glu	Val	Asp	Thr	Asn	Pro
	195						200					205			
Leu	Leu	Gly	Ala	Gly	Lys	Phe	Ala	Thr	Asp	Pro	Ala	Val	Thr	Leu	Ala
	210					215					220				
His	Glu	Leu	Ile	His	Ala	Gly	His	Arg	Leu	Tyr	Gly	Ile	Ala	Ile	Asn
225					230					235					240
Pro	Asn	Arg	Val	Phe	Lys	Val	Asn	Thr	Asn	Ala	Tyr	Tyr	Glu	Met	Ser
			245						250					255	
Gly	Leu	Glu	Val	Ser	Phe	Glu	Glu	Leu	Arg	Thr	Phe	Gly	Gly	His	Asp
		260						265					270		
Ala	Lys	Phe	Ile	Asp	Ser	Leu	Gln	Glu	Asn	Glu	Phe	Arg	Leu	Tyr	Tyr
	275						280					285			
Tyr	Asn	Lys	Phe	Lys	Asp	Ile	Ala	Ser	Thr	Leu	Asn	Lys	Ala	Lys	Ser
	290					295					300				
Ile	Val	Gly	Thr	Thr	Ala	Ser	Leu	Gln	Tyr	Met	Lys	Asn	Val	Phe	Lys
305					310					315					320
Glu	Lys	Tyr	Leu	Leu	Ser	Glu	Asp	Thr	Ser	Gly	Lys	Phe	Ser	Val	Asp
			325						330					335	
Lys	Leu	Lys	Phe	Asp	Lys	Leu	Tyr	Lys	Met	Leu	Thr	Glu	Ile	Tyr	Thr
		340						345					350		
Glu	Asp	Asn	Phe	Val	Lys	Phe	Phe	Lys	Val	Leu	Asn	Arg	Lys	Thr	Tyr
	355						360					365			
Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val
	370					375					380				
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala
385					390					395					400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr
			405						410					415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys
		420					425						430		
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Asp	Asp	Asp	Asp	Lys
	435						440				445				
Tyr	Gly	Gly	Phe	Leu	Arg	Arg	Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn
	450					455					460				
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu
465					470					475					480
Val	Leu	Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro
			485						490					495	
Ser	Glu	Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr
			500					505						510	

-continued

Ser	Asp	Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu
	515						520					525			
Ile	Gln	Gln	Tyr	Tyr	Leu	Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn
	530					535				540					
Ile	Ser	Ile	Glu	Asn	Leu	Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu
545				550					555					560	
Met	Pro	Asn	Ile	Glu	Arg	Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp
			565						570					575	
Lys	Tyr	Thr	Met	Phe	His	Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly
		580						585					590		
Lys	Ser	Arg	Ile	Ala	Leu	Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn
	595						600					605			
Pro	Ser	Arg	Val	Tyr	Thr	Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val
	610					615				620					
Asn	Lys	Ala	Thr	Glu	Ala	Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu
625				630						635				640	
Val	Tyr	Asp	Phe	Thr	Asp	Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys
			645					650						655	
Ile	Ala	Asp	Ile	Thr	Ile	Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn
		660						665					670		
Ile	Gly	Asn	Met	Leu	Tyr	Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe
	675					680					685				
Ser	Gly	Ala	Val	Ile	Leu	Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro
	690				695					700					
Val	Leu	Gly	Thr	Phe	Ala	Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu
705				710					715					720	
Thr	Val	Gln	Thr	Ile	Asp	Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp
			725					730						735	
Asp	Glu	Val	Tyr	Lys	Tyr	Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn
		740					745					750			
Thr	Gln	Ile	Asp	Leu	Ile	Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn
	755					760					765				
Gln	Ala	Glu	Ala	Thr	Lys	Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr
770					775						780				
Thr	Glu	Glu	Glu	Lys	Asn	Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser
785				790					795					800	
Ser	Lys	Leu	Asn	Glu	Ser	Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys
			805					810					815		
Phe	Leu	Asn	Gln	Cys	Ser	Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro
		820					825					830			
Tyr	Gly	Val	Lys	Arg	Leu	Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala
		835				840						845			
Leu	Leu	Lys	Tyr	Ile	Tyr	Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val
	850				855						860				
Asp	Arg	Leu	Lys	Asp	Lys	Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro
865				870					875					880	
Phe	Gln	Leu	Ser	Lys	Tyr	Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu
			885					890						895	

Asp

-continued

```

<211> LENGTH: 907
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 94

Gly Ser Met Glu Phe Val Asn Lys Gln Phe Asn Tyr Lys Asp Pro Val
1          5          10          15

Asn Gly Val Asp Ile Ala Tyr Ile Lys Ile Pro Asn Ala Gly Gln Met
20          25          30

Gln Pro Val Lys Ala Phe Lys Ile His Asn Lys Ile Trp Val Ile Pro
35          40          45

Glu Arg Asp Thr Phe Thr Asn Pro Glu Glu Gly Asp Leu Asn Pro Pro
50          55          60

Pro Glu Ala Lys Gln Val Pro Val Ser Tyr Tyr Asp Ser Thr Tyr Leu
65          70          75          80

Ser Thr Asp Asn Glu Lys Asp Asn Tyr Leu Lys Gly Val Thr Lys Leu
85          90          95

Phe Glu Arg Ile Tyr Ser Thr Asp Leu Gly Arg Met Leu Leu Thr Ser
100         105         110

Ile Val Arg Gly Ile Pro Phe Trp Gly Gly Ser Thr Ile Asp Thr Glu
115         120         125

Leu Lys Val Ile Asp Thr Asn Cys Ile Asn Val Ile Gln Pro Asp Gly
130         135         140

Ser Tyr Arg Ser Glu Glu Leu Asn Leu Val Ile Ile Gly Pro Ser Ala
145         150         155         160

Asp Ile Ile Gln Phe Glu Cys Lys Ser Phe Gly His Glu Val Leu Asn
165         170         175

Leu Thr Arg Asn Gly Tyr Gly Ser Thr Gln Tyr Ile Arg Phe Ser Pro
180         185         190

Asp Phe Thr Phe Gly Phe Glu Glu Ser Leu Glu Val Asp Thr Asn Pro
195         200         205

Leu Leu Gly Ala Gly Lys Phe Ala Thr Asp Pro Ala Val Thr Leu Ala
210         215         220

His Glu Leu Ile His Ala Gly His Arg Leu Tyr Gly Ile Ala Ile Asn
225         230         235         240

Pro Asn Arg Val Phe Lys Val Asn Thr Asn Ala Tyr Tyr Glu Met Ser
245         250         255

Gly Leu Glu Val Ser Phe Glu Glu Leu Arg Thr Phe Gly Gly His Asp
260         265         270

Ala Lys Phe Ile Asp Ser Leu Gln Glu Asn Glu Phe Arg Leu Tyr Tyr
275         280         285

Tyr Asn Lys Phe Lys Asp Ile Ala Ser Thr Leu Asn Lys Ala Lys Ser
290         295         300

Ile Val Gly Thr Thr Ala Ser Leu Gln Tyr Met Lys Asn Val Phe Lys
305         310         315         320

Glu Lys Tyr Leu Leu Ser Glu Asp Thr Ser Gly Lys Phe Ser Val Asp
325         330         335

Lys Leu Lys Phe Asp Lys Leu Tyr Lys Met Leu Thr Glu Ile Tyr Thr
340         345         350

Glu Asp Asn Phe Val Lys Phe Phe Lys Val Leu Asn Arg Lys Thr Tyr
355         360         365

```

-continued

Leu	Asn	Phe	Asp	Lys	Ala	Val	Phe	Lys	Ile	Asn	Ile	Val	Pro	Lys	Val	370	375	380	
Asn	Tyr	Thr	Ile	Tyr	Asp	Gly	Phe	Asn	Leu	Arg	Asn	Thr	Asn	Leu	Ala	385	390	395	400
Ala	Asn	Phe	Asn	Gly	Gln	Asn	Thr	Glu	Ile	Asn	Asn	Met	Asn	Phe	Thr	405	410	415	
Lys	Leu	Lys	Asn	Phe	Thr	Gly	Leu	Phe	Glu	Phe	Tyr	Lys	Leu	Leu	Cys	420	425	430	
Val	Asp	Gly	Ile	Ile	Thr	Ser	Lys	Thr	Lys	Ser	Asp	Asp	Asp	Asp	Lys	435	440	445	
Tyr	Gly	Gly	Phe	Leu	Arg	Arg	Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn	450	455	460	
Gln	Ala	Leu	Ala	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	465	470	475	480
Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu	Gln	Cys	Ile	Lys	485	490	495	
Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu	Asp	Asn	Phe	Thr	500	505	510	
Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp	Thr	Asn	Ile	Glu	515	520	525	
Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln	Gln	Tyr	Tyr	Leu	530	535	540	
Thr	Phe	Asn	Phe	Asp	Asn	Glu	Pro	Glu	Asn	Ile	Ser	Ile	Glu	Asn	Leu	545	550	555	560
Ser	Ser	Asp	Ile	Ile	Gly	Gln	Leu	Glu	Leu	Met	Pro	Asn	Ile	Glu	Arg	565	570	575	
Phe	Pro	Asn	Gly	Lys	Lys	Tyr	Glu	Leu	Asp	Lys	Tyr	Thr	Met	Phe	His	580	585	590	
Tyr	Leu	Arg	Ala	Gln	Glu	Phe	Glu	His	Gly	Lys	Ser	Arg	Ile	Ala	Leu	595	600	605	
Thr	Asn	Ser	Val	Asn	Glu	Ala	Leu	Leu	Asn	Pro	Ser	Arg	Val	Tyr	Thr	610	615	620	
Phe	Phe	Ser	Ser	Asp	Tyr	Val	Lys	Lys	Val	Asn	Lys	Ala	Thr	Glu	Ala	625	630	635	640
Ala	Met	Phe	Leu	Gly	Trp	Val	Glu	Gln	Leu	Val	Tyr	Asp	Phe	Thr	Asp	645	650	655	
Glu	Thr	Ser	Glu	Val	Ser	Thr	Thr	Asp	Lys	Ile	Ala	Asp	Ile	Thr	Ile	660	665	670	
Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Gly	Asn	Met	Leu	Tyr	675	680	685	
Lys	Asp	Asp	Phe	Val	Gly	Ala	Leu	Ile	Phe	Ser	Gly	Ala	Val	Ile	Leu	690	695	700	
Leu	Glu	Phe	Ile	Pro	Glu	Ile	Ala	Ile	Pro	Val	Leu	Gly	Thr	Phe	Ala	705	710	715	720
Leu	Val	Ser	Tyr	Ile	Ala	Asn	Lys	Val	Leu	Thr	Val	Gln	Thr	Ile	Asp	725	730	735	
Asn	Ala	Leu	Ser	Lys	Arg	Asn	Glu	Lys	Trp	Asp	Glu	Val	Tyr	Lys	Tyr	740	745	750	
Ile	Val	Thr	Asn	Trp	Leu	Ala	Lys	Val	Asn	Thr	Gln	Ile	Asp	Leu	Ile	755	760	765	

-continued

Arg	Lys	Lys	Met	Lys	Glu	Ala	Leu	Glu	Asn	Gln	Ala	Glu	Ala	Thr	Lys
770						775					780				
Ala	Ile	Ile	Asn	Tyr	Gln	Tyr	Asn	Gln	Tyr	Thr	Glu	Glu	Glu	Lys	Asn
785					790					795					800
Asn	Ile	Asn	Phe	Asn	Ile	Asp	Asp	Leu	Ser	Ser	Lys	Leu	Asn	Glu	Ser
			805						810					815	
Ile	Asn	Lys	Ala	Met	Ile	Asn	Ile	Asn	Lys	Phe	Leu	Asn	Gln	Cys	Ser
			820					825					830		
Val	Ser	Tyr	Leu	Met	Asn	Ser	Met	Ile	Pro	Tyr	Gly	Val	Lys	Arg	Leu
		835					840					845			
Glu	Asp	Phe	Asp	Ala	Ser	Leu	Lys	Asp	Ala	Leu	Leu	Lys	Tyr	Ile	Tyr
	850					855					860				
Asp	Asn	Arg	Gly	Thr	Leu	Ile	Gly	Gln	Val	Asp	Arg	Leu	Lys	Asp	Lys
865					870					875					880
Val	Asn	Asn	Thr	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Gln	Leu	Ser	Lys	Tyr
			885						890					895	
Val	Asp	Asn	Gln	Arg	Leu	Leu	Ser	Thr	Leu	Asp					
		900						905							

<210> SEQ ID NO 95

<211> LENGTH: 913

<212> TYPE: PRT

<213> ORGANISM: Artificial Sequence

<220> FEATURE:

<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 95

Gly	Ser	Glu	Phe	Met	Pro	Ile	Thr	Ile	Asn	Asn	Phe	Asn	Tyr	Ser	Asp
1				5					10					15	
Pro	Val	Asp	Asn	Lys	Asn	Ile	Leu	Tyr	Leu	Asp	Thr	His	Leu	Asn	Thr
			20				25						30		
Leu	Ala	Asn	Glu	Pro	Glu	Lys	Ala	Phe	Arg	Ile	Thr	Gly	Asn	Ile	Trp
	35					40					45				
Val	Ile	Pro	Asp	Arg	Phe	Ser	Arg	Asn	Ser	Asn	Pro	Asn	Leu	Asn	Lys
	50					55				60					
Pro	Pro	Arg	Val	Thr	Ser	Pro	Lys	Ser	Gly	Tyr	Tyr	Asp	Pro	Asn	Tyr
65				70					75					80	
Leu	Ser	Thr	Asp	Ser	Asp	Lys	Asp	Thr	Phe	Leu	Lys	Glu	Ile	Ile	Lys
			85					90						95	
Leu	Phe	Lys	Arg	Ile	Asn	Ser	Arg	Glu	Ile	Gly	Glu	Glu	Leu	Ile	Tyr
	100						105						110		
Arg	Leu	Ser	Thr	Asp	Ile	Pro	Phe	Pro	Gly	Asn	Asn	Asn	Thr	Pro	Ile
	115					120						125			
Asn	Thr	Phe	Asp	Phe	Asp	Val	Asp	Phe	Asn	Ser	Val	Asp	Val	Lys	Thr
	130					135					140				
Arg	Gln	Gly	Asn	Asn	Trp	Val	Lys	Thr	Gly	Ser	Ile	Asn	Pro	Ser	Val
145					150					155					160
Ile	Ile	Thr	Gly	Pro	Arg	Glu	Asn	Ile	Ile	Asp	Pro	Glu	Thr	Ser	Thr
			165					170						175	
Phe	Lys	Leu	Thr	Asn	Asn	Thr	Phe	Ala	Ala	Gln	Glu	Gly	Phe	Gly	Ala
	180						185						190		
Leu	Ser	Ile	Ile	Ser	Ile	Ser	Pro	Arg	Phe	Met	Leu	Thr	Tyr	Ser	Asn
	195						200					205			

-continued

Ala	Thr	Asn	Asp	Val	Gly	Glu	Gly	Arg	Phe	Ser	Lys	Ser	Glu	Phe	Cys
210					215						220				
Met	Asp	Pro	Ile	Leu	Ile	Leu	Met	His	Glu	Leu	Asn	His	Ala	Met	His
225				230						235				240	
Asn	Leu	Tyr	Gly	Ile	Ala	Ile	Pro	Asn	Asp	Gln	Thr	Ile	Ser	Ser	Val
			245						250					255	
Thr	Ser	Asn	Ile	Phe	Tyr	Ser	Gln	Tyr	Asn	Val	Lys	Leu	Glu	Tyr	Ala
		260						265					270		
Glu	Ile	Tyr	Ala	Phe	Gly	Gly	Pro	Thr	Ile	Asp	Leu	Ile	Pro	Lys	Ser
	275						280					285			
Ala	Arg	Lys	Tyr	Phe	Glu	Glu	Lys	Ala	Leu	Asp	Tyr	Tyr	Arg	Ser	Ile
290					295						300				
Ala	Lys	Arg	Leu	Asn	Ser	Ile	Thr	Thr	Ala	Asn	Pro	Ser	Ser	Phe	Asn
305				310						315				320	
Lys	Tyr	Ile	Gly	Glu	Tyr	Lys	Gln	Lys	Leu	Ile	Arg	Lys	Tyr	Arg	Phe
			325						330					335	
Val	Val	Glu	Ser	Ser	Gly	Glu	Val	Thr	Val	Asn	Arg	Asn	Lys	Phe	Val
		340						345					350		
Glu	Leu	Tyr	Asn	Glu	Leu	Thr	Gln	Ile	Phe	Thr	Glu	Phe	Asn	Tyr	Ala
	355						360					365			
Lys	Ile	Tyr	Asn	Val	Gln	Asn	Arg	Lys	Ile	Tyr	Leu	Ser	Asn	Val	Tyr
370					375						380				
Thr	Pro	Val	Thr	Ala	Asn	Ile	Leu	Asp	Asp	Asn	Val	Tyr	Asp	Ile	Gln
385				390						395				400	
Asn	Gly	Phe	Asn	Ile	Pro	Lys	Ser	Asn	Leu	Asn	Val	Leu	Phe	Met	Gly
			405						410					415	
Gln	Asn	Leu	Ser	Arg	Asn	Pro	Ala	Leu	Arg	Lys	Val	Asn	Pro	Glu	Asn
		420						425					430		
Met	Leu	Tyr	Leu	Phe	Thr	Lys	Phe	Cys	Val	Asp	Gly	Ile	Ile	Thr	Ser
	435						440					445			
Lys	Thr	Lys	Ser	Asp	Asp	Asp	Asp	Lys	Tyr	Gly	Gly	Phe	Leu	Arg	Arg
450					455						460				
Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn	Gln	Ala	Leu	Ala	Gly	Gly	Gly
465				470						475				480	
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu
			485					490					495		
Gln	Cys	Arg	Glu	Leu	Leu	Val	Lys	Asn	Thr	Asp	Leu	Pro	Phe	Ile	Gly
		500						505					510		
Asp	Ile	Ser	Asp	Val	Lys	Thr	Asp	Ile	Phe	Leu	Arg	Lys	Asp	Ile	Asn
	515						520					525			
Glu	Glu	Thr	Glu	Val	Ile	Tyr	Tyr	Pro	Asp	Asn	Val	Ser	Val	Asp	Gln
530					535						540				
Val	Ile	Leu	Ser	Lys	Asn	Thr	Ser	Glu	His	Gly	Gln	Leu	Asp	Leu	Leu
545				550						555				560	
Tyr	Pro	Ser	Ile	Asp	Ser	Glu	Ser	Glu	Ile	Leu	Pro	Gly	Glu	Asn	Gln
			565					570						575	
Val	Phe	Tyr	Asp	Asn	Arg	Thr	Gln	Asn	Val	Asp	Tyr	Leu	Asn	Ser	Tyr
		580						585					590		
Tyr	Tyr	Leu	Glu	Ser	Gln	Lys	Leu	Ser	Asp	Asn	Val	Glu	Asp	Phe	Thr
	595						600					605			
Phe	Thr	Arg	Ser	Ile	Glu	Glu	Ala	Leu	Asp	Asn	Ser	Ala	Lys	Val	Tyr

-continued

610					615					620					
Thr	Tyr	Phe	Pro	Thr	Leu	Ala	Asn	Lys	Val	Asn	Ala	Gly	Val	Gln	Gly
625					630					635					640
Gly	Leu	Phe	Leu	Met	Trp	Ala	Asn	Asp	Val	Val	Glu	Asp	Phe	Thr	Thr
				645					650					655	
Asn	Ile	Leu	Arg	Lys	Asp	Thr	Leu	Asp	Lys	Ile	Ser	Asp	Val	Ser	Ala
			660					665					670		
Ile	Ile	Pro	Tyr	Ile	Gly	Pro	Ala	Leu	Asn	Ile	Ser	Asn	Ser	Val	Arg
		675					680					685			
Arg	Gly	Asn	Phe	Thr	Glu	Ala	Phe	Ala	Val	Thr	Gly	Val	Thr	Ile	Leu
	690					695					700				
Leu	Glu	Ala	Phe	Pro	Glu	Phe	Thr	Ile	Pro	Ala	Leu	Gly	Ala	Phe	Val
705					710					715					720
Ile	Tyr	Ser	Lys	Val	Gln	Glu	Arg	Asn	Glu	Ile	Ile	Lys	Thr	Ile	Asp
				725					730					735	
Asn	Cys	Leu	Glu	Gln	Arg	Ile	Lys	Arg	Trp	Lys	Asp	Ser	Tyr	Glu	Trp
			740					745					750		
Met	Met	Gly	Thr	Trp	Leu	Ser	Arg	Ile	Ile	Thr	Gln	Phe	Asn	Asn	Ile
		755					760					765			
Ser	Tyr	Gln	Met	Tyr	Asp	Ser	Leu	Asn	Tyr	Gln	Ala	Gly	Ala	Ile	Lys
	770					775					780				
Ala	Lys	Ile	Asp	Leu	Glu	Tyr	Lys	Lys	Tyr	Ser	Gly	Ser	Asp	Lys	Glu
785					790					795					800
Asn	Ile	Lys	Ser	Gln	Val	Glu	Asn	Leu	Lys	Asn	Ser	Leu	Asp	Val	Lys
				805					810					815	
Ile	Ser	Glu	Ala	Met	Asn	Asn	Ile	Asn	Lys	Phe	Ile	Arg	Glu	Cys	Ser
			820					825					830		
Val	Thr	Tyr	Leu	Phe	Lys	Asn	Met	Leu	Pro	Lys	Val	Ile	Asp	Glu	Leu
		835					840					845			
Asn	Glu	Phe	Asp	Arg	Asn	Thr	Lys	Ala	Lys	Leu	Ile	Asn	Leu	Ile	Asp
	850					855					860				
Ser	His	Asn	Ile	Ile	Leu	Val	Gly	Glu	Val	Asp	Lys	Leu	Lys	Ala	Lys
865					870					875					880
Val	Asn	Asn	Ser	Phe	Gln	Asn	Thr	Ile	Pro	Phe	Asn	Ile	Phe	Ser	Tyr
				885					890					895	
Thr	Asn	Asn	Ser	Leu	Leu	Lys	Asp	Ile	Ile	Asn	Glu	Tyr	Phe	Asn	Leu
			900					905					910		

Asp

<210> SEQ ID NO 96
 <211> LENGTH: 1432
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 96

Gly	Ser	Leu	Val	Arg	Asp	Asp	Val	Asp	Tyr	Gln	Ile	Phe	Arg	Asp	Phe
1				5					10					15	
Ala	Glu	Asn	Lys	Gly	Lys	Phe	Phe	Val	Gly	Ala	Thr	Asp	Leu	Ser	Val
			20					25					30		
Lys	Asn	Lys	Arg	Gly	Gln	Asn	Ile	Gly	Asn	Ala	Leu	Ser	Asn	Val	Pro
		35					40					45			

-continued

Met	Ile	Asp	Phe	Ser	Val	Ala	Asp	Val	Asn	Lys	Arg	Ile	Ala	Thr	Val	50	55	60
Val	Asp	Pro	Gln	Tyr	Ala	Val	Ser	Val	Lys	His	Ala	Lys	Ala	Glu	Val	65	70	75
His	Thr	Phe	Tyr	Tyr	Gly	Gln	Tyr	Asn	Gly	His	Asn	Asp	Val	Ala	Asp	85	90	95
Lys	Glu	Asn	Glu	Tyr	Arg	Val	Val	Glu	Gln	Asn	Asn	Tyr	Glu	Pro	His	100	105	110
Lys	Ala	Trp	Gly	Ala	Ser	Asn	Leu	Gly	Arg	Leu	Glu	Asp	Tyr	Asn	Met	115	120	125
Ala	Arg	Phe	Asn	Lys	Phe	Val	Thr	Glu	Val	Ala	Pro	Ile	Ala	Pro	Thr	130	135	140
Asp	Ala	Gly	Gly	Gly	Leu	Asp	Thr	Tyr	Lys	Asp	Lys	Asn	Arg	Phe	Ser	145	150	155
Ser	Phe	Val	Arg	Ile	Gly	Ala	Gly	Arg	Gln	Leu	Val	Tyr	Glu	Lys	Gly	165	170	175
Val	Tyr	His	Gln	Glu	Gly	Asn	Glu	Lys	Gly	Tyr	Asp	Leu	Arg	Asp	Leu	180	185	190
Ser	Gln	Ala	Tyr	Arg	Tyr	Ala	Ile	Ala	Gly	Thr	Pro	Tyr	Lys	Asp	Ile	195	200	205
Asn	Ile	Asp	Gln	Thr	Met	Asn	Thr	Glu	Gly	Leu	Ile	Gly	Phe	Gly	Asn	210	215	220
His	Asn	Lys	Gln	Tyr	Ser	Ala	Glu	Glu	Leu	Lys	Gln	Ala	Leu	Ser	Gln	225	230	235
Asp	Ala	Leu	Thr	Asn	Tyr	Gly	Val	Leu	Gly	Asp	Ser	Gly	Ser	Pro	Leu	245	250	255
Phe	Ala	Phe	Asp	Lys	Gln	Lys	Asn	Gln	Trp	Val	Phe	Leu	Gly	Thr	Tyr	260	265	270
Asp	Tyr	Trp	Ala	Gly	Tyr	Gly	Lys	Lys	Ser	Trp	Gln	Glu	Trp	Asn	Ile	275	280	285
Tyr	Lys	Lys	Glu	Phe	Ala	Asp	Lys	Ile	Lys	Gln	His	Asp	Asn	Ala	Gly	290	295	300
Thr	Val	Lys	Gly	Asn	Gly	Glu	His	His	Trp	Lys	Thr	Thr	Gly	Thr	Asn	305	310	315
Ser	His	Ile	Gly	Ser	Thr	Ala	Val	Arg	Leu	Ala	Asn	Asn	Glu	Gly	Asp	325	330	335
Ala	Asn	Asn	Gly	Gln	Asn	Val	Thr	Phe	Glu	Asp	Asn	Gly	Thr	Leu	Val	340	345	350
Leu	Asn	Gln	Asn	Ile	Asn	Gln	Gly	Ala	Gly	Gly	Leu	Phe	Phe	Lys	Gly	355	360	365
Asp	Tyr	Thr	Val	Lys	Gly	Ala	Asn	Asn	Asp	Ile	Thr	Trp	Leu	Gly	Ala	370	375	380
Gly	Ile	Asp	Val	Ala	Asp	Gly	Lys	Lys	Val	Val	Trp	Gln	Val	Lys	Asn	385	390	395
Pro	Asn	Gly	Asp	Arg	Leu	Ala	Lys	Ile	Gly	Lys	Gly	Thr	Leu	Glu	Ile	405	410	415
Asn	Gly	Thr	Gly	Val	Asn	Gln	Gly	Gln	Leu	Lys	Val	Gly	Asp	Gly	Thr	420	425	430
Val	Ile	Leu	Asn	Gln	Lys	Ala	Asp	Ala	Asp	Lys	Lys	Val	Gln	Ala	Phe	435	440	445

-continued

Ser	Gln	Val	Gly	Ile	Val	Ser	Gly	Arg	Gly	Thr	Leu	Val	Leu	Asn	Ser
450						455					460				
Ser	Asn	Gln	Ile	Asn	Pro	Asp	Asn	Leu	Tyr	Phe	Gly	Phe	Arg	Gly	Gly
465					470					475					480
Arg	Leu	Asp	Ala	Asn	Gly	Asn	Asp	Leu	Thr	Phe	Glu	His	Ile	Arg	Asn
				485					490					495	
Val	Asp	Glu	Gly	Ala	Arg	Ile	Val	Asn	His	Asn	Thr	Asp	His	Ala	Ser
			500					505					510		
Thr	Ile	Thr	Leu	Thr	Gly	Lys	Ser	Leu	Ile	Thr	Asn	Pro	Asn	Ser	Leu
		515					520					525			
Ser	Val	His	Ser	Ile	Gln	Asn	Asp	Tyr	Asp	Glu	Asp	Asp	Tyr	Ser	Tyr
	530					535					540				
Tyr	Tyr	Arg	Pro	Arg	Arg	Pro	Ile	Pro	Gln	Gly	Lys	Asp	Leu	Tyr	Tyr
545					550					555					560
Lys	Asn	Tyr	Arg	Tyr	Tyr	Ala	Leu	Lys	Ser	Gly	Gly	Arg	Leu	Asn	Ala
				565					570					575	
Pro	Met	Pro	Glu	Asn	Gly	Val	Ala	Glu	Asn	Asn	Asp	Trp	Ile	Phe	Met
			580					585					590		
Gly	Tyr	Thr	Gln	Glu	Glu	Ala	Arg	Lys	Asn	Ala	Met	Asn	His	Lys	Asn
		595					600					605			
Asn	Arg	Arg	Ile	Gly	Asp	Phe	Gly	Gly	Phe	Phe	Asp	Glu	Glu	Asn	Gly
	610					615					620				
Lys	Gly	His	Asn	Gly	Ala	Leu	Asn	Leu	Asn	Phe	Asn	Gly	Lys	Ser	Ala
625					630					635					640
Gln	Lys	Arg	Phe	Leu	Leu	Thr	Gly	Gly	Ala	Asn	Leu	Asn	Gly	Lys	Ile
				645					650					655	
Ser	Val	Thr	Gln	Gly	Asn	Val	Leu	Leu	Ser	Gly	Arg	Pro	Thr	Pro	His
			660				665						670		
Ala	Arg	Asp	Phe	Val	Asn	Lys	Ser	Ser	Ala	Arg	Lys	Asp	Ala	His	Phe
		675					680					685			
Ser	Lys	Asn	Asn	Glu	Val	Val	Phe	Glu	Asp	Asp	Trp	Ile	Asn	Arg	Thr
	690					695					700				
Phe	Lys	Ala	Ala	Glu	Ile	Ala	Val	Asn	Gln	Ser	Ala	Ser	Phe	Ser	Ser
705					710					715					720
Gly	Arg	Asn	Val	Ser	Asp	Ile	Thr	Ala	Asn	Ile	Thr	Ala	Thr	Asp	Asn
			725					730					735		
Ala	Lys	Val	Asn	Leu	Gly	Tyr	Lys	Asn	Gly	Asp	Glu	Val	Cys	Val	Arg
		740						745					750		
Ser	Asp	Tyr	Thr	Gly	Tyr	Val	Thr	Cys	Asn	Thr	Gly	Asn	Leu	Ser	Asp
	755						760					765			
Lys	Ala	Leu	Asn	Ser	Phe	Asp	Ala	Thr	Arg	Ile	Asn	Gly	Asn	Val	Asn
	770					775					780				
Leu	Asn	Gln	Asn	Ala	Ala	Leu	Val	Leu	Gly	Lys	Ala	Ala	Leu	Trp	Gly
785					790					795					800
Lys	Ile	Gln	Gly	Gln	Gly	Asn	Ser	Arg	Val	Ser	Leu	Asn	Gln	His	Ser
				805					810					815	
Lys	Trp	His	Leu	Thr	Gly	Asp	Ser	Gln	Val	His	Asn	Leu	Ser	Leu	Ala
			820					825					830		
Asp	Ser	His	Ile	His	Leu	Asn	Asn	Ala	Ser	Asp	Ala	Gln	Ser	Ala	Asn
		835					840					845			
Lys	Tyr	His	Thr	Ile	Lys	Ile	Asn	His	Leu	Ser	Gly	Asn	Gly	His	Phe

-continued

850	855	860
His Tyr Leu Thr Asp 865	Leu Ala Lys Asn Leu Gly 870	Asp Lys Val Leu Val 875
Lys Glu Ser Ala Ser 885	Gly His Tyr Gln Leu His Val 890	Gln Asn Lys Thr 895
Gly Glu Pro Asn Gln Glu Gly Leu Asp 900	Leu Phe Asp Ala Ser Ser Val 905	
Gln Asp Arg Ser Arg Leu Phe Val Ser 915	Leu Ala Asn His Tyr Val Asp 920	
Leu Gly Ala Leu Arg Tyr Thr Ile Lys Thr 930	Glu Asn Gly Ile Thr Arg 935	
Leu Tyr Asn Pro Tyr Ala Gly Asn Gly Arg 945	Pro Val Lys Pro Ala Pro 950	
Cys Val Asp Gly Ile Ile Thr Ser Lys Thr 965	Lys Ser Leu Asp Asp Asp 970	
Asp Lys Tyr Gly Gly Phe Leu Arg Arg Ile Arg 980	Pro Lys Leu Lys Trp 985	
Asp Asn Gln Ala Leu Ala Gly Gly Gly Gly Ser 995	Gly Gly Gly Gly Ser 1000	
Gly Gly Gly Gly Ser Ala Leu Val Leu Gln Cys Ile 1010	Lys Val Asn 1015	
Asn Trp Asp Leu Phe Phe Ser Pro Ser Glu Asp 1025	Asn Phe Thr Asn 1030	
Asp Leu Asn Lys Gly Glu Glu Ile Thr Ser Asp Thr 1040	Asn Ile Glu 1045	
Ala Ala Glu Glu Asn Ile Ser Leu Asp Leu Ile Gln 1055	Gln Tyr Tyr 1060	
Leu Thr Phe Asn Phe Asp Asn Glu Pro Glu Asn Ile 1070	Ser Ile Glu 1075	
Asn Leu Ser Ser Asp Ile Ile Gly Gln Leu Glu Leu 1085	Met Pro Asn 1090	
Ile Glu Arg Phe Pro Asn Gly Lys Lys Tyr Glu Leu 1100	Asp Lys Tyr 1105	
Thr Met Phe His Tyr Leu Arg Ala Gln Glu Phe Glu 1115	His Gly Lys 1120	
Ser Arg Ile Ala Leu Thr Asn Ser Val Asn Glu Ala 1130	Leu Leu Asn 1135	
Pro Ser Arg Val Tyr Thr Phe Phe Ser Ser Asp Tyr 1145	Val Lys Lys 1150	
Val Asn Lys Ala Thr Glu Ala Ala Met Phe Leu Gly 1160	Trp Val Glu 1165	
Gln Leu Val Tyr Asp Phe Thr Asp Glu Thr Ser Glu 1175	Val Ser Thr 1180	
Thr Asp Lys Ile Ala Asp Ile Thr Ile Ile Ile Pro 1190	Tyr Ile Gly 1195	
Pro Ala Leu Asn Ile Gly Asn Met Leu Tyr Lys Asp 1205	Asp Phe Val 1210	
Gly Ala Leu Ile Phe Ser Gly Ala Val Ile Leu Leu 1220	Glu Phe Ile 1225	
Pro Glu Ile Ala Ile Pro Val Leu Gly Thr Phe Ala 1235	Leu Val Ser 1240	

-continued

Tyr Ile Ala Asn Lys Val Leu Thr Val Gln Thr Ile Asp Asn Ala
 1250 1255 1260
 Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu Val Tyr Lys Tyr Ile
 1265 1270 1275
 Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln Ile Asp Leu Ile
 1280 1285 1290
 Arg Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala Glu Ala Thr
 1295 1300 1305
 Lys Ala Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu Glu Glu
 1310 1315 1320
 Lys Asn Asn Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys Leu
 1325 1330 1335
 Asn Glu Ser Ile Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu
 1340 1345 1350
 Asn Gln Cys Ser Val Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr
 1355 1360 1365
 Gly Val Lys Arg Leu Glu Asp Phe Asp Ala Ser Leu Lys Asp Ala
 1370 1375 1380
 Leu Leu Lys Tyr Ile Tyr Asp Asn Arg Gly Thr Leu Ile Gly Gln
 1385 1390 1395
 Val Asp Arg Leu Lys Asp Lys Val Asn Asn Thr Leu Ser Thr Asp
 1400 1405 1410
 Ile Pro Phe Gln Leu Ser Lys Tyr Val Asp Asn Gln Arg Leu Leu
 1415 1420 1425
 Ser Thr Leu Asp
 1430

<210> SEQ ID NO 97
 <211> LENGTH: 911
 <212> TYPE: PRT
 <213> ORGANISM: Artificial Sequence
 <220> FEATURE:
 <223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 97

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

-continued

Asp	Pro	Ser	Gly	Ala	Thr	Thr	Lys	Ser	Ala	Met	Leu	Thr	Asn	Leu	Ile
145					150					155					160
Ile	Phe	Gly	Pro	Gly	Pro	Val	Leu	Asn	Lys	Asn	Glu	Val	Arg	Gly	Ile
				165					170					175	
Val	Leu	Arg	Val	Asp	Asn	Lys	Asn	Tyr	Phe	Pro	Cys	Arg	Asp	Gly	Phe
			180					185					190		
Gly	Ser	Ile	Met	Gln	Met	Ala	Phe	Cys	Pro	Glu	Tyr	Val	Pro	Thr	Phe
		195				200						205			
Asp	Asn	Val	Ile	Glu	Asn	Ile	Thr	Ser	Leu	Thr	Ile	Gly	Lys	Ser	Lys
	210					215					220				
Tyr	Phe	Gln	Asp	Pro	Ala	Leu	Leu	Leu	Met	His	Glu	Leu	Ile	His	Val
225					230					235					240
Leu	His	Gly	Leu	Tyr	Gly	Met	Gln	Val	Ser	Ser	His	Glu	Ile	Ile	Pro
			245					250						255	
Ser	Lys	Gln	Glu	Ile	Tyr	Met	Gln	His	Thr	Tyr	Pro	Ile	Ser	Ala	Glu
			260					265					270		
Glu	Leu	Phe	Thr	Phe	Gly	Gly	Gln	Asp	Ala	Asn	Leu	Ile	Ser	Ile	Asp
	275						280					285			
Ile	Lys	Asn	Asp	Leu	Tyr	Glu	Lys	Thr	Leu	Asn	Asp	Tyr	Lys	Ala	Ile
	290					295					300				
Ala	Asn	Lys	Leu	Ser	Gln	Val	Thr	Ser	Cys	Asn	Asp	Pro	Asn	Ile	Asp
305					310					315					320
Ile	Asp	Ser	Tyr	Lys	Gln	Ile	Tyr	Gln	Gln	Lys	Tyr	Gln	Phe	Asp	Lys
			325					330						335	
Asp	Ser	Asn	Gly	Gln	Tyr	Ile	Val	Asn	Glu	Asp	Lys	Phe	Gln	Ile	Leu
			340					345					350		
Tyr	Asn	Ser	Ile	Met	Tyr	Gly	Phe	Thr	Glu	Ile	Glu	Leu	Gly	Lys	Lys
	355						360					365			
Phe	Asn	Ile	Lys	Thr	Arg	Leu	Ser	Tyr	Phe	Ser	Met	Asn	His	Asp	Pro
	370					375					380				
Val	Lys	Ile	Pro	Asn	Leu	Leu	Asp	Asp	Thr	Ile	Tyr	Asn	Asp	Thr	Glu
385					390					395					400
Gly	Phe	Asn	Ile	Glu	Ser	Lys	Asp	Leu	Lys	Ser	Glu	Tyr	Lys	Gly	Gln
			405						410					415	
Asn	Met	Arg	Val	Asn	Thr	Asn	Ala	Phe	Arg	Asn	Val	Asp	Gly	Ser	Gly
			420					425					430		
Leu	Val	Ser	Lys	Leu	Ile	Gly	Leu	Cys	Val	Asp	Gly	Ile	Ile	Thr	Ser
	435						440					445			
Lys	Thr	Lys	Ser	Asp	Asp	Asp	Lys	Tyr	Gly	Gly	Phe	Leu	Arg	Arg	
	450					455				460					
Ile	Arg	Pro	Lys	Leu	Lys	Trp	Asp	Asn	Gln	Ala	Leu	Ala	Gly	Gly	Gly
465					470					475					480
Gly	Ser	Gly	Gly	Gly	Gly	Ser	Gly	Gly	Gly	Gly	Ser	Ala	Leu	Val	Leu
			485					490						495	
Gln	Cys	Ile	Lys	Val	Asn	Asn	Trp	Asp	Leu	Phe	Phe	Ser	Pro	Ser	Glu
			500					505					510		
Asp	Asn	Phe	Thr	Asn	Asp	Leu	Asn	Lys	Gly	Glu	Glu	Ile	Thr	Ser	Asp
	515						520					525			
Thr	Asn	Ile	Glu	Ala	Ala	Glu	Glu	Asn	Ile	Ser	Leu	Asp	Leu	Ile	Gln
	530					535					540				

-continued

Gln Tyr Tyr Leu Thr	Phe Asn Phe Asp Asn	Glu Pro Glu Asn Ile Ser
545	550	555 560
Ile Glu Asn Leu Ser	Ser Asp Ile Ile Gly Gln Leu Glu Leu Met Pro	
565	570	575
Asn Ile Glu Arg Phe	Pro Asn Gly Lys Lys Tyr Glu Leu Asp Lys Tyr	
580	585	590
Thr Met Phe His Tyr	Leu Arg Ala Gln Glu Phe Glu His Gly Lys Ser	
595	600	605
Arg Ile Ala Leu Thr	Asn Ser Val Asn Glu Ala Leu Leu Asn Pro Ser	
610	615	620
Arg Val Tyr Thr Phe	Phe Ser Ser Asp Tyr Val Lys Lys Val Asn Lys	
625	630	635 640
Ala Thr Glu Ala Ala	Met Phe Leu Gly Trp Val Glu Gln Leu Val Tyr	
645	650	655
Asp Phe Thr Asp Glu	Thr Ser Glu Val Ser Thr Thr Asp Lys Ile Ala	
660	665	670
Asp Ile Thr Ile Ile	Ile Pro Tyr Ile Gly Pro Ala Leu Asn Ile Gly	
675	680	685
Asn Met Leu Tyr Lys	Asp Asp Phe Val Gly Ala Leu Ile Phe Ser Gly	
690	695	700
Ala Val Ile Leu Leu	Glu Phe Ile Pro Glu Ile Ala Ile Pro Val Leu	
705	710	715 720
Gly Thr Phe Ala Leu	Val Ser Tyr Ile Ala Asn Lys Val Leu Thr Val	
725	730	735
Gln Thr Ile Asp Asn	Ala Leu Ser Lys Arg Asn Glu Lys Trp Asp Glu	
740	745	750
Val Tyr Lys Tyr Ile	Val Thr Asn Trp Leu Ala Lys Val Asn Thr Gln	
755	760	765
Ile Asp Leu Ile Arg	Lys Lys Met Lys Glu Ala Leu Glu Asn Gln Ala	
770	775	780
Glu Ala Thr Lys Ala	Ile Ile Asn Tyr Gln Tyr Asn Gln Tyr Thr Glu	
785	790	795 800
Glu Glu Lys Asn Asn	Ile Asn Phe Asn Ile Asp Asp Leu Ser Ser Lys	
805	810	815
Leu Asn Glu Ser Ile	Asn Lys Ala Met Ile Asn Ile Asn Lys Phe Leu	
820	825	830
Asn Gln Cys Ser Val	Ser Tyr Leu Met Asn Ser Met Ile Pro Tyr Gly	
835	840	845
Val Lys Arg Leu Glu	Asp Phe Asp Ala Ser Leu Lys Asp Ala Leu Leu	
850	855	860
Lys Tyr Ile Tyr Asp	Asn Arg Gly Thr Leu Ile Gly Gln Val Asp Arg	
865	870	875 880
Leu Lys Asp Lys Val	Asn Asn Thr Leu Ser Thr Asp Ile Pro Phe Gln	
885	890	895
Leu Ser Lys Tyr Val	Asp Asn Gln Arg Leu Leu Ser Thr Leu Asp	
900	905	910

What is claimed is:

1. A single chain, polypeptide fusion protein, comprising:
 - a. a non-cytotoxic protease, or a fragment thereof, which protease or protease fragment cleaves a protein of the exocytic fusion apparatus of a nociceptive sensory afferent;
 - b. a dynorphin Targeting Moiety that binds to a Binding Site on the nociceptive sensory afferent, which Binding Site endocytoses to be incorporated into an endosome within the nociceptive sensory afferent;
 - c. a protease cleavage site at which site the fusion protein is cleavable by a protease, wherein the protease cleavage site is located between the non-cytotoxic protease or fragment thereof and the Targeting Moiety;
 - d. a translocation domain that translocates the protease or protease fragment from within an endosome, across the endosomal membrane and into the cytosol of the nociceptive sensory afferent; and
 wherein the Targeting Moiety is located between the protease cleavage site and the translocation domain.
2. The fusion protein according to claim 1, wherein the Targeting Moiety and the protease cleavage site are separated by at most 10 amino acid residues, or by at most 5 amino acid residues, or by at most zero amino acid residues.
3. The fusion protein according to claim 1, wherein the non-cytotoxic protease is a clostridial neurotoxin L-chain or an IgA protease.
4. The fusion protein according to claim 1, wherein the translocation domain is the H_N domain of a clostridial neurotoxin.
5. The fusion protein according to claim 1, wherein the Targeting Moiety binds specifically to the ORL_1 receptor.
6. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 14 or 16 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 5 or 6 conservative amino acid substitutions.
7. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 14 or 16 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 3 or 4 conservative amino acid substitutions.
8. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 14 or 16 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 1 or 2 conservative amino acid substitutions.
9. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 10 or 12 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 5 or 6 conservative amino acid substitutions.
10. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 10 or 12 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 3 or 4 conservative amino acid substitutions.

11. The fusion protein according to claim 1, wherein the Targeting Moiety comprises or consists of an amino acid sequence according to SEQ ID NO. 89 or a fragment comprising or consisting of at least 10 or 12 contiguous amino acid residues thereof, or a variant amino acid sequence of said SEQ ID NO: 89 or said fragment having a maximum of 1 or 2 conservative amino acid substitutions.

12. A polypeptide fusion protein comprising any one of SEQ ID NOs: 91, 92, 93, 94, 95, or 96, or a fusion protein having at least 90% or 95% sequence identity therewith.

13. A polynucleotide molecule comprising a nucleic acid sequence encoding the polypeptide fusion protein according to claim 1.

14. An expression vector, which comprises a promoter, the polynucleotide molecule according to claim 13, wherein said polynucleotide molecule is located downstream of the promoter, and a terminator located downstream of the polynucleotide molecule.

15. A polynucleotide molecule comprising a nucleic acid sequence that is the complement of the nucleic acid sequence according to claim 13.

16. A method for preparing a single-chain polypeptide fusion protein, comprising:

- a. transfecting a host cell with the expression vector of claim 14, and
- b. culturing said host cell under conditions promoting expressing of the polypeptide fusion protein by the expression vector.

17. A method of preparing a non-cytotoxic agent, comprising:

- a. contacting a single-chain polypeptide fusion protein according to claim 1 with a protease capable of cleaving the protease cleavage site;
- b. cleaving the protease cleavage site; and thereby forming a di-chain fusion protein.

18. A non-cytotoxic polypeptide, obtained by the method of claim 17, wherein the polypeptide is a di-chain polypeptide, and wherein:

- a. the first chain comprises the non-cytotoxic protease, or a fragment thereof, which protease or protease fragment is capable of cleaving a protein of the exocytic fusion apparatus of a nociceptive sensory afferent;
- b. the second chain comprises the dynorphin TM and the translocation domain that is capable of translocating the protease or protease fragment from within an endosome, across the endosomal membrane and into the cytosol of the nociceptive sensory afferent; and

the first and second chains are disulphide linked together.

19. A method of treating, preventing or ameliorating pain in a subject, comprising administering to said patient a therapeutically effective amount of the fusion protein according to claim 1.

20. A method according to claim 19, wherein the pain is chronic pain selected from neuropathic pain, inflammatory pain, headache pain, somatic pain, visceral pain, and referred pain.

21. A method of treating, preventing or ameliorating pain in a subject, comprising administering to said patient a therapeutically effective amount of a polypeptide according to claim 18.

22. A method according to claim 21, wherein the pain is chronic pain selected from neuropathic pain, inflammatory pain, headache pain, somatic pain, visceral pain, and referred pain.