

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
27 April 2006 (27.04.2006)

PCT

(10) International Publication Number
WO 2006/044378 A2

(51) International Patent Classification:
G06F 19/00 (2006.01)

(21) International Application Number:
PCT/US2005/036521

(22) International Filing Date: 12 October 2005 (12.10.2005)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/618,211 12 October 2004 (12.10.2004) US
60/676,500 29 April 2005 (29.04.2005) US

(71) Applicant (for all designated States except US): **UNIVERSITY OF IOWA RESEARCH FOUNDATION** [US/US]; 100 Oakdale Campus, Iowa City, IA 52242 (US).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **ELCOCK, Adrian, H.** [US/US]; 1551 Ranier Drive, Iowa City, IA 52246 (US). **ROCKEY, William, M.** [US/US]; 2865 Coral Ct., Apt. 104, Coralville, IA 52241 (US).

(74) Agents: **HUIZENGA, David, E.** et al.; Needle & Rosenberg, P.C., 999 Peachtree Street Suite 1000, Atlanta, GA 30309-3915 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— without international search report and to be republished upon receipt of that report

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: RAPID COMPUTATIONAL IDENTIFICATION OF TARGETS

(57) Abstract: Disclosed are compositions and methods for rapid computational identification of targets.

WO 2006/044378 A2

Rapid Computational Identification of Targets

CROSS REFERENCE TO RELATED PATENT APPLICATIONS

1. This application claims priority to U.S. Provisional Application No. 60/618,211 filed October 12, 2004 and U.S. Provisional Application No. 60/676,500 filed April 29, 2005, both of which are herein incorporated by reference in their entireties.

I. BACKGROUND

2. The vast majority of drugs show a high correlation of structure and specificity to produce pharmacological effects. Experimental evidence indicates that drugs interact with receptor sites localized in macromolecules which have protein-like properties and specific three dimensional shapes. Often three points of attachment or interaction of a drug to a receptor site are preferred. In most cases a rather specific chemical structure is required for the receptor site and a complementary drug structure. Slight changes in the molecular structure of the drug can drastically change specificity.

3. It is desirable to be able to identify new targets for existing drugs. Current experimental approaches to this problem have included 'protein chips', on which protein targets are arrayed and assayed (3), and proteomics techniques capable of detecting proteins that bind to drug analogues covalently attached to a column (4). What is needed in the art is a computational method for identifying the protein receptors likely to bind a drug, which can provide accurate predictions of the drug's ability to bind to each homologue of the receptor.

II. SUMMARY

4. Disclosed are methods related to the identification of targets for a given molecule. Also disclosed are methods of inhibiting a receptor with a molecule and identifying molecules that interact and modulate receptors. Also disclosed are methods of making a pharmaceutical composition.

III. BRIEF DESCRIPTION OF THE DRAWINGS

5. The accompanying drawings, which are incorporated in and constitute a part of this specification, illustrate several embodiments and together with the description illustrate the disclosed compositions and methods.

5 6. **Figure 1** shows the dependence of the classification efficiency on the atomic van der Waals radius used in the computations. Results are shown for the seven drugs studied.

7. **Figure 2** shows the distribution of binding energies obtained from a computational screen of imatinib with 493 human protein kinases. The shaded area
10 indicates the 22 kinases predicted to bind the drug on the basis of their computed binding energies.

8. **Figure 3** shows a subset of the sidechain rotamers sampled around the drug imatinib. Millions of rotamer combinations of the residues within 5.0 Å of the drug (shown in green) are sampled with Monte Carlo methods in order to compute
15 each drug-receptor binding energy.

9. **Figure 4** shows phylogenetic trees showing the relationships between 'sequences' constructed from the drug-binding residues of each of the ~20 tested kinases. Results are shown for the five drugs for which the computations appeared successful. Those kinases correctly predicted by the computations to be targets of a given drug are
20 boxed red; those kinases falsely predicted to be targets are boxed blue.

10. **Figure 5** is similar to Figure 4, but shows results for the two drugs for which the computations appeared unsuccessful.

11. **Figure 6** shows the distribution of binding energies obtained from a computational screen of (left) Purvalanol B, and (right) SB 203580 with 493 human
25 protein kinases.

12. **Figure 7** shows the correlation between computed binding energies and experimental IC₅₀ values for (left) Purvalanol B, and (right) hymenialdisine. TP denotes true positive, FP false positive, TN true negative, and FN false negative.

13. **Figure 8** shows an illustrative example of the dependence of computed results
30 on the energy function.

14. **Figure 9** shows the distribution of classification efficiencies of the testing sets for (A) SB 203580, (B) purvalanol B, and (C) imatinib. The computed testing set

classification efficiencies computed from the model are shown as dark bars; those obtained from randomized trials are shown by the white bars.

15. **Figure 10** shows a flow diagram illustrating exemplary steps in a disclosed method.

5 16. **Figure 11** shows a flow diagram illustrating exemplary steps in a disclosed method.

17. **Figure 12** shows a flow diagram illustrating exemplary steps in a disclosed method.

10 18. **Figure 13** shows a flow diagram illustrating exemplary steps in a disclosed method.

19. **Figure 14** shows a distribution of classification efficiencies of the testing sets obtained from SCR calculations.

20. **Figure 15** shows a distribution of testing set classification efficiencies obtained from AutoDock calculations with the fixed inhibitor assumption.

15 21. **Figure 16** shows a superimposed views of SB203580 in a model of the p38 α binding site before and after GROMACS energy minimization.

22. **Figure 17** shows a comparison of the inhibitor-kinase contacts made by SB203580 and the crystal structures 1PME (mutant Erk2) and 1A9U (p38ALPHA).

20 23. **Figure 18** shows a superimposed view of SB203580 in the binding sites of the crystal structures 1PME and 1A9U.

24. **Figure 19** shows an ordered list of the computed binding energies obtained with SCR for five inhibitors for which the calculations were successful.

25. **Figure 20** shows an ordered list of the computed binding energies obtained from docking calculations with AutoDock.

25 26. **Figure 21** shows an ordered list of inhibitors that have experimental data in the form of percentage activity.

27. **Figure 22** shows an ordered list of the computed binding energies obtained with SCR (Side Chain Rotamer program) for five inhibitors for which the calculations were successful.

30 28. **Figure 23** shows a summary of training/setting results for various screening protocols applied to five inhibitors and their respective kinase panels.

29. **Figure 24** shows an ordered list of the computed binding energies obtained using AutoDock and the fixed-inhibitor assumption.

30. **Figure 25** shows an ordered list of the computed binding energies obtained from docking calculations conducted with AutoDock.

31. **Figure 26** shows an ordered list of the computed binding energies obtained by applying AutoDock's energy function to complexes that were first energy-minimized by GROMACS.

32. **Figure 27** shows an ordered list of the computed binding energies obtained with SCR (Side Chain Rotamer program).

IV. DETAILED DESCRIPTION

33. Before the present compounds, compositions, articles, devices, and/or methods are disclosed and described, it is to be understood that they are not limited to specific synthetic methods or specific recombinant biotechnology methods unless otherwise specified, or to particular reagents unless otherwise specified, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting.

A. Definitions

34. As used in the specification and the appended claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a pharmaceutical carrier" includes mixtures of two or more such carriers, and the like.

35. Ranges can be expressed herein as from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, another embodiment includes from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, by use of the antecedent "about," it will be understood that the particular value forms another embodiment. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint. It is also understood that there are a number of values disclosed herein, and that each value is also herein disclosed as "about" that particular value in addition to the value itself. For example, if the value "10" is disclosed, then "about 10" is also disclosed. It is also understood that when a value is disclosed that "less than or equal to" the value, "greater than or equal to the value" and possible ranges between values are also disclosed, as appropriately understood by the skilled artisan. For example, if the value "10" is disclosed then "less than or equal to 10"

as well as “greater than or equal to 10” is also disclosed. It is also understood that throughout the application, data is provided in a number of different formats, and that this data, represents endpoints and starting points, and ranges for any combination of the data points. For example, if a particular data point “10” and a particular data point “15” are disclosed, it is understood that greater than, greater than or equal to, less than, less than or equal to, and equal to 10 and 15 are considered disclosed as well as between 10 and 15.

36. In this specification and in the claims which follow, reference will be made to a number of terms which shall be defined to have the following meanings:

37. “Optional” or “optionally” means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

38. The terms “higher,” “increases,” “elevates,” or “elevation” refer to increases above basal levels, or as compared to a control. The terms “low,” “lower,” “inhibits,” “reduces,” or “reduction” refer to decreases below basal levels, or as compared to a control. For example, basal levels are normal *in vivo* levels prior to, or in the absence of, addition of an agent that binds a receptor.

39. As used throughout, “potential target” refers to any molecule capable of interacting with another molecule. Examples of potential targets include, but are not limited to, kinases, nuclear receptors, phosphatases, phosphodiesterases, transferases (such as methyl transferases and glycotransferases), serine proteases, oxidoreductases, hydrolases, esterases, glycosyl hydrolases, ribonucleases, lyases, isomerases, G-coupled protein receptors, and ligases.

40. As used throughout, “molecule” refers to any compound which is capable of interacting with another molecule. An example of a “molecule” used in this context includes, but is not limited, to proteins and drugs. The terms “protein,” “drug” and “molecule” can be used interchangeably throughout, except where explicitly indicated otherwise.

41. As used throughout, “known target” refers to any molecule whose interaction with a molecule as described above, is known. Typically a ‘known target’ is a protein. Also typically the interaction between the molecule and the known target is sufficiently strong to produce a therapeutic response.

42. The term "associated with" means that there has been a link or correlation between the items discussed. For example, a particular receptor might be associated with a disease. This would mean that the receptor has been linked or is correlated with the presence of the disease. It can also mean that the receptor has been shown to be wholly or in part causative of the disease.

43. Throughout this application, various publications are referenced. The disclosures of these publications in their entireties are hereby incorporated by reference into this application in order to more fully describe the state of the art to which this pertains. The references disclosed are also individually and specifically incorporated by reference herein for the material contained in them that is discussed in the sentence in which the reference is relied upon.

B. Methods

44. In general the methods disclosed allow for the identification of targets for a molecule. It is understood that a target can be a receptor, protein, or any other type of molecule, but often is an amino acid based molecule, such as a protein. It is understood that the molecule can also be anything that can interact with the target, meaning it could be a small molecule, a nucleic acid, or even, an amino acid based molecule, such as a protein. Often the molecule, can be, for example, a drug. Often the molecule will have some type of activity, such as modulation of a protein activity, such as reduction or activation, such as an antagonist or agonist. In these instances, for example, the molecule could be referred to as an active molecule. While the examples disclosed herein, and the discussion regarding the compositions and methods, may use one or more different descriptions, such as molecule or drug or target or receptor in describing a particular embodiment, it is understood that the general nature of the methods applies to any two compositions regardless of what they are called, provided they function as in the methods as disclosed herein.

45. Many therapeutic drugs act by binding a protein receptor (target). Drugs that are designed to activate a receptor are known as agonists. Drugs that are designed to inactivate a receptor are known as antagonists, or blockers, and often act by inhibiting the protein-receptor interaction that would have otherwise occurred at that site. Often, a drug known to bind one receptor also binds other receptors in a subject. Generally speaking, the more closely related the receptors are, the higher the probability of the drug binding

the related receptor. This degree of relatedness can be measured by comparing homology or sequence similarity between the known target and potential targets.

46. The binding of related receptors by a drug can either be an advantage or a disadvantage. When advantageous, a drug known to bind one receptor, and therefore treat one condition or disease, can also bind another receptor and therefore treat another condition or disease. This is of enormous advantage because often the drug has already been shown to be safe and has been approved for use by the FDA. The binding of related receptors becomes a disadvantage when the binding does not serve a useful purpose and instead causes unwanted or adverse side effects. Identifying these interactions can also be useful because the structure of the drug can then be modified to minimize the unwanted interactions. Also, since drugs react differently in different subjects, identifying the target of a drug in a subject with unwanted side effects can help establish a population that should not, or on the other hand, should have the drug administered to them.

47. Therefore, identifying other receptors that would interact with a drug is of enormous importance, both to identify potentially useful new treatments, as well as to identify potentially harmful or unwanted side effects. It is also useful in drug customization and design.

48. Several chemical forces can result in the binding of the drug to the receptor. Essentially any type of bond can be involved with the drug-receptor interaction. Covalent bonds are very tight and practically irreversible. Most drug-receptor interactions are non-covalent; covalent bond formation is rather rare. Since many drugs contain acid or amine functional groups which are ionized at physiological pH, ionic bonds can be formed by the attraction of opposite charges in the receptor site.

49. Polar-polar interactions as in hydrogen bonding are a further extension of the attraction of opposite charges. The drug-receptor reaction is essentially an exchange of the hydrogen bond between a drug molecule, surrounding water, and the receptor site.

50. Finally, hydrophobic bonds can be formed between non-polar hydrocarbon groups on the drug and those in the receptor site. These bonds are not very specific but can make a major contribution to the strength of the drug/receptor interaction.

51. Repulsive forces which decrease the stability of the drug-receptor interaction include repulsion of like charges and steric hindrance. Steric hindrance refers to certain

3-dimensional features where repulsion occurs between electron clouds, inflexible chemical bonds, or bulky alkyl groups.

1. Identifying Potential Targets

52. Described herein are methods of identifying potential targets of a molecule.

5 The methods involve some basic similarities. Typically the method first utilizes a 3-dimensional structure of the known target with the molecule, such as a drug. This known structure can have been determined using any known means, such as crystallography or solution NMR spectroscopy. That structure can also be obtained through computer molecular modeling simulation programs, such as AutoDock. The methods typically
10 involve determining the amount of binding, such as determining the binding energy, between a molecule, such as an active molecule, such as a drug, and a potential target for that molecule. An active molecule, is a molecule that has some activity against a target, such as inhibiting a target's activity or enhancing the target's activity. In addition, the potential target is typically a composition, such as a receptor, which has some genetic
15 relationship, such as homology or identity, to a known target for the molecule.

53. Typically, the percentage identity of the sequences of the known target and potential target can be viewed in number of ways. For example, one can look at the identity between the entire known target and the potential target. One can also look at the identity between the potential target and the known target only in the domain where the
20 drug or molecule binds, for example, a kinase domain. One can also look at the identity between the potential target molecule and the known target at the level of a sub-domain, such as only those residues in the potential target which are within 7Å, 6Å, 5Å, 4Å, 3Å, or 2Å of a residue which is in contact with the molecule in the known target. Generally, the more specific the sub-domain the higher the identity will be between the
25 amino acids of the potential target and the known target. For example, in one embodiment there may be 30% (could be 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%) or greater identity between the known target and potential target as a whole, 50% (could be 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%) or greater identity between the drug binding domain of the known target and the potential
30 target, and 70% (could be 75%, 80%, 85%, 90%, 95%) or greater identity between the residues of the potential target that correspond to the residues of the known target which are within 5Å of a residue which interacts with the drug. Another sub-domain is a sub-

domain of residues which actually contact the drug. In this case the identity is typically greater than 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or higher.

54. Typically, the potential target exists in a family of potential targets, i.e. a set of potential targets, all of which have some genetic relationship, such as homology or identity, to the known target for the molecule. A family consisting of any number of members may be screened. The maximum number of members in the family is only limited by the amount of computer power available to screen each member in a desired amount of time. The methods involve at least one template structure of the molecule and a target, often this would be with a known target. It is not required that this structure be existent, as it can be generated, in some cases during the disclosed methods, using standard structure determination techniques. It is preferred that a real structure exist at the time the methods are employed.

55. "High resolution" means a resolution of perhaps 3.0Å or smaller in a crystal structure. Structures of any resolution, such as, 6.0Å, 5.0Å, 4.0Å, 3.0Å, 2.0Å or smaller can be employed in the disclosed methods. For example, structures of resolutions of 1.75 Å (1OPJ), 2.0 Å (1PME), 2.05 Å (1CKP), 2.10 Å (1DM2), and 2.30 Å have all been successfully used.

56. It is also typical that the methods involve modeling the structure of the potential target, using information from the structure of the known target. This modeling can be performed in any way, and as described herein.

57. Often, the backbone of the region which has the genetic relationship and which is in the region of the known target that interacts with the molecule, is held constant in the potential target, relative to the backbone of the known target, when the potential target is modeled using the structure information of the known target. The structure of the entire backbone of the potential receptor is not required: all that is required is a structure for the backbone for residues that are within 7Å, 6Å, 5Å, 4Å, 3Å, or 2Å of an atom of the drug. For example the backbone residues in the immediate vicinity of the drug in the high resolution structure of the drug in complex with a known target. "Immediate vicinity" means any receptor residue that has an atom within 5Å of an atom of the drug.

58. The sidechains of the amino acids can be added initially to the fixed backbone using a simple sidechain-adding program such as SCWRL3.0 (A. A. Canutescu, A. A. Shelenkov, and R. L. Dunbrack, Jr. A graph theory algorithm for protein side-chain

prediction. *Protein Science* 12, 2001-2014 (2003). A program such as SCWRL can be used to build an initial model of the target receptor. Once this has been constructed, one can decide which sidechains should be allowed to move during the binding energy calculations.

5 59. One parameter that is decided at some point during the disclosed methods is the parameter called side chain movement. In the disclosed methods, certain side chains are held fixed and certain side chains are allowed to move, such as to be sampled. Thus, one way of determining if a side chain is a fixed side chain is by determining the distance the side chain is away from an atom of the drug. For example, sidechains that have all
10 atoms more than 7Å, 6Å, 5Å, 4Å, or 3Å from any atom of the drug can be side chains that are fixed. Another example, sidechains that have an atom within 7Å, 6Å, 5Å, 4Å, or 3Å of any atom of the drug can be allowed to move, and sidechains that do not meet this criterion are held fixed.

15 60. In other embodiments, the methods involve holding fixed the side chains of the amino acids of the potential and known targets that are not directly involved in binding the drug. Sidechains that have at least one atom within 7Å, 6Å, 5Å, 4Å, or 3Å of any atom of the drug in the initial model constructed as discussed herein, are side chains which can be considered involved in drug binding. Side chains which do not meet the criteria for an involved side chain are considered side chains not involved in drug
20 binding.

 61. Side chains determined to be involved in binding can be allowed to move and can sample different conformational positions from rotamer libraries, by for example, a Monte Carlo sampling procedure. Side chains determined to not to be involved in drug binding can be held fixed.

25 62. The conformation and position of the drug can be held fixed during the calculations; that is, it may be assumed that the drug binds in exactly the same orientation to the potential target as it does to a known target. For flexible drug molecules, rotamer libraries similar to those used for describing receptor sidechain flexibility can be used to model alternative drug conformations.

30 63. Then, a binding energy can be determined between the molecule and the potential target, and if the binding energy meets certain criteria, then the potential target can be designated as an actual target, i.e. one that is likely to be biologically modulated by a molecule-actual target interaction. The criterion can be that the computed binding

energy of the molecule with the potential target is similar to, or more favorable than, the computed binding energy of the same molecule with a known target. For example, an actual target can be a target where the computed binding energy as discussed herein is, for example, at least 60%, 65%, 70%, 75%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 100%, 101%, 102%, 103%, 104%, 105%, 106%, 107%, 108%, 109%, 110%, 120%, 130%, 140%, 150%, 200%, 250%, 300%, 350%, 400%, 450%, 500%, 600%, 700%, 800%, 900%, 1000%, or greater than that of the known target binding energy. An actual target can also be a target which after ordering all potential targets in terms of the strength of their binding energies, are the targets which are in the top 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1% of computed binding strengths, of for example, a set of potential targets where the set is at least 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 125, 150, 175, 200, 225, 250, 275, 300, 350, 400, 500, 700, or a 1000 potential targets.

64. It is also understood that once a potential target is identified, as disclosed herein, traditional testing and analysis can be performed, such as performing a biological assay using the molecule and the actual target to further define the ability of the molecule to modulate the actual target. The disclosed methods can include the step of assaying the biological activity of the molecule and potential target, as well as performing, for example, combinatorial chemistry studies using libraries based on the molecule, for example.

65. Energy calculations can be based on molecular or quantum mechanics. Molecular mechanics approximates the energy of a system by summing a series of empirical functions representing components of the total energy like bond stretching, van der Waals forces, or electrostatic interactions. Quantum mechanics methods use various degrees of approximation to solve the Schrodinger equation. These methods deal with electronic structure, allowing for the characterization of chemical reactions.

66. Potential targets of the molecule can be identified. This can occur by selecting potential targets with a given similarity to the known target. For example, sequence information can be used to compare relative homologies or similarities. Homologous, or similar, sequences can be identified, for example, using *SWISS-PROT*, *PIR* (1-3), *GenBank* and *NRL-3D*. *SWISS-PROT*. The sequences can be compared using,

for example, <http://www.bioinfo.biocenter.helsinki.fi:8080/dali/index.html>, or
<http://us.expasy.org/spdbv/>. Alternatively, targets in the same family as the known target
can be selected. For example, if a known molecule-target interaction occurs wherein the
target is a kinase, other members of the kinase family can be selected as well as potential
5 targets.

67. To prepare each drug structure for calculation, atoms can be built in that were
unresolved or absent from the crystal structures of the drug. This can be done, for
example, using the PRODRG webserver

<http://www.davapc1.bioch.dundee.ac.uk./programs/prodrgr>, or standard molecular

10 modeling programs such as InsightII or Quanta (both at www.accelrys.com), or any other
molecular modeling system capable of preparing the drug structure.

68. An accurate structural model of the potential target can then be elucidated.

Typically, the potential targets to be tested are modeled with the backbone in an identical
conformation to that of the known target-molecule crystal structure or solution structure.

15 Typically the next step is to construct structural models using, for example, sequence
alignment. For certain families of receptors, sequence alignments can be taken directly
from a sequence database dedicated to that particular family. For example, the Kinase
Sequence Database (KSD) contains a curated alignment of the ATP-catalytic domains of
over 7000 kinases. Other such databases exist for other families of receptors. Examples
20 of these databases include but are not limited to the Cytochrome P450 Homepage
(<http://drnelson.utmem.edu/CytochromeP450.html>) for cytochrome P450s; The EF-Hand
Calcium Binding Proteins Data Library

(http://structbio.vanderbilt.edu/cabp_database/cabp.html) for calcium-binding proteins;

The Glucocorticoid Receptor Resource (<http://nrr.georgetown.edu/GRR/GRR.HTML>)

25 for the glucocorticoid receptor; The Kinesin Homepage

(<http://www.proweb.org/kinesin/>) for kinesins; Alignments of RecA Genes and Proteins

(<http://www.tigr.org/~jeisen/RecA/RecA.Alignment.html>) for the RecA protein; and the

GPCRDB (<http://www.gpcr.org/7tm/>) for G protein coupled receptors. If a pre-existing
sequence alignment is not available, or if the sequence of the potential or known targets

30 are not present in the pre-existing sequence alignment, a sequence alignment of the
potential and known target sequences can be constructed using standard multiple
sequence alignment programs such as CLUSTALW (J. D. Thompson et al. CLUSTAL-

W – Improving the sensitivity of progressive multiple sequence alignment through

sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673-4680 (1994)) or any similar method known to those skilled in the art such as MAFFT (K. Katoh *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* **30**, 3059-3066 (2002)) or DCA (J. Stoye. Multiple sequence alignment with the Divide-and-Conquer method. *Gene* **211**, GC45-GC56 (1998)). With the sequence of the potential target aligned with that of the known, structurally-characterized target, a structure file for each potential target to be tested can be created.

69. To complete preparation of the structure of the potential target, sidechains can be added. This can be done, for example, by using the rotamer-modeling program SCWRL 3.0 (A. A. Canutescu, A. A. Shelenkov, and R. L. Dunbrack, Jr. A graph theory algorithm for protein side-chain prediction. *Protein Science* **12**, 2001-2014 (2003), or any similar method known to those skilled in the art, for example, the method of Liang & Grishin (S. D. Liang, and N. V. Grishin. Side-chain modeling with an optimized scoring function. *Protein Science* **11**, 322-331 (2004) or the SCAP method of Xiang & Honig (Z. X. Xiang, and B. Honig. Extending the accuracy limits of prediction for side-chain conformations. *J. Mol. Biol.* **311**, 421-430 (2001)). Fast and accurate side-chain conformation prediction is important for homology modeling, ab initio protein structure prediction, and protein design applications. The SCWRL program is an example of one method, and is widely used because of its speed, accuracy, and ease of use, and any program performing functions such as those performed by SCWRL can be used. Some of the functions performed by SCWRL are, for example, SCWRL uses results from graph theory to solve the combinatorial problem encountered in the side-chain prediction problem. In this method, side chains are represented as vertices in an undirected graph. Any two residues that have rotamers with nonzero interaction energies are considered to have an edge in the graph. The resulting graph can be partitioned into connected subgraphs with no edges between them. These subgraphs can in turn be broken into biconnected components, which are graphs that cannot be disconnected by removal of a single vertex. The combinatorial problem is reduced to finding the minimum energy of these small biconnected components and combining the results to identify the global minimum energy conformation. This algorithm is able to complete predictions on a set of 180 proteins with 34342 side chains in <7 min of computer time. The total chi(1) and chi(1 + 2) dihedral angle accuracies are 82.6% and 73.7% using a

simple energy function based on the backbone-dependent rotamer library and a linear repulsive steric energy. The new algorithm allows for use of SCWRL in sequence design and ab initio structure prediction, as well addition of complex energy function and conformational flexibility.

5 70. Hydrogens can also be added using methods such as the hydrogen bond optimization module (HBOND) of the modeling program WHATIF or corresponding modules in any standard molecular modeling program known to those skilled in the art such as InsightII (Accelrys) or Sybyl (Tripos, Inc.). When WHATIF determines if a hydrogen bond can be formed between the hydrogen of the donor atom and the lone pair
10 of the acceptor atom, it uses four parameters. These are: 1) Distance between the donor and acceptor atom. 2) Distance between the (calculated) hydrogen position, and the acceptor atom. 3) Angle from donor atom over the hydrogen to the acceptor atom. And 4) Angle from the hydrogen over the acceptor to a 'virtual' atom. If the acceptor is only covalently bound to one atom, this atom is the so-called virtual atom. If the acceptor is
15 covalently bound to two atoms, the virtual atom is on the bisector of those two.

71. Hydrogen bonds can be placed according to the following algorithm: If the geometry fixes the hydrogen position, this position is used, whereby the donor hydrogen distance is set to 1.0 Angstrom. If the hydrogen has a degree of rotational freedom, then the cone on which the hydrogen can potentially be found is calculated. This cone has a
20 top angle of one hundred twenty degrees. The hydrogen is now placed on the two points that this cone has in common with the plane through the donor, a point on the rotation axis of the cone and the acceptor. WHAT IF only uses hydrogens that can be involved in hydrogen bonds. The cysteine side chain is not considered for hydrogen bond calculations.

25 72. Any constellation that creates a donor/hydrogen/acceptor triplet that falls within the four values described above can be accepted as a hydrogen bond. This program, as well as the accompanying manual, can be found at <http://www.cmbi.kun.nl/whatif/> (*WHAT IF: A molecular modeling and drug design program*. G.Vriend, J. Mol. Graph. (1990) 8, 52-56.)

30 73. The binding affinity of the potential target and molecule can then be calculated. There are numerous means for carrying this out. For example, the sampling of sidechain positions and the computation of the binding thermodynamics can be accomplished using an empirical function that models the energy of the potential target-

molecule as a sum of electrostatic and van der Waals interactions between all pairs of atoms within the model. Any other computationally fast method for scoring the binding affinity of the drug with the potential target molecule can be used (H. Gohlke, & G. Klebe. Approaches to the description and prediction of the binding affinity of small-molecule ligands to macromolecular receptors. *Angew. Chem. Int. Ed.* **41**, 2644-4676 (2002)). Examples of such scoring methods include, but are not limited to, those implemented in programs such as AutoDock (G. M. Morris *et al.* Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639-1662 (1998)), Gold (G. Jones *et al.* Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation. *J. Mol. Biol.* **245**, 43-53 (1995)), Chem-Score (M. D. Eldridge *et al.* *J. Comput.-Aided Mol. Des.* **11**, 425-445 (1997)) and Drug-Score (H. Gohlke *et al.* Knowledge-based scoring function to predict protein-ligand interactions. *J. Mol. Biol.* **295**, 337-356 (2000)).

74. Flexibility can be incorporated into the sidechains of residues of the potential target that are close to the molecule through the use of rotamer libraries that are sampled by Monte Carlo (MC) methods or can be incorporated by sampling sidechain conformations with molecular dynamics (MD) simulations. In the case of using MC methods to sample sidechain rotamers, a typical simulation step can comprise (a) selecting one of the residues close to the drug at random, (b) selecting a new rotamer (conformation) for the sidechain of the selected residue at random, (c) evaluating the energy of the drug-receptor complex with the new conformation of the receptor using one of the methods listed above, and (d) applying a Metropolis test, known to those skilled in the art, to determine whether or not to accept the newly generated sidechain conformation based on the difference in energy between the newly generated conformation and the conformation generated in the previous simulation step. An entire simulation can comprise millions of such simulation steps, with the calculated energy being some average of the individual energies computed at each step of the simulation. The computed binding energy of the drug with the potential target can then be the difference between the average energy of the drug-target complex and the average energy of the target alone.

75. Rotamer libraries are known to those of skill in the art and can be obtained from a variety of sources, including the internet. Rotamers are low energy side-chain conformations. The use of a library of rotamers allows for the modeling of a structure to

try the most likely side-chain conformations, saving time and producing a structure that is more likely to be correct. The use of a library of rotamers can be restricted to those residues that are within a given region of the potential target, for example, at the drug binding site, or within a specified distance of the drug. The latter distance can be set at any desired length, for example, the potential target can be 2, 3, 4, 5, 6, 7, 8, or 9 Å from any atom of the molecule.

76. Electrostatic interactions between every pair of atoms can be calculated, for example, using a Coulombic model with the formula:

77. $E_{\text{elec}} = 332.08 \frac{q_1 q_2}{\epsilon r}$, where q_1 and q_2 are partial atomic charges, r is the distance between them, and ϵ is the dielectric constant.

78. Partial atomic charges can be taken from existing parameter sets that have been developed to describe charge distributions in proteins. Example parameter sets include, but are not limited to, PARSE (D. A. Sitkoff *et al.* Accurate calculation of hydration free-energies using macroscopic solvent models. *J. Phys. Chem.* **98**, 1978-1988 (1994)), CHARMM (MacKerell *et al.* All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586-3616, 1998) and AMBER (W. D. Cornell *et al.* A 2nd generation force-field for the simulation of proteins, nucleic-acids, and organic-molecules. *J. Am. Chem. Soc.* **117**, 5179-5195 (1995)). Partial charges for atoms of the drug molecule can be assigned either by analogy with those of similar functional groups found in proteins, or by empirical assignment methods such as that implemented in the PRODRG server (D. M. F. van Aalten *et al.* PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput.-Aided Mol. Design* **10**, 255-262 (1996)), or by the use of standard quantum mechanical calculation methods (for example, C. I. Bayly *et al.* A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges – the RESP model. *J. Phys. Chem.* **97**, 10269-10280, (1993)).

79. The electrostatic interaction can also be calculated by more elaborate methodologies that incorporate electrostatic desolvation effects. These can include explicit solvent and implicit solvent models: in the former, water molecules are directly included in the calculations, whereas in the latter, the effects of water are described by a dielectric continuum approach. Specific examples of implicit solvent methods for calculating electrostatic interactions include but are not limited to: Poisson-Boltzmann based methods and Generalized Born methods (M. Feig & C. L. Brooks. Recent

advances in the development and application of implicit solvent models in biomolecule simulations. *Curr. Opin. Struct. Biol.* **14**, 217-224 (2004)).

80. van der Waals and hydrophobic interactions between pairs of atoms (where both atoms are either sulfur or carbon) can be calculated using a simple Lennard-Jones formalism with the following equation:

81. $E_{vdw} = \epsilon \{ \sigma_{att}^{12}/r^{12} - \sigma_{att}^6/r^6 \}$. where ϵ is an energy, r is the distance between the two atoms and σ_{att} is the distance at which the energy of interaction is zero.

82. van der Waals interactions between pairs of atoms (where one or both atoms are neither sulfur nor carbon) can be calculated using a simple repulsive energy term:

83. $E_{vdw} = \epsilon \{ \sigma_{rep}^{12}/r^{12} \}$. where ϵ is an energy, r is the distance between the two atoms and σ_{rep} determines the distance at which the repulsive interaction is equal to ϵ .

84. Hydrophobic interactions between atoms can also be calculated using a variety of other methods known to those skilled in the art. For example, the energetic contribution can be calculated as being proportional to the amount of solvent accessible surface area of the ligand and receptor that is buried when the complex is formed. Such contributions can be expressed in terms of interactions between pairs of atoms, such as in the method proposed by Street & Mayo (A. G. Street & S. L. Mayo. Pairwise calculation of protein solvent-accessible surface areas. *Folding & Design* **3**, 253-258 (1998)). Any other implementation of a formalism for describing hydrophobic or van der Waals or other energetic contributions can be included in the calculations.

85. Binding energies can be calculated for each potential target-molecule interaction. For example, Monte Carlo sampling of the flexible sidechains in the receptor can be conducted in the presence and absence of the molecule, and the average energy in each simulation calculated. A binding energy for the ligand (molecule) with the receptor can then be calculated as the difference between the two calculated average energies.

86. The computed binding energy of a potential target with the drug can be compared with the computed binding energy of a known target with the drug to determine if the potential target is likely to be a real target. These results can then be confirmed using experimental data, wherein the actual interaction between the molecule and potential target can be measured. Examples of methods that can be used to determine an actual interaction between the molecule and the potential target include but are not limited to: equilibrium dialysis measurements (wherein binding of a radioactive form of drug to the target is detected), enzyme inhibition assays (wherein the enzymatic activity

of a receptor enzyme can be monitored in the presence and absence of the drug), and chemical shift perturbation measurements (wherein binding of the drug to the receptor is monitored by observing changes in NMR chemical shifts of atoms in the receptor).

87. Described herein, and illustrated in FIG. 10, is a method of identifying a target for a molecule comprising the steps: a) modeling the molecule in complex with a known target for the molecule (1001), b) obtaining potential target molecules by selecting potential target molecules with a defined homology to the known target (1002), and c) determining the binding affinity of a potential target with the molecule by modeling the potential target with the molecule, wherein side chain rotamers are sampled during homology modeling (1003).

88. As illustrated in FIG. 11, also disclosed is a method of identifying a target for a molecule comprising the steps: a) obtaining a structural model of the molecule and a known target, wherein the known target comprises a known target-molecule binding domain (1101), b) obtaining a potential target by identifying potential targets having a defined homology with the known target (1102), c) performing homology modeling with the identified potential target, wherein during the homology modeling the backbone conformations are held identical to the known target, wherein the sidechains are sampled from a library of rotamers (1103), and d) calculating a binding energy of the molecule and the identified potential target (1104).

89. Also disclosed, and illustrated in FIG. 12, is a method of identifying a desired protein-molecule interaction comprising: a) determining structural information for a protein known to interact with the molecule of interest (1201); b) identifying which residues of the protein of step a) interact with the molecule (1202); c) comparing the residues identified in step b) with a database of proteins (1203); d) selecting proteins having an area of similarity to the residues identified in step b) (1204); e) calculating interaction energies between the proteins of step d) and the molecule of interest (1205); and f) determining which proteins are capable of interacting in a desired fashion with the molecule of interest (1206).

2. Methods of making a pharmaceutical composition

a) Compositions identified by screening with disclosed compositions / combinatorial chemistry

(1) Combinatorial chemistry

5 90. The disclosed methods and systems can be used for any combinatorial technique to identify molecules or macromolecular molecules that interact with the disclosed compositions in a desired way. For example, the disclosed methods for identifying targets for molecules, can identify a molecule-target pair, and this molecule-target pair interaction or activity can be modified, such as enhanced, by using the disclosed combinatorial techniques with a library related to the molecule to identify variants of the molecule that have even better or more desirable activity between the original molecule and target. Once the target is identified, the disclosed methods can also be used to identify molecules, such as a functional nucleic acid, which would have characteristics similar or more desirable, for example, than the original molecule and identified target. The nucleic acids, peptides, and related molecules disclosed herein can be used as targets for the combinatorial approaches.

10 91. It is understood that when using the disclosed compositions in combinatorial techniques or screening methods, molecules, such as macromolecular molecules, will be identified that have particular desired properties such as inhibition or stimulation or the target molecule's function. The molecules identified and isolated when using the disclosed compositions, such as kinases and other proteins and systems, are also disclosed. Thus, the products produced using the combinatorial or screening approaches that involve the disclosed compositions, such as kinases, are also considered herein disclosed.

20 92. Combinatorial chemistry includes but is not limited to all methods for isolating small molecules or macromolecules that are capable of binding either a small molecule or another macromolecule, typically in an iterative process. Proteins, oligonucleotides, and sugars are examples of macromolecules. For example, oligonucleotide molecules with a given function, catalytic or ligand-binding, can be isolated from a complex mixture of random oligonucleotides in what has been referred to as "in vitro genetics" (Szostak, TIBS 19:89, 1992). One synthesizes a large pool of molecules bearing random and defined sequences and subjects that complex mixture, for example, approximately 10^{15} individual sequences in 100 μ g of a 100 nucleotide RNA, to

some selection and enrichment process. Through repeated cycles of affinity chromatography and PCR amplification of the molecules bound to the ligand on the column, Ellington and Szostak (1990) estimated that 1 in 10^{10} RNA molecules folded in such a way as to bind a small molecule dyes. DNA molecules with such ligand-binding behavior have been isolated as well (Ellington and Szostak, 1992; Bock et al, 1992). Techniques aimed at similar goals exist for small organic molecules, proteins, antibodies and other macromolecules known to those of skill in the art. Screening sets of molecules for a desired activity whether based on small organic libraries, oligonucleotides, or antibodies is broadly referred to as combinatorial chemistry. Combinatorial techniques are particularly suited for defining binding interactions between molecules and for isolating molecules that have a specific binding activity, often called aptamers when the macromolecules are nucleic acids.

93. There are a number of methods for isolating proteins which either have de novo activity or a modified activity. For example, phage display libraries have been used to isolate numerous peptides that interact with a specific target. (See for example, United States Patent No. 6,031,071; 5,824,520; 5,596,079; and 5,565,332 which are herein incorporated by reference at least for their material related to phage display and methods relate to combinatorial chemistry)

94. A preferred method for isolating proteins that have a given function is described by Roberts and Szostak (Roberts R.W. and Szostak J.W. Proc. Natl. Acad. Sci. USA, 94(23)12997-302 (1997). This combinatorial chemistry method couples the functional power of proteins and the genetic power of nucleic acids. An RNA molecule is generated in which a puromycin molecule is covalently attached to the 3'-end of the RNA molecule. An *in vitro* translation of this modified RNA molecule causes the correct protein, encoded by the RNA to be translated. In addition, because of the attachment of the puromycin, a peptidyl acceptor which cannot be extended, the growing peptide chain is attached to the puromycin which is attached to the RNA. Thus, the protein molecule is attached to the genetic material that encodes it. Normal *in vitro* selection procedures can now be done to isolate functional peptides. Once the selection procedure for peptide function is complete traditional nucleic acid manipulation procedures are performed to amplify the nucleic acid that codes for the selected functional peptides. After amplification of the genetic material, new RNA is transcribed with puromycin at the 3'-end, new peptide is translated and another functional round of selection is performed.

Thus, protein selection can be performed in an iterative manner just like nucleic acid selection techniques. The peptide which is translated is controlled by the sequence of the RNA attached to the puromycin. This sequence can be anything from a random sequence engineered for optimum translation (i.e. no stop codons etc.) or it can be a degenerate sequence of a known RNA molecule to look for improved or altered function of a known peptide. The conditions for nucleic acid amplification and in vitro translation are well known to those of ordinary skill in the art and are preferably performed as in Roberts and Szostak (Roberts R.W. and Szostak J.W. Proc. Natl. Acad. Sci. USA, 94(23)12997-302 (1997)).

95. Another preferred method for combinatorial methods designed to isolate peptides is described in Cohen et al. (Cohen B.A., et al., Proc. Natl. Acad. Sci. USA 95(24):14272-7 (1998)). This method utilizes and modifies two-hybrid technology. Yeast two-hybrid systems are useful for the detection and analysis of protein:protein interactions. The two-hybrid system, initially described in the yeast *Saccharomyces cerevisiae*, is a powerful molecular genetic technique for identifying new regulatory molecules, specific to the protein of interest (Fields and Song, Nature 340:245-6 (1989)). Cohen et al., modified this technology so that novel interactions between synthetic or engineered peptide sequences could be identified which bind a molecule of choice. The benefit of this type of technology is that the selection is done in an intracellular environment. The method utilizes a library of peptide molecules that attached to an acidic activation domain. A peptide of choice, for example a portion of a kinase is attached to a DNA binding domain of a transcriptional activation protein, such as Gal 4. By performing the Two-hybrid technique on this type of system, molecules that bind the portion of a kinase can be identified.

96. Using methodology well known to those of skill in the art, in combination with various combinatorial libraries, one can isolate and characterize those small molecules or macromolecules, which bind to or interact with the desired target. The relative binding affinity of these compounds can be compared and optimum compounds identified using competitive binding studies, which are well known to those of skill in the art.

97. Techniques for making combinatorial libraries and screening combinatorial libraries to isolate molecules which bind a desired target are well known to those of skill in the art. Representative techniques and methods can be found in but are not limited to

United States patents 5,084,824, 5,288,514, 5,449,754, 5,506,337, 5,539,083, 5,545,568, 5,556,762, 5,565,324, 5,565,332, 5,573,905, 5,618,825, 5,619,680, 5,627,210, 5,646,285, 5,663,046, 5,670,326, 5,677,195, 5,683,899, 5,688,696, 5,688,997, 5,698,685, 5,712,146, 5,721,099, 5,723,598, 5,741,713, 5,792,431, 5,807,683, 5,807,754, 5,821,130, 5,831,014, 5,834,195, 5,834,318, 5,834,588, 5,840,500, 5,847,150, 5,856,107, 5,856,496, 5,859,190, 5,864,010, 5,874,443, 5,877,214, 5,880,972, 5,886,126, 5,886,127, 5,891,737, 5,916,899, 5,919,955, 5,925,527, 5,939,268, 5,942,387, 5,945,070, 5,948,696, 5,958,702, 5,958,792, 5,962,337, 5,965,719, 5,972,719, 5,976,894, 5,980,704, 5,985,356, 5,999,086, 6,001,579, 6,004,617, 6,008,321, 6,017,768, 6,025,371, 6,030,917, 6,040,193, 6,045,671, 6,045,755, 6,060,596, and 6,061,636.

98. Combinatorial libraries can be made from a wide array of molecules using a number of different synthetic techniques. For example, libraries containing fused 2,4-pyrimidinediones (United States patent 6,025,371) dihydrobenzopyrans (United States Patent 6,017,768 and 5,821,130), amide alcohols (United States Patent 5,976,894), hydroxy-amino acid amides (United States Patent 5,972,719) carbohydrates (United States patent 5,965,719), 1,4-benzodiazepin-2,5-diones (United States patent 5,962,337), cyclics (United States patent 5,958,792), biaryl amino acid amides (United States patent 5,948,696), thiophenes (United States patent 5,942,387), tricyclic Tetrahydroquinolines (United States patent 5,925,527), benzofurans (United States patent 5,919,955), isoquinolines (United States patent 5,916,899), hydantoin and thiohydantoin (United States patent 5,859,190), indoles (United States patent 5,856,496), imidazol-pyrido-indole and imidazol-pyrido-benzothiophenes (United States patent 5,856,107) substituted 2-methylene-2, 3-dihydrothiazoles (United States patent 5,847,150), quinolines (United States patent 5,840,500), PNA (United States patent 5,831,014), containing tags (United States patent 5,721,099), polyketides (United States patent 5,712,146), morpholino-subunits (United States patent 5,698,685 and 5,506,337), sulfamides (United States patent 5,618,825), and benzodiazepines (United States patent 5,288,514).

99. As used herein combinatorial methods and libraries included traditional screening methods and libraries as well as methods and libraries used in iterative processes.

100. Also disclosed herein are methods of making a pharmaceutical composition. In the methods described herein, interactions between potential targets and molecules can be found. These interactions can indicate, for example, a drug-target interaction. Once this

interaction is established, pharmaceutical compositions can be made that interact with the target. One example of a method of making a pharmaceutical comprises a) modeling the pharmaceutical in complex with a known target for the molecule; b) obtaining potential target molecules by selecting potential target molecules with a defined homology to the known target, c) determining the binding affinity of a potential target with the pharmaceutical by modeling the potential target with the pharmaceutical, wherein a Monte Carlo function is used for sampling of side chain rotamers; d) identifying target molecules of the pharmaceutical; e) synthesizing the pharmaceutical; and f) testing the pharmaceutical for binding to the target molecule.

101. It is understood that there are numerous ways in which the disclosed methods can be combined with other drug discovery mechanisms. For example, once potential targets are identified for a known molecule, as described herein, other drug discovery and selection techniques can be employed to create a library of molecules related to the known molecule with the identified target to optimize the manipulation of the identified target. These varied molecules can be tested and identified in the disclosed methods, just as the identified targets were identified using the disclosed methods.

102. For example, once a potential target of a drug has been identified, structures of closely related molecules can also be constructed using methods outlined earlier and tested for their ability to bind to the potential target with more selectivity. This can be done by for example, adding small functional groups (e.g. methyl, hydroxyl, t-butyl) to the original molecule using standard molecular modeling methods known to those skilled in the art. It can be assumed in this process that the positions of those atoms that are common to both the original drug and the modified drug will remain the same. The binding energy of the newly modified molecule with the potential target and other known targets can then be computed in order to identify molecules that bind with greater selectivity for the potential target of interest. Large numbers of possible modifications to an existing molecule can be investigated individually. In this way, a drug can be developed that binds strongly to a desired target without also binding strongly to other, undesired targets.

103. Once a modified molecule whose computed selectivity for a desired target has been designed computationally, it can be synthesized by standard organic chemistry. The computational nature of the design process will lessen the need for expensive efforts to be directed toward synthesizing modified molecules that ultimately do not have the desired selectivity.

3. Methods of inhibiting a receptor

104. Disclosed herein are methods of inhibiting a specific receptor with a known drug. Using the methods disclosed herein, various drug-target interactions can be elucidated

5 105. For example, disclosed are methods of inhibiting a receptor selected from the group consisting of MAK, FLT4, MUSK, CDK3, KDR, PCTAIRE2, CDK2, PCTAIRE1, CDC2, FLT3, CDKL1, Erk3, ICK, CDK7, TRKA, PCTAIRE3, CDC7, Erk4, GCN2, ROR1, NEK3, FLT1, NEK6, PDGFRa, FGFR2, CASK, ROR2, Erk7, NEK7, CCRK, TRKB, CDK5, DYRK1A, TRKC, MPSK1, AurA, MAP3K4, RET, 10 DYRK1B, CDK9, CDKL3, AurB, JAK2, TIE1, AurC, MSK1, PEK, MER, PFTAIRE2, PIM2, SGK, ABL, Wee1, PFTAIRE1, LMR2, CDKL2, Wee1B, PAK5, CLK3, TLK1, TLK2, PAK4, EphA1, EphA7, JAK1, MSK2, DDR1, KIT, CDK11, CDK8, FGFR3, PKCt, DDR2, SRPK2, PDGFRb, FGFR1, DYRK4, EphB3, TIE2, CDK6, Fused, PKACg, NEK9, SRPK1, TYK2, RSK1, RSK3, HCK, RSK2, RSK4, EphA6, PKCz, 15 CHED, GSK3B, DMPK2, JAK3, MRCKb, PYK2, ITK, IRAK1, PKCi, MRCKa, MLK1, MAP2K5, HRI, EphA10, DMPK1, CDKL4, YES, EphB6, and SYK comprising incubating the receptor with the drug purvalanol. Purvalanol is a known selective inhibitor of the human CDK2/cyclin A and Cdc2/cyclin B kinase complex.

20 106. Also disclosed are methods of inhibiting a receptor selected from the group consisting of EphA1, EphB3, EphB1, EphB4, RIPK3, EphB2, DDR1, FRK, DDR2, EphA8, PDGFRa, YES, BRK, BLK, MAP2K5, QIK, LYN, QSK, FGR, EphA6, HCK, PDGFRb, LCK, YANK2, EphA3, SIK, EphA4, MOK, p38a, EphA5, SRM, YANK3, YANK1, SRC, FYN, p38b, RiPK2, MLK4, EphA2, EphB6, RSK4 (Domain 2), GAK, RET, RSK3 (Domain 2), TGFbR2, BRAF, CSK, ACK, RAF1, CaMKK1, 25 HER4/ErbB4, BTK, KDR, FLT4, and KIT comprising incubating the receptor with the drug SB 203580. SB 203580 is a pyridinyl imidazole which acts as a specific inhibitor of p38 MAP Kinase. It has the chemical formula C₂₁H₁₆N₃FOS, and the chemical name 4-(4-Fluorophenyl)-2-(4-methylsulfinyl phenyl)-5-(4-pyridyl) 1H-imidazole.

30 107. Also disclosed are methods of inhibiting a receptor selected from the group consisting of FMS, TEC, MYT1, IKKb, RiPK2, RET, YES, BMX, CSK, HCK, FRK, BLK, FGR, ABL, SRC, LCK, LYN, ACK, PDGFRa, IKKa, PDGFRb, KIT, FGFR2, HER4/ErbB4, FYN, FLT1, SYK, FGFR4, FAK, FLT4, Wnk2, and Wnk3 comprising incubating the receptor with the drug imatinib. Imatinib mesylate is

designated chemically as 4-[(4-Methyl-1-piperazinyl)methyl]-N-[4-methyl-3-[[4-(3-pyridinyl)-2-pyrimidinyl]amino]-phenyl]benzamide methanesulfonate.

108. The above methods of inhibiting a receptor can also comprise the step of identifying the receptor as a target for the drug prior to inhibiting the receptor, identifying a subject in need of modulating the particular receptor, identifying a subject as having a disease where the particular receptor is involved, or diagnosing a need for modulation of the receptor, or indicating an understanding of a need for modulating the receptor or treating the subject for any of the targets or receptors or compositions described herein, alone or in any combination.

109. Table 9 shows the sequence of a number of kinases identified and discussed herein.

Name	Accession	Group	Family	Pseudogene ?	SEQ ID NO Protein	SEQ ID NO Kinase Domain
ABL	SK006	TK	Abl	N	1	33
ACK	SK009	TK	Ack	N	2	34
Wnk2	SK016	Other	Wnk	N	3	35
BLK	SK049	TK	Src	N	4	36
FMS	SK094	TK	PDGFR	N	5	37
CSK	SK095	TK	Csk	N	6	38
FAK	SK138	TK	Fak	N	7	39
FGFR2	SK144	TK	FGFR	N	8	40
FGFR4	SK147	TK	FGFR	N	9	41
FGR	SK148	TK	Src	N	10	42
FLT1	SK150	TK	VEGFR	N	11	43
FLT4	SK151	TK	VEGFR	N	12	44
FYN	SK153	TK	Src	N	13	45
HCK	SK164	TK	Src	N	14	46
HER4/ErbB4	SK168	TK	EGFR	N	15	47
IKKa	SK175	Other	IKK	N	16	48
IKKb	SK176	Other	IKK	N	17	49
KIT	SK201	TK	PDGFR	N	18	50
LCK	SK206	TK	Src	N	19	51
LYN	SK210	TK	Src	N	20	52
MYT1	SK248	Other	WEE	N	21	53
PDGFRa	SK274	TK	PDGFR	N	22	54
PDGFRb	SK275	TK	PDGFR	N	23	55
RET	SK326	TK	Ret	N	24	56
RIPK2	SK329	TKL	RIPK	N	25	57
SRC	SK357	TK	Src	N	26	58
SYK	SK363	TK	Syk	N	27	59
TEC	SK366	TK	Tec	N	28	60
YES	SK393	TK	Src	N	29	61
BMX	SK417	TK	Tec	N	30	62
FRK	SK419	TK	Src	N	31	63
Wnk3	SK641	Other	Wnk	N	32	64

4. Summary of therapeutic relevance for top 50 imatinib targets

110. The targets identified for imatinib are shown in Table 9. These targets have therapeutic relevance. For example, Table 10 shows a list of targets and their binding energy to imatinib as disclosed herein along with a non-limiting list of diseases the target is associated with.

111. Table 10

Target Name	KSD ID	Binding Energy	Commercial Availability	Therapeutic Targets
RET	5419753	-18.64	Yes	<i>Familial medullary thyroid carcinoma:</i> Kameyama K, Okinaga H, Takami H RET oncogene mutations in 75 cases of familial medullary thyroid carcinoma in Japan BIOMEDICINE & PHARMACOTHERAPY 58 (6-7): 345-347 JUL-AUG 2004: <i>Human papillary thyroid carcinoma:</i> Melillo RM, Cirafo AM, De Falco V, et al. The oncogenic activity of RET point mutants for follicular thyroid cells may account for the occurrence of papillary thyroid carcinoma in patients affected by familial medullary thyroid carcinoma AMERICAN JOURNAL OF PATHOLOGY 165 (2): 511-521 AUG 2004 <i>Multiple endocrine neoplasia type 2:</i> Arighi E, Popsueva A, Degl'Innocenti D, et al. Biological effects of the dual phenotypic Janus mutation of ret cosegregating with both multiple endocrine neoplasia type 2 and Hirschsprung's disease MOLECULAR ENDOCRINOLOGY 18 (4): 1004-1017 APR 1 2004

FMS	Sk094	-22.20	Yes	Choriocarcinoma: Chambers SK, Ivins CM, Kacinski BM, et al. An unexpected effect of glucocorticoids on stimulation of c-fms proto-oncogene expression in choriocarcinoma cells that express little glucocorticoid receptor AMERICAN JOURNAL OF OBSTETRICS AND GYNECOLOGY 190 (4): 974-983 APR 2004 Preeclampsia: Chang E, Robinson C, Johnson D, et al. Soluble FMS-like tyrosine kinase 1 levels are elevated in preeclampsia and correlate with disease severity AMERICAN JOURNAL OF OBSTETRICS AND GYNECOLOGY 189 (6): S95-S95 110 Suppl. S DEC 2003 Breast cancer: Flick MB, Sapi E, Kacinski BM Hormonal regulation of the c-fms proto-oncogene in breast cancer cells is mediated by a composite glucocorticoid response element JOURNAL OF CELLULAR BIOCHEMISTRY 85 (1): 10-23 2002 Kawasaki disease: Yasukawa K, Terai M, Shulman ST, et al. Systemic production of vascular endothelial growth factor and fms-like tyrosine kinase-1 receptor in acute Kawasaki disease CIRCULATION 105 (6): 766-769 FEB 12 2002 Colon carcinoma: Cui J, Yang DH, Bi XJ, et al. Methylation status of c-fms oncogene in HCC and its relationship with clinical pathology WORLD JOURNAL OF GASTROENTEROLOGY 7 (1): 136-139 FEB 2001 Pathological bone resorption: Al-Saffar N, Revell PA Differential expression of transforming growth factor-alpha and macrophage colony-stimulating factor/colony-stimulating factor-1R (c-fms) by multinucleated giant cells involved in pathological bone resorption at the site of orthopaedic implants JOURNAL OF ORTHOPAEDIC RESEARCH 18 (5): 800-807 SEP 2000 Myeloid malignancy: NCBI entry: Mutations in this gene have been associated with a predisposition to myeloid malignancy. http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?27262659:NCBI:5994077
-----	-------	--------	-----	---

TEC	4507429	-20.31	No	Cancer: Maltais A, Filion C, Labelle Y The AF2 domain of the orphan nuclear receptor TEC is essential for the transcriptional activity of the oncogenic fusion protein EWS/TEC CANCER LETTERS 183 (1): 87-94 SEP 8 2002 Myelodysplastic syndrome: OMIM entry: Interestingly, high expression of TEC was seen in each of 3 patients examined with myelodysplastic syndrome. http://www.ncbi.nlm.nih.gov/entrez/dispmim.cgi?id=600583
MYT1	2460023	-19.56	No	Cervical cancer (Patent 6,391,894)
IKK β	3213217	019.52	Yes	Inflammation; Scheidereit C Mechanims of gene activation by pathogenic and inflammatory stimuli through IKK and NF-kB pathways SHOCK 21: 45-45 177 Suppl. 1 2004 Breast cancer: Monks NR, Pardee AB, Baile SR, et al. Specific inhibition of NF-kB activation by dominant negative IKK-beta and siRNA in ER-negative human breast cancer cells. CLINICAL CANCER RESEARCH 9 (16): 6102S-6103S Part 2 Suppl. S DEC 1 2003 Prostate cancer: Gasparian AV, Yao YJ, Lu JX, et al. Selenium compounds inhibit I kappa B kinase (IKK) and nuclear factor-kappa B (NF-kappa B) in prostate cancer cells MOLECULAR CANCER THERAPEUTICS 1 (12): 1079-1087 OCT 2002
RIPk2	4505637	-18.66	No	Crohn's disease; arthrocutaneouveal granulomatosis (Blau syndrome); granulomatous synovitis with uveitis and cranial neuropathies (Blau syndrome) http://genecards.bcgsc.ca/cgi-bin/carddisp?CARD15&snpcount=25 Christian Stehlik , Hideki Hayashi , Frederick Pio, Adam Godzik and John C. Reed CARD6 Is a Modulator of NF-B Activation by Nod1- and Cardiak-mediated Pathways J. Biol. Chem., Vol. 278, Issue 34, 31941-31949, August 22, 2003

YES	125870	-18.38	Yes	Parkinson's disease: Nakaso K, Yoshimoto Y, Nakano T, et al. Transcriptional activation of p62/A170/ZIP during the formation of the aggregates: possible mechanisms and the role in Lewy body formation in Parkinson's disease BRAIN RESEARCH 1012 (1-2): 42-51 JUN 25 2004 Prostate cancer : Kitamura H, Torigoe T, Asanuma H, et al. Cytosolic overexpression of p62 sequestosome 1 is a novel characteristic of neoplastic prostate tissue. Differential immunohistochemical pattern of p62 expression in selected pathologic entities of the prostate gland JOURNAL OF UROLOGY 171 (4): 106-106 Suppl. S APR 2004 Breast cancer: Thompson HGR, Harris JW, Wold BJ, et al. p62 overexpression in breast tumors and regulation by prostate-derived Ets factor in breast cancer cells ONCOGENE 22 (15): 2322-2333 APR 17 2003 Paget's disease of the bone : Ciani B, Layfield R, Cavey JR, et al. Structure of the ubiquitin-associated domain of p62 (SQSTM1) and implications for mutations that cause Paget's disease of bone JOURNAL OF BIOLOGICAL CHEMISTRY 278 (39): 37409-37412 SEP 26 2003 Frontotemporal dementia: Arai T, Nonaka T, Hasegawa M, et al. Neuronal and glial inclusions in frontotemporal dementia with or without motor neuron disease are immunopositive for p62 NEUROSCIENCE LETTERS 342 (1-2): 41-44 MAY 15 2003 Mixed connective tissue disease : Kraemer DM, Kraus MR, Kneitz C, et al. Nucleoporin p62 antibodies in a case of mixed connective tissue disease CLINICAL AND DIAGNOSTIC LABORATORY IMMUNOLOGY 10 (2): 329-331 MAR 2003 Inflammation : Moscat J, Diaz-Meco MT The atypical PKC scaffold protein P62 is a novel target for anti-inflammatory and anti-cancer therapies ADVANCES IN ENZYME REGULATION 42: 173-179 2002 Alzheimer's disease: Kuusisto E, Salminen A, Alafuzoff I Early accumulation of p62 in neurofibrillary tangles in Alzheimer's disease: possible role in tangle formation NEUROPATHOLOGY AND APPLIED NEUROBIOLOGY 28 (3): 228-237 JUN 2002
-----	--------	--------	-----	--

BMX	4502435	-18.35	Yes	Breast cancer: Chen XY, Huang LM, Kung HJ, et al. The role of tyrosine kinase Etk/Bmx in EGF-induced apoptosis of MDA-MB-468 breast cancer cells ONCOGENE 23 (10): 1854-1862 MAR 11 2004 Prostate cancer: Lee LF, Guan JL, Qiu Y, et al. Neuropeptide-induced androgen independence in prostate cancer cells: Roles of nonreceptor tyrosine kinases Etk/Bmx, Src, and focal adhesion kinase MOLECULAR AND CELLULAR BIOLOGY 21 (24): 8385-8397 DEC 2001
CSK	729887	-18.32	Yes	Colon cancer: Rengifo-Cam W, Konishi A, Morishita N, et al. Csk defines the ability of integrin-mediated cell adhesion and migration in human colon cancer cells: implication for a potential role in cancer metastasis ONCOGENE 23 (1): 289-297 JAN 8 2004 Breast cancer: McShan GD, Zagozdzon R, Park SY, et al. Csk homologous kinase associates with RAFTK/Pyk2 in breast cancer cells and negatively regulates its activation and breast cancer cell migration INTERNATIONAL JOURNAL OF ONCOLOGY 21 (1): 197-205 JUL 2002
HCK	2194103	-18.14	Yes	HIV: Hanna Z, Weng XD, Kay DG, et al. The pathogenicity of human immunodeficiency virus (HIV) type 1 Nef in CD4C/HIV transgenic mice is abolished by mutation of its SH3-binding domain, and disease development is delayed in the absence of Hck JOURNAL OF VIROLOGY 75 (19): 9378-9392 OCT 2001 Encephalomyocarditis virus-induced diabetes: Choi KS, Jun HS, Kim HN, et al. Role of Hck in the pathogenesis of encephalomyocarditis virus-induced diabetes in mice JOURNAL OF VIROLOGY 75 (4): 1949-1957 FEB 2001 Cancer: Howlett CJ, Bisson SA, Resek ME, et al. The proto-oncogene p120(Cbl) is a downstream substrate of the Hck protein-tyrosine kinase BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 257 (1): 129-138 APR 2 1999

FRK	4503787	-18.10	No	OMIM entry: PCR analysis detected highest expression of FRK in breast and lung cancer cell lines, with minimal expression in hemopoietic cell lines. Expression in lung tumors was variable. <u>Anneren et al. (2000)</u> showed that expression of Gtk, the rodent homolog of FRK, in a rat pheochromocytoma cell line used as a model for neuronal cell differentiation induced nerve growth factor (see <u>162030</u>)-independent neurite outgrowth and Rap1 (<u>179520</u>) activation, probably through activation of the CrkII (<u>164762</u>)-C3G (GRF2; <u>600303</u>) pathway. http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=606573
BLK	914204	-18.03	Yes	OMIM entry: Expression of BLK in immature T cells suggested that it may play an important role in thymopoiesis. <u>Appel et al. (2002)</u> constructed a physical and transcription map of the critical region for keratolytic winter erythema (KWE; <u>148370</u>), an autosomal dominant skin disorder mapped to chromosome 8p23-p22. The BLK gene was identified in the BAC contig between microsatellite markers D8S1695 and D8S1759. The gene is comprised of 13 exons spanning about 70 kb. In contrast to the results of <u>Drebin et al. (1995)</u>, <u>Appel et al. (2002)</u> detected transcription of BLK in lymphoblastoid cell lines, spleen, liver, leukocytes, ovary, muscle, and testis. Mutation screening of each exon by direct sequencing of genomic DNA from KWE patients did not reveal any pathogenic mutation. Because BLK is a member of the SRC family, which is thought to play an important role in the signaling pathways controlling cell proliferation and differentiation, the authors considered the gene to be a good positional candidate for the cancers mapping to this region. http://www.ncbi.nlm.nih.gov/entrez/dispomim.cgi?id=191305

FGR	125358	-17.95	Yes	Small cell lung cancer: Zarn JA, Zimmermann SM, Pass MK, et al. Association of CD24 with the kinase c-fgr in a small cell lung cancer cell line and with the kinase lyn in an erythroleukemia cell line BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 225 (2): 384-391 AUG 14 1996 Epstein-Barr Virus: CHEAH MSC, PARMLEY RT, SUMAYA CV, et al. C-FGR ONCOGENE EXPRESSION IN HEMOPHILIA WITH ELEVATED EPSTEIN-BARR-VIRUS (EBV) SEROLOGY AND AIDS-LYMPHOMAS PEDIATRIC RESEARCH 27 (4): A140-A140 Part 2 APR 1990 AIDS: CHEAH MSC, PARMLEY RT, SUMAYA CV, et al. C-FGR ONCOGENE EXPRESSION IN HEMOPHILIA WITH ELEVATED EPSTEIN-BARR-VIRUS (EBV) SEROLOGY AND AIDS-LYMPHOMAS PEDIATRIC RESEARCH 27 (4): A140-A140 Part 2 APR 1990 Burkitt's lymphoma: SHARP NA, LUSCOMBE MJ, CLEMENS MJ REGULATION OF C-FGR PROTO-ONCOGENE EXPRESSION IN BURKITT'S-LYMPHOMA CELLS - EFFECT OF INTERFERON TREATMENT AND RELATIONSHIP TO EBV STATUS AND C-MYC MESSENGER-RNA LEVELS ONCOGENE 4 (8): 1043-1046 AUG 1989
ABL	625609	-17.91	Yes	Chronic myelogenous leukemia: Tsukahara F, Maru Y Mechanism of stability and degradation of the chronic myelogenous leukemia oncoprotein, BCR-ABL JOURNAL OF PHARMACOLOGICAL SCIENCES 94: 105P-105P Suppl. 1 2004

SRC	125711	-17.90	Yes	Breast cancer: Rossi AM, Capiati DA, Picotto G, et al. MAPK inhibition by 1 alpha,25(OH)(2)-vitamin D-3 in breast cancer cells. Evidence on the participation of the VDR and Src JOURNAL OF STEROID BIOCHEMISTRY AND MOLECULAR BIOLOGY 89-90 (1-5): 287-290 Sp. Iss. SI MAY 2004 Medullary thyroid cancer: Liu ZJ, Falola J, Zhu XD, et al. Antiproliferative effects of Src inhibition on medullary thyroid cancer JOURNAL OF CLINICAL ENDOCRINOLOGY AND METABOLISM 89 (7): 3503-3509 JUL 2004 Colon cancer: Thamilselvan V, Patel A, van Zyp JV, et al. Colon cancer cell adhesion in response to Src kinase activation and actin-cytoskeleton by non-laminar shear stress JOURNAL OF CELLULAR BIOCHEMISTRY 92 (2): 361-371 MAY 15 2004 Prostate cancer: Paronetto MP, Farini D, Sammarco I, et al. Expression of a truncated form of the c-kit tyrosine kinase receptor and activation of Src kinase in human prostatic cancer AMERICAN JOURNAL OF PATHOLOGY 164 (4): 1243-1251 APR 2004
LCK	8569365	-17.90	Yes	Coxsackievirus B3-mediated heart disease: Liu P, Aitken K, Kong YY, et al. The tyrosine kinase p56(lck) is essential in coxsackievirus B3-mediated heart disease NATURE MEDICINE 6 (4): 429-434 APR 2000 Small cell lung cancer: Krystal GW, Litz J Activation of LCK by Kit is necessary for SCF-mediated growth of small cell lung cancer. BLOOD 90 (10): 1375-1375 Part 1 Suppl. 1 NOV 15 1997 Colorectal cancer: Mayer K, Ballhausen WG Expression of alternatively spliced lck transcripts from the proximal promoter in colorectal cancer derived cell lines ANTICANCER RESEARCH 16 (4A): 1733-1737 JUL-AUG 1996
LYN	125480	-17.90	Yes	Prostate cancer: Goldenberg-Furmanov M, Stein I, Pikarsky E, et al. Lyn is a target gene for prostate cancer: Sequence-based inhibition induces regression of human tumor xenografts CANCER RESEARCH 64 (3): 1058-1066 FEB 1 2004 Erythroleukemia : Zarn JA, Zimmermann SM, Pass MK, et al. Association of CD24 with the kinase c-fgr in a small cell lung cancer cell line and with the kinase lyn in an erythroleukemia cell line BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 225 (2): 384-391 AUG 14 1996

PDGF R α	5453870	-17.07	Yes	Metastatic cancer : Tsutsumimoto T, Yi B, Williams P, et al. Paracrine activation of PDGFR tyrosine kinase in Osteoblasts is critical to osteosclerotic bone Metastases in breast cancer producing PDGF-BB. JOURNAL OF BONE AND MINERAL RESEARCH 18: S198-S198 Suppl. 2 SEP 2003 Breast cancer: Kleer CG, Shen R, Wolf J, et al. Characterization of platelet derived growth factor receptor (PDGFR) expression in breast cancer identifies inflammatory breast cancer as a potential target for treatment with PDGFR inhibitors MODERN PATHOLOGY 16 (1): 153 JAN 2003 Myeloproliferative disease: Cools J, Stover EH, Boulton CL, et al. PKC412 overcomes resistance to imatinib in a murine model of FIP1L1-PDGFR alpha-induced myeloproliferative disease CANCER CELL 3 (5): 459-469 MAY 2003 Medulloblastoma: MacDonald TJ, Brown KM, LaFleur B, et al. Expression profiling of medulloblastoma: PDGFRA and the RAS/MAPK pathway as therapeutic targets for metastatic disease NATURE GENETICS 29 (2): 143-152 OCT 2001
IKKa	4502843	-17.04	Yes	Lung cancer: Sanlioglu S, Luleci G, Thomas KW Simultaneous inhibition of Rac1 and IKK pathways sensitizes lung cancer cells to TNF alpha-mediated apoptosis CANCER GENE THERAPY 8 (11): 897-905 NOV 2001
PDGF Rb	4505683	-16.97	YES	Myeloproliferative disease: Cain JA, Grisolan JL, Laird AD, et al. Complete remission of TEL-PDGFRB-induced myeloproliferative disease in mice by receptor tyrosine kinase inhibitor SU11657 BLOOD 104 (2): 561-564 JUL 15 2004 Breast cancer: Kleer CG, Shen R, Wolf J, et al. Characterization of platelet derived growth factor receptor (PDGFR) expression in breast cancer identifies inflammatory breast cancer as a potential target for treatment with PDGFR inhibitors MODERN PATHOLOGY 16 (1): 153 JAN 2003

KIT	18117734	-16.78	No	Small cell lung cancer: Rohr UP, Rehfeld N, Pflugfelder L, et al. Expression of the tyrosine kinase C-kit is an independent prognostic factor in patients with small cell lung cancer INTERNATIONAL JOURNAL OF CANCER 111 (2): 259-263 AUG 20 2004 Systemic mast cell disease associated with eosinophilia: Cervantes F, Camos M, Boque C, et al. FIP1L1-PDGFRA and c-kit D816V mutation-based clonality studies in systemic mast cell disease associated with eosinophilia HAEMATOLOGICA 89 (7): 871-873 JUL 2004 Prostate cancer: Paronetto MP, Farini D, Sammarco I, et al. Expression of a truncated form of the c-kit tyrosine kinase receptor and activation of Src kinase in human prostatic cancer AMERICAN JOURNAL OF PATHOLOGY 164 (4): 1243-1251 APR 2004 Gastrointestinal stromal tumors: de Silva MV, Reid R Gastrointestinal stromal tumors (GIST): C-kit mutations, CD117 expression, differential diagnosis and targeted cancer therapy with imatinib PATHOLOGY & ONCOLOGY RESEARCH 9 (1): 13-19 2003
FGFR 2	533220	-16.66	No	Breast cancer: Koziczak M, Holbro T, Hynes NE Blocking of FGFR signaling inhibits breast cancer cell proliferation through downregulation of D-type cyclins ONCOGENE 23 (20): 3501-3508 APR 29 2004 Uterine cervical cancer: Kurban G, Ishiwata T, Kudo M, et al. Expression of keratinocyte growth factor receptor (KGFR/FGFR2 IIIb) in human uterine cervical cancer ONCOLOGY REPORTS 11 (5): 987-991 MAY 2004 Non-small cell lung cancer: Beau-Faller M, Gaub MP, Schneider A, et al. Allelic imbalance at loci containing FGFR, FGF, c-Met and HGF candidate genes in non-small cell lung cancer subtypes, implication for progression EUROPEAN JOURNAL OF CANCER 39 (17): 2538-2547 NOV 2003 Gastric cancer: Ferrin LJ, Lee HC Structure of the fibroblast growth factor receptor 2 (FGFR2) amplicon in a gastric cancer cell line. GASTROENTEROLOGY 118 (4): 2818 Part 1 Suppl. 2 APR 2000

HER4/ ErbB4	422854	-16.10	Yes	Transitional cell carcinoma of the bladder: Junttila TT, Laato M, Vahlberg T, et al. Identification of patients with transitional cell carcinoma of the bladder overexpressing ErbB2, ErbB3, or specific ErbB4 isoforms: Real-time reverse transcription-PCR analysis in estimation of ErbB receptor status from cancer patients CLINICAL CANCER RESEARCH 9 (14): 5346-5357 NOV 1 2003 Alzheimer's disease: Chaudhury AR, Gerecke KM, Wyss JM, et al. Neuregulin-1 and ErbB4 immunoreactivity is associated with neuritic plaques in Alzheimer disease brain and in a transgenic model of Alzheimer disease JOURNAL OF NEUROPATHOLOGY AND EXPERIMENTAL NEUROLOGY 62 (1): 42-54 JAN 2003 Anti-glomerular basement membrane disease: Levidiotis V, Khong TF, Katerelos M, et al. Increased expression of HB-EGF and its receptor erbB4/HER4 in accelerated anti-GBM disease NEPHROLOGY 7 (5): 233-238 OCT 2002 Breast cancer: Sartor CI, Zhou H, Kozłowska E, et al. HER4 mediates ligand-dependent antiproliferative and differentiation responses in human breast cancer cells MOLECULAR AND CELLULAR BIOLOGY 21 (13): 4265-4275 JUL 2001 Non-melanoma skin cancer: Krahn G, Leiter U, Kaskel P, et al. Coexpression patterns of EGFR, HER2, HER3 and HER4 in non-melanoma skin cancer EUROPEAN JOURNAL OF CANCER 37 (2): 251-259 JAN 2001
FYN	6978863	-15.96	Yes	Alzheimer's disease: Lee G, Thangavel R, Sharma VM, et al. Phosphorylation of tau by fyn: Implications for Alzheimer's disease JOURNAL OF NEUROSCIENCE 24 (9): 2304-2312 MAR 3 2004 Oral cancer: Li XW, Yang YJ, Hu YM, et al. alpha(nu)beta(6)-Fyn signaling promotes oral cancer progression JOURNAL OF BIOLOGICAL CHEMISTRY 278 (43): 41646-41653 OCT 24 2003: Fibrosarcoma: TAKAYAMA T, MOGI Y, KOGAWA K, et al. A ROLE FOR THE FYN ONCOGENE IN METASTASIS OF METHYLCHOLANTHRENE-INDUCED FIBROSARCOMA-A CELLS INTERNATIONAL JOURNAL OF CANCER 54 (5): 875-879 JUL 9 1993

FLT1	4503749	-15.92	Yes	Breast cancer: Caine GJ, Lip GY, Blann AD Platelet-derived VEGF, Flt-1, angiopoietin-1 and P-selectin in breast and prostate cancer: further evidence for a role of platelets in tumour angiogenesis ANNALS OF MEDICINE 36 (4): 273-277 2004 Prostate cancer: Oyama N, Ponde DE, Higashikubo R, et al. In vitro and in vivo assessment of 18F-FLT as a proliferation marker in prostate cancer tumor model JOURNAL OF UROLOGY 171 (4): 291-291 Suppl. S APR 2004 Diffuse large B cell lymphoma: Aref S, Mabed M, Zalata K, et al. The interplay between C-Myc oncogene expression and circulating vascular endothelial growth factor (sVEGF), its antagonist receptor, soluble Flt-1 in diffuse large B cell lymphoma (DLBCL): Relationship to patient outcome (vol 45, pg 499, 2004) LEUKEMIA & LYMPHOMA 45 (6): 1311-1311 JUN 2004
SYK	2136036	-15.89	Yes	Breast cancer: Zhang XY, Shrikande U, Zhou Q, et al. Localization of Syk in breast cancer cells and a role for the tyrosine kinase in cell-cell adhesion FASEB JOURNAL 18 (8): C107-C107 Suppl. S MAY 14 2004
FGFR 4	120051	-15.71	Yes	Breast cancer: Jezequel P, Campion L, Joalland MP, et al. G388R mutation of the FGFR4 gene is not relevant to breast cancer prognosis BRITISH JOURNAL OF CANCER 90 (1): 189-193 JAN 1 2004
FAK	3183518	-15.70	Yes	Ovarian cancer: Sood AK, Coffin JE, Schneider GB, et al. Role of FAK in ovarian cancer invasion and migration. JOURNAL OF THE SOCIETY FOR GYNECOLOGIC INVESTIGATION 11 (2): 181A-181A 323 Suppl. S FEB 2004 Pancreatic cancer: Ito H, Gardner-Thorpe J, Duxbury MS, et al. Focal adhesion kinase (FAK) gene silencing reduces pancreatic cancer MMP activity and invasiveness. JOURNAL OF THE AMERICAN COLLEGE OF SURGEONS 197 (3): S76-S76 Suppl. S SEP 2003 Colon cancer: Walsh MF, Thamilselvan V, Grotelueschen R, et al. Absence of adhesion triggers differential FAK and SAPKp38 signals in SW620 human colon cancer cells that may inhibit adhesiveness and lead to cell death CELLULAR PHYSIOLOGY AND BIOCHEMISTRY 13 (3): 135-146 2003
FLT4	388522	-15.63	Yes	Thyroid disease: Karamysheva A, Shushanov S, Bronstein M, et al. Vascular endothelial growth factor c (VEGFc) and its receptor FLT4 expression in human thyroid pathology. JOURNAL OF INVESTIGATIVE MEDICINE 47 (4): 195A-195A Suppl. S APR 1999

112. It is understood that each of the diseases listed in Table 10 is a disease for which imatinib and its derivatives can be used to treat. Thus, subjects having these diseases would be candidates for treatment with imatinib and its derivatives or purvalanol or SB 203580 or their derivatives depending on which receptor is target by which drug. Methods of treatment comprising administering imatinib or a derivative to treat these diseases alone or in combination with other treatments for these diseases, such as radiation, surgical, or other chemotherapy protocols are also disclosed.

C. Computer Systems and Methods

1. Systems

113. It is understood that the methods disclosed herein are useful with computer systems for implementing the steps described herein. For example, disclosed herein is a computer system having a memory means, a processing means, a data input means, and a visual display means, the memory means containing sequence information for a known target capable of interacting with a molecule, such as a drug, and modules containing information to be compared with the sequence information of the known target, and the processing means being operable to compute molecule-potential target binding energy using the methods of identifying a target disclosed herein, and display the structures of molecules based on input atomic structure information with a visual display means.

114. Also disclosed herein, and illustrated in FIG. 13, are methods of displaying a representation of a potential target-molecule interaction on a computer having a processing means, a memory means, a visual display means, an input means and an output means comprising: receiving the three-dimensional coordinates of atoms of the potential target (1301); producing a representation of the potential target based upon the received coordinates (1302); and displaying the representation of the potential target-molecule interaction on the visual display means, wherein the potential target in the potential target-molecule interaction comprises a potential target which is a homologue of the known target of the molecule (1303).

115. Also disclosed herein is an apparatus comprising: (a) a system data store capable of storing coordinate sets; and (b) a system processor in communication with the system data store that carries out the following steps: (i) modeling a molecule in complex with a known target for the molecule, (ii) obtaining potential target molecules by

selecting potential target molecules with a defined homology to the known target, and (iii) determining the binding affinity of a potential target with the molecule by modeling the potential target with the molecule, wherein a Monte Carlo function is used for sampling of side chain rotamers.

5 116. It is also understood that the proteins disclosed herein can be represented as a sequence consisting of the nucleotides of amino acids, or as the amino acids themselves. There are a variety of ways to display these sequences, for example the nucleotide guanosine can be represented by G or g. Likewise the amino acid valine can be represented by Val or V. Those of skill in the art understand how to display and
10 express any nucleic acid or protein sequence in any of the variety of ways that exist, each of which is considered herein disclosed. Specifically contemplated herein is the display of these sequences on computer readable mediums, such as, commercially available floppy disks, tapes, chips, hard drives, compact disks, and video disks, or other computer readable mediums. Also disclosed are the binary code representations of the disclosed
15 sequences. Those of skill in the art understand what is meant by computer readable media. Thus, computer readable mediums on which the nucleic acids or protein sequences are recorded, stored, or saved.

a) Machine Readable Storage Media

20 117. Disclosed are machine-readable storage mediums, also referred to as computer readable media, comprising a data storage material encoded with machine readable data. Furthermore, the data can be extracted and manipulated by machines configured to read the data stored on the machine readable storage media, and in fact, when performing the molecular modeling, such as displaying a configuration of the disclosed compositions, as discussed herein, typically the data will be retrieved or stored
25 on a machine readable storage media.

 118. Disclosed are machine readable storage media comprising the coordinates set forth herein or obtained or coordinates producing equivalent configurations of the disclosed compositions or their variants as discussed herein.

30 119. For example, a system for reading a data storage medium may include a computer comprising a central processing unit ("CPU"), a working memory which may be, e.g., RAM (random access memory) or "core" memory, mass storage memory (such as one or more disk drives or CD-ROM drives), one or more display devices (e.g., cathode-ray tube ("CRT") displays, light emitting diode ("LED") displays, liquid crystal

displays ("LCDs"), electroluminescent displays, vacuum fluorescent displays, field emission displays ("FEDs"), plasma displays, projection panels, etc.), one or more user input devices (e.g., keyboards, microphones, mice, touch screens, etc.), one or more input lines, and one or more output lines, all of which are interconnected by a conventional
5 bidirectional system bus. The system may be a stand-alone computer, or may be networked (e.g., through local area networks, wide area networks, intranets, extranets, or the internet) to other systems (e.g., computers, hosts, servers, etc.). The system may also include additional computer controlled devices such as consumer electronics and appliances. Input hardware may be coupled to the computer by input lines and may be
10 implemented in a variety of ways. Machine-readable data of this invention may be inputted via the use of a modem or modems connected by a telephone line or dedicated data line. Alternatively or additionally, the input hardware may comprise CD-ROM drives or disk drives. In conjunction with a display terminal, a keyboard may also be used as an input device. Output hardware may be coupled to the computer by output lines and
15 may similarly be implemented by conventional devices. By way of example, the output hardware may include a display device for displaying a graphical representation of a binding pocket of this invention using a program such as QUANTA as described herein. Output hardware might also include a printer, so that hard copy output may be produced, or a disk drive, to store system output for later use.

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

30 121. Machine-readable storage devices useful in the present invention include, but are not limited to, magnetic devices, electrical devices, optical devices, and combinations thereof. Examples of such data storage devices include, but are not limited to, hard disk devices, CD devices, digital video disk devices, floppy disk devices, removable hard disk devices, magneto-optic disk devices, magnetic tape devices, flash memory devices, bubble memory devices, holographic storage devices, and any other mass storage peripheral device. It should be understood that these storage devices include

necessary hardware (e.g., drives, controllers, power supplies, etc.) as well as any necessary media (e.g., disks, flash cards, etc.) to enable the storage of data.

2. Structures

122. The disclosed methods can be performed on computers and molecular structures are displayed and created.

123. Also disclosed are scalable three dimensional configurations of points derived from structure coordinates of at least a portion of the molecules used herein. In one embodiment, the scalable three dimensional set of points is derived from structure coordinates of a model.

124. Also disclosed are scalable three dimensional set of points derived from structure coordinates of at least a portion of a molecule or a molecular complex that is structurally homologous to a disclosed composition.

125. Also disclosed are molecules or molecular complexes and their cognate coordinates that are structurally homologous to a disclosed composition.

126. Also disclosed are methods involving molecular replacement, substitution, deletion, or alteration to obtain structural information about a molecule or molecular complex of unknown structure, but which is related to the disclosed structures, through for example, amino acid identity. The methods include producing a solution of the molecule or molecular complex, generating a solution structure aided by the information disclosed herein, and applying at least a portion of the structure coordinates obtained to the data related to the molecule or molecule complex to generate a three-dimensional structure of at least a portion of the molecule or molecular complex.

127. Also disclosed are methods for homology modeling.

128. Each of the constituent amino acids of a protein can be defined by a set of structure coordinates. The term "structure coordinates" refers to a Cartesian set of coordinates.

129. Disclosed are representation of variations in structure coordinates which can be generated by mathematically manipulating the disclosed structure coordinates. For example, the structure coordinates obtained for a given protein could be manipulated by permutations of the structure coordinates, fractionalization of the structure coordinates, integer additions or subtractions to sets of the structure coordinates, inversion of the structure coordinates, rotation of the structure coordinates about an arbitrary axis, or any combination of the above. Alternatively, modifications in the crystal

structure due to mutations, additions, substitutions, and/or deletions of amino acids, or other changes in any of the components that make up the composition from which the coordinates were produced, could also yield variations in structure coordinates. Such variations in the individual coordinates will have little effect on the global shape.

5 Furthermore, when variations are made in a concentrated region of the composition or in the structural representation of the composition, the effect on other structural regions of the molecule typically is minimal. One way of judging the effect the variations in one part of the composition or structure have on another part of the composition or structure is to compare the cognate regions to the disclosed structures standard error and judge
10 whether the differences are in an acceptable range, for example, within the same error range for that region in the original structure. If the error is within the disclosed error ranges, the structures or compared regions of the structure can be said to be equivalent. The alterations and modifications discussed herein, in connection with the discussion of protein modifications and discussed herein indicate that modifications which will not
15 alter, for example the properties of the disclosed compositions, such as binding between a molecule and a target, can be made and are disclosed.

a) Coordinates

130. Structure coordinates define a unique configuration of points in space. Those of skill in the art understand that a set of structure coordinates for protein or an
20 protein/ligand complex, or a portion thereof, define a relative set of points that, in turn, define a configuration in three dimensions. A key piece of information obtained from the coordinates is the position of the atoms that make up the composition. The position of the atoms is defined in a Cartesian form, such that there are x-y-z positions which allow for a determination of distances and angles between two or more atoms. Thus, a similar
25 or identical configuration, i.e. structure, can be defined by an entirely different set of coordinates, provided the distances and angles between coordinates remain essentially the same. By manipulating the distances and angles in a like manner a scalable representation can be obtained.

131. Disclosed are scalable three-dimensional configurations derived from
30 structure coordinates obtained for the proteins and molecules discussed herein, or portion thereof, or from coordinates producing a configuration with essentially the same angles and distances between the atoms. Also disclosed are scalable three-dimensional configurations derived from the structure coordinates obtained from the protein structure

database, such as the RCSB protein databank found at <http://www.rcsb.org/pdb>, and the NCBI structure database found at <http://www.ncbi.nlm.nih.gov/Structure/>. It is understood that in certain situations, the structures and information needed to produce these structures disclosed in these databases are incorporated by reference for material related to the structures of proteins and protein complexes for the coordinate material. In certain situations this incorporation is only for the material present in these databases as of the time of filing of this application.

132. Also disclosed are scalable three-dimensional configurations of points derived from structure coordinates of molecules or molecular complexes that are structurally homologous to the disclosed proteins, as well as structurally equivalent configurations.

133. The configurations of points in space derived from structure coordinates according to the invention can be visualized as, for example, a holographic image, a stereodiagram, a model or a computer-displayed image, and the invention thus includes such images, diagrams or models.

134. Comparisons between different structures, different conformations of the same structure, and different parts of the same structure can be performed in a variety of ways. For example, typically the structures (coordinates making up the structure) are loaded, the atom equivalences in these structures are defined; the structures are fit, and then the resulting comparisons are reviewed.

135. Modeling programs typically also allow for a determination of the variances, the root mean square deviations, and statistical significance of the various structures.

136. The term "root mean square deviation" means the square root of the arithmetic mean of the squares of the deviations. This allows for comparison of two sets of data for example or the cognate position in two configurations or structures.

3. Modeling and modeling of variants

137. Computational techniques can be used to screen, identify, select and design chemical entities capable of associating with the identified targets or molecules or structurally homologous targets or molecules. The disclosed coordinates and those that produce similar homologous structures, i.e. having RMS deviations of less than or equal to 5, 4, 3, 2, or 1 angstroms can be used to model potential molecule-target interactions. Atoms of the potential ligand can be included in modeling simulation involving the

known target or identified target and or molecule complex as disclosed herein, and the contacts that arise between the potential ligand in a variety of positions with the targets or with a region, such as the molecule binding site, can be investigated. Energy minimization of these contacts between the potential ligand and the molecule can indicate potential ligands having, for example a desired affinity or a desired specificity. FIG. 16 shows superimposed views of SB203580 in a model of the p38 α binding site before and after GROMACS energy minimization. Structures prior to minimization are shown in red, those after minimization are shown in green. The kinase sidechains have been removed for clarity. The ligands identified as having a desired number of contacts, with atoms of the target as positioned by the coordinates or homologs disclosed herein, can be chosen and then optionally further tested by synthesizing or making the ligand and target and performing standard biochemistry to assay binding activity or functional activity, such as those that use kinetic or thermodynamic methodology, such as, equilibrium dialysis, microcalorimetry, circular dichroism, capillary zone electrophoresis, nuclear magnetic resonance spectroscopy, fluorescence spectroscopy, and combinations thereof.

138. Drug designing typically involves computer-assisted design of chemical entities that associate with a target, its homologs, or portions thereof. Chemical entities can be designed in a step-wise fashion, one fragment at a time, or may be designed as a whole or "de novo."

139. The binding sites of targets and molecules as disclosed herein set forth the position of target atoms for interaction with ligands which will be able to bind or inhibit the interaction. The conformation of the binding site allows for a precise three dimensional map for rationally designing molecules that will form, for example, a set number of contacts with the atoms defining the binding regions as disclosed herein.

140. A contact as used herein means any position between two atoms, typically one atom of a molecule, such as a ligand, and one atom of the target, such as a receptor, that when position by an energy minimization program, for example, are less than 5 Å°, 4 Å°, 3 Å°, 2 Å°, or 1 Å° apart. Thus, a contact can for example, correlate with, for example, non-covalent interactions, such as a hydrogen bonds, Vander Walls interactions, hydrophobic interactions, and electrostatic interactions, between two atoms. Typically a contact will add to the binding energy between two atoms, but it can also be repulsive, typically more repulsive the closer the two atoms become. It is understood that

for a ligand to be a potential therapeutic candidate, it must have an appropriate level or quality of contacts, such that an interaction occurs, but that it should not cause steric and energetic problems. Conformational considerations include the overall three-dimensional *structure* and orientation of the chemical entity in relation to the binding pocket, and the spacing between various functional groups of an entity that directly interact with the binding pocket or homologs thereof.

141. The modeling and display of the disclosed compositions can be accomplished using any modeling program, such as QUANTA, SYBYL, Insight II/Discover (Molecular Simulations, Inc., San Diego, Calif. 92121). These programs may be implemented, for example, using a Silicon Graphics workstation such as an Indigo² with "IMPACT" graphics. Other hardware systems and software packages will be known to those skilled in the art. Drug design programs, such as, GRID (P. J. Goodford, J. Med. Chem. 28:849-857 (1985); available from Oxford University, Oxford, UK); MCSS (A. Miranker et al., Proteins: Struct. Funct. Gen., 11:29-34 (1991); available from Molecular Simulations, San Diego, Calif.); AUTODOCK (D. S. Goodsell et al., Proteins: Struct. Funct. Genet. 8:195-202 (1990); available from Scripps Research Institute, La Jolla, Calif.); and DOCK (I. D. Kuntz et al., J. Mol. Biol. 161:269-288 (1982); available from University of California, San Francisco, Calif.), LUDI (H.-J. Bohm, J. Comp. Aid. Molec. Design. 6:61-78 (1992); available from Molecular Simulations Inc., San Diego, Calif.); LEGEND (Y. Nishibata et al., Tetrahedron, 47:8985 (1991); available from Molecular Simulations Inc., San Diego, Calif.); LeapFrog (available from Tripos Associates, St. Louis, Mo.); and SPROUT (V. Gillet et al., J. Comput. Aided Mol. Design 7:127-153 (1993); available from the University of Leeds, UK), can also be used.

142. The efficiency of a potential ligand's interaction with a target can be evaluated and optimized. For example, typically a preferred ligand will cause little perturbation to the three dimensional positioning of the atoms of target that are in the vicinity of the interaction or are somehow allosterically affected. The level of perturbation can be determined by comparing the energy state of the structural conformation for the bound and unbound states. Typically the smaller the change the less perturbation and the less perturbation the higher the likelihood that the ligand will be desirable as for example, a competitive inhibitor. This perturbation energy can be, for example, less than or equal to about 30 kcal/mole, 20 kcal/mole, 15 kcal/mole, 10 kcal/mole, 8 kcal/mole, 6 kcal/mole, 5 kcal/mole, 4 kcal/mole, 3 kcal/mole, 2 kcal/mole,

or 1 kcal.mole. Ligands may interact with the target molecule in more than one conformation that is similar in overall binding energy. In those cases, the perturbation energy of binding can be taken as the difference between the energy of the free entity and the average energy of the conformations observed when the ligand binds to the target molecule.

143. An entity designed or selected as binding to a target may be further computationally optimized so that in its bound state it would preferably lack repulsive electrostatic interaction with the target enzyme and with the surrounding water molecules. Such non-complementary electrostatic interactions include repulsive charge-charge, dipole-dipole, and charge-dipole interactions.

144. Specific computer software is available in the art to evaluate compound deformation energy and electrostatic interactions. Examples of programs designed for such uses include: Gaussian 94, revision C (M. J. Frisch, Gaussian, Inc., Pittsburgh, Pa. 15106); AMBER, version 4.1 (P. A. Kollman, University of California at San Francisco, 94143); QUANTA/CHARMM (Molecular Simulations, Inc., San Diego, Calif. 92121);

145. The disclosed structures and coordinates can also be used to screen potential ligands, for example, as drug candidates, which interact with, i.e. form contacts with, the identified target. Small molecule databases, such as structure databases can be used for this. Not only whole molecules can be screened, but subparts of molecule, for example, various functional groups can also be screen to find preferred functional groups for forming contacts with the identified target structures disclosed herein. Functional groups that make a desired set of contacts, for example, with a desired or particular region of the target molecule, can then be used to further build combinations of these and other types of functional groups to design ligands containing the functional groups or combinations of functional groups.

146. It is understood that also disclosed are iterative approaches which use successive performance of the various steps disclosed herein to optimize molecules and/or isolate molecules from sets of molecules. This can also be done with multiple coordinate sets that have been obtained, for example, from the solution of structures involving a ligand or series of structures involving a series of ligands. For example, molecules known to have preferred biochemical properties can be solved in a co-structure, and then the structure information obtained from this can be used to select potential ligands for function.

147. A compound that is identified or designed as a result of any of these methods disclosed herein can be obtained (or synthesized) and tested for its biological activity, e.g., inhibition of target activity or enhancement of target activity.

148. Structures of variant molecules or proteins can be produced without obtaining individual coordinates for the variant. In essence the coordinates of the molecule or protein, disclosed herein or coordinates that produce a similar structure are used as a starting point and the variant atom or atoms of the variant molecule or protein are substituted into the simulated structure and their relative position to the original unchanging atoms, i.e. coordinates, are determined through any of a variety of energy minimization functions. Thus, sequence alignment, secondary *structure* prediction, the screening of structural libraries of the disclosed molecules and proteins produced from the disclosed coordinates, or any combination of these can be used to overlay the variant structure. For example, the variant atom or atoms can also be modeled from any structural library having coordinates of similar or identical atoms. Thus, the initial structure to undergo energy minimization can be arrived at by modeling known coordinates for a given for the given atom or atoms. These libraries of structures can be screened for the optimal structure. A side chain rotamer library can be used to model a given side chain or set of side chains. After initial energy minimization iterative or new energy minimizations may be necessary if the structure produced after energy minimization violates a physical constraint, such as correct stereochemistry.

D. Compositions

149. Disclosed herein are compositions to be used with the methods disclosed herein, such as proteins and nucleic acids encoding the proteins, as well as molecules such as drugs. These and other materials are disclosed herein, and it is understood that when combinations, subsets, interactions, groups, etc. of these materials are disclosed that while specific reference of each various individual and collective combinations and permutation of these compounds may not be explicitly disclosed, each is specifically contemplated and described herein. For example, if a particular known target is disclosed and discussed and a number of modifications that can be made to a number of molecules including the amino acids are discussed, specifically contemplated is each and every combination and permutation of amino acids, and the modifications that are possible unless specifically indicated to the contrary. Thus, if a class of molecules A, B, and C are disclosed as well as a class of molecules D, E, and F and an example of a combination

molecule, A-D is disclosed, then even if each is not individually recited each is individually and collectively contemplated meaning combinations, A-E, A-F, B-D, B-E, B-F, C-D, C-E, and C-F are considered disclosed. Likewise, any subset or combination of these is also disclosed. Thus, for example, the sub-group of A-E, B-F, and C-E would be considered disclosed. This concept applies to all aspects of this application including, but not limited to, steps in methods of making and using the disclosed compositions. Thus, if there are a variety of additional steps that can be performed it is understood that each of these additional steps can be performed with any specific embodiment or combination of embodiments of the disclosed methods.

1. Homology/identity

150. It is understood that one way to define any known variants and derivatives of the target, or those that might arise, to be used with the methods disclosed herein, is through defining the variants and derivatives in terms of homology to specific known sequences. For example a known target such as a protein has a particular sequence, and there is a particular nucleic acid corresponding to that particular sequence. Those of skill in the art readily understand how to determine the homology of two or more proteins or nucleic acids, such as genes. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

151. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. Mol Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A. 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

152. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 which are herein incorporated by reference for at least material related to nucleic acid alignment.

2. Sequence similarities

153. It is understood that as discussed herein the use of the terms homology and identity mean the same thing as similarity. Thus, for example, if the use of the word homology is used between two non-natural sequences it is understood that this is not necessarily indicating an evolutionary relationship between these two sequences, but rather is looking at the similarity or relatedness between their nucleic acid sequences. Many of the methods for determining homology between two evolutionarily related molecules are routinely applied to any two or more nucleic acids or proteins for the purpose of measuring sequence similarity regardless of whether they are evolutionarily related or not.

154. In general, it is understood that one way to define any known variants and derivatives or those that might arise, of the disclosed genes and proteins herein, is through defining the variants and derivatives in terms of homology to specific known sequences. This identity of particular sequences disclosed herein is also discussed elsewhere herein. In general, variants of genes and proteins herein disclosed typically have at least, about 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, or 99 percent homology to the stated sequence or the native sequence. Those of skill in the art readily understand how to determine the homology of two proteins or nucleic acids, such as genes. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

155. Another way of calculating homology can be performed by published algorithms. Optimal alignment of sequences for comparison may be conducted by the local homology algorithm of Smith and Waterman Adv. Appl. Math. 2: 482 (1981), by the homology alignment algorithm of Needleman and Wunsch, J. Mol Biol. 48: 443 (1970), by the search for similarity method of Pearson and Lipman, Proc. Natl. Acad. Sci. U.S.A. 85: 2444 (1988), by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI), or by inspection.

156. For example, as used herein, a sequence recited as having a particular percent homology to another sequence refers to sequences that have the recited homology as calculated by any one or more of the calculation methods described above. For

example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using the Zuker calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by any of the other calculation methods. As another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using both the Zuker calculation method and the Pearson and Lipman calculation method even if the first sequence does not have 80 percent homology to the second sequence as calculated by the Smith and Waterman calculation method, the Needleman and Wunsch calculation method, the Jaeger calculation methods, or any of the other calculation methods. As yet another example, a first sequence has 80 percent homology, as defined herein, to a second sequence if the first sequence is calculated to have 80 percent homology to the second sequence using each of calculation methods (although, in practice, the different calculation methods will often result in different calculated homology percentages).

a) Sequences

157. There are a variety of sequences related to, for example, the kinase receptors described herein, as well as any other protein disclosed herein that are disclosed on Genbank, and these sequences and others are herein incorporated by reference in their entireties as well as for individual subsequences contained therein.

158. A variety of sequences are provided herein and these and others can be found in Genbank, at www.pubmed.gov. Those of skill in the art understand how to resolve sequence discrepancies and differences and to adjust the compositions and methods relating to a particular sequence to other related sequences. Primers and/or probes can be designed for any sequence given the information disclosed herein and known in the art.

3. Peptides

a) Protein variants

159. As discussed herein there are numerous variants of a known target, or a potential target (such as a protein), that are known and herein contemplated. Also disclosed are specific receptors whose sequences are known in the art. In addition to the known functional variants there are derivatives of the proteins which also function in the disclosed methods and compositions. Protein variants and derivatives are well

understood to those of skill in the art and can involve amino acid sequence modifications. For example, amino acid sequence modifications typically fall into one or more of three classes: substitutional, insertional or deletional variants. Insertions include amino and/or carboxyl terminal fusions as well as intrasequence insertions of single or multiple amino acid residues. Insertions ordinarily will be smaller insertions than those of amino or carboxyl terminal fusions, for example, on the order of one to four residues.

Immunogenic fusion protein derivatives, such as those described in the examples, are made by fusing a polypeptide sufficiently large to confer immunogenicity to the target sequence by cross-linking in vitro or by recombinant cell culture transformed with DNA encoding the fusion. Deletions are characterized by the removal of one or more amino acid residues from the protein sequence. Typically, no more than about from 2 to 6 residues are deleted at any one site within the protein molecule. These variants ordinarily are prepared by site specific mutagenesis of nucleotides in the DNA encoding the protein, thereby producing DNA encoding the variant, and thereafter expressing the DNA in recombinant cell culture. Techniques for making substitution mutations at predetermined sites in DNA having a known sequence are well known, for example M13 primer mutagenesis and PCR mutagenesis. Amino acid substitutions are typically of single residues, but can occur at a number of different locations at once; insertions usually will be on the order of about from 1 to 10 amino acid residues; and deletions will range about from 1 to 30 residues. Deletions or insertions preferably are made in adjacent pairs, i.e. a deletion of 2 residues or insertion of 2 residues. Substitutions, deletions, insertions or any combination thereof may be combined to arrive at a final construct. The mutations must not place the sequence out of reading frame and preferably will not create complementary regions that could produce secondary mRNA structure. Substitutional variants are those in which at least one residue has been removed and a different residue inserted in its place. Such substitutions generally are made in accordance with the following Tables 1 and 2 and are referred to as conservative substitutions.

160. TABLE 1: Amino Acid Abbreviations

Amino Acid	Abbreviations
alanine	AlaA
alloseucine	Alle
arginine	ArgR
asparagine	AsnN
aspartic acid	AspD
cysteine	CysC
glutamic acid	GluE
glutamine	GlnK

Amino Acid	Abbreviations
glycine	GlyG
histidine	HisH
isoleucine	IleI
leucine	LeuL
lysine	LysK
phenylalanine	PheF
proline	ProP
pyroglutamic acidp	Glu
serine	SerS
threonine	ThrT
tyrosine	TyrY
tryptophan	TrpW
valine	ValV

TABLE 2: Amino Acid Substitutions	
Original Residue Exemplary Conservative Substitutions, others are known in the art.	
Ala	ser
Arg	lys, gln
Asn	gln; his
Asp	glu
Cys	ser
Gln	asn, lys
Glu	asp
Gly	pro
His	asn; gln
Ile	leu; val
Leu	ile; val
Lys	arg; gln;
Met	Leu; ile
Phe	met; leu; tyr
Ser	thr
Thr	ser
Trp	tyr
Tyr	trp; phe
Val	ile; leu

161. Substantial changes in function or immunological identity are made by selecting substitutions that are less conservative than those in Table 2, i.e., selecting residues that differ more significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site or (c) the bulk of the side chain. The substitutions which in general are expected to produce the greatest changes in the protein properties will be those in which (a) a hydrophilic residue, e.g. seryl or threonyl, is substituted for (or by) a hydrophobic residue, e.g. leucyl, isoleucyl, phenylalanyl, valyl or alanyl; (b) a cysteine or proline is substituted for (or by) any other residue; (c) a residue having an electropositive side chain, e.g., lysyl, arginyl, or histidyl, is substituted for (or by) an electronegative residue, e.g., glutamyl or aspartyl; or

(d) a residue having a bulky side chain, e.g., phenylalanine, is substituted for (or by) one not having a side chain, e.g., glycine, in this case, (e) by increasing the number of sites for sulfation and/or glycosylation.

162. For example, the replacement of one amino acid residue with another that is biologically and/or chemically similar is known to those skilled in the art as a conservative substitution. For example, a conservative substitution would be replacing one hydrophobic residue for another, or one polar residue for another. The substitutions include combinations such as, for example, Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; and Phe, Tyr. Such conservatively substituted variations of each explicitly disclosed sequence are included within the mosaic polypeptides provided herein.

163. Substitutional or deletional mutagenesis can be employed to insert sites for N-glycosylation (Asn-X-Thr/Ser) or O-glycosylation (Ser or Thr). Deletions of cysteine or other labile residues also may be desirable. Deletions or substitutions of potential proteolysis sites, e.g. Arg, is accomplished for example by deleting one of the basic residues or substituting one by glutaminyl or histidyl residues.

164. Certain post-translational derivatizations are the result of the action of recombinant host cells on the expressed polypeptide. Glutaminyl and asparaginyl residues are frequently post-translationally deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Other post-translational modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the o-amino groups of lysine, arginine, and histidine side chains (T.E. Creighton, *Proteins: Structure and Molecular Properties*, W. H. Freeman & Co., San Francisco pp 79-86 [1983]), acetylation of the N-terminal amine and, in some instances, amidation of the C-terminal carboxyl.

165. It is understood that one way to define the variants and derivatives of the disclosed proteins herein is through defining the variants and derivatives in terms of homology/identity to specific known sequences. Specifically disclosed are variants of both the target molecules and known targets herein disclosed which have at least, 30%, 40%, 50% or 60% or 70% or 75% or 80% or 85% or 90% or 95% homology to the stated sequence. Those of skill in the art readily understand how to determine the homology of

two proteins. For example, the homology can be calculated after aligning the two sequences so that the homology is at its highest level.

166. The same types of homology can be obtained for nucleic acids by for example the algorithms disclosed in Zuker, M. *Science* 244:48-52, 1989, Jaeger et al. *Proc. Natl. Acad. Sci. USA* 86:7706-7710, 1989, Jaeger et al. *Methods Enzymol.* 183:281-306, 1989 which are herein incorporated by reference for at least material related to nucleic acid alignment.

167. It is understood that the description of conservative mutations and homology can be combined together in any combination, such as embodiments that have at least 70% homology to a particular sequence wherein the variants are conservative mutations.

168. It is understood that there are numerous amino acid and peptide analogs which can be incorporated into the disclosed compositions. For example, there are numerous D amino acids or amino acids which have a different functional substituent then the amino acids shown in Table 1 and Table 2. The opposite stereo isomers of naturally occurring peptides are disclosed, as well as the stereo isomers of peptide analogs. These amino acids can readily be incorporated into polypeptide chains by charging tRNA molecules with the amino acid of choice and engineering genetic constructs that utilize, for example, amber codons, to insert the analog amino acid into a peptide chain in a site specific way (Thorson et al., *Methods in Molec. Biol.* 77:43-73 (1991), Zoller, *Current Opinion in Biotechnology*, 3:348-354 (1992); Ibba, *Biotechnology & Genetic Engineering Reviews* 13:197-216 (1995), Cahill et al., *TIBS*, 14(10):400-403 (1989); Benner, *TIB Tech*, 12:158-163 (1994); Ibba and Hennecke, *Bio/technology*, 12:678-682 (1994) all of which are herein incorporated by reference at least for material related to amino acid analogs).

169. Molecules can be produced that resemble peptides, but which are not connected via a natural peptide linkage. For example, linkages for amino acids or amino acid analogs can include $\text{CH}_2\text{NH--}$, $\text{--CH}_2\text{S--}$, $\text{--CH}_2\text{--CH}_2\text{--}$, --CH=CH-- (cis and trans), $\text{--COCH}_2\text{--}$, $\text{--CH(OH)CH}_2\text{--}$, and $\text{--CHH}_2\text{SO--}$ (These and others can be found in Spatola, A. F. in *Chemistry and Biochemistry of Amino Acids, Peptides, and Proteins*, B. Weinstein, eds., Marcel Dekker, New York, p. 267 (1983); Spatola, A. F., *Vega Data* (March 1983), Vol. 1, Issue 3, Peptide Backbone Modifications (general review); Morley, *Trends Pharm Sci* (1980) pp. 463-468; Hudson, D. et al., *Int J Pept Prot Res*

14:177-185 (1979) (--CH₂NH--, CH₂CH₂--); Spatola et al. Life Sci 38:1243-1249 (1986) (--CH H₂--S); Hann J. Chem. Soc Perkin Trans. I 307-314 (1982) (--CH--CH--, cis and trans); Almquist et al. J. Med. Chem. 23:1392-1398 (1980) (--COCH₂--); Jennings-White et al. Tetrahedron Lett 23:2533 (1982) (--COCH₂--); Szelke et al. European Appln, EP 45665 CA (1982): 97:39405 (1982) (--CH(OH)CH₂--); Holladay et al. Tetrahedron. Lett 24:4401-4404 (1983) (--C(OH)CH₂--); and Hruby Life Sci 31:189-199 (1982) (--CH₂--S--); each of which is incorporated herein by reference. A particularly preferred non-peptide linkage is --CH₂NH--. It is understood that peptide analogs can have more than one atom between the bond atoms, such as b-alanine, g-aminobutyric acid, and the like.

170. Amino acid analogs and analogs and peptide analogs often have enhanced or desirable properties, such as, more economical production, greater chemical stability, enhanced pharmacological properties (half-life, absorption, potency, efficacy, etc.), altered specificity (e.g., a broad-spectrum of biological activities), reduced antigenicity, and others.

171. D-amino acids can be used to generate more stable peptides, because D amino acids are not recognized by peptidases and such. Systematic substitution of one or more amino acids of a consensus sequence with a D-amino acid of the same type (e.g., D-lysine in place of L-lysine) can be used to generate more stable peptides. Cysteine residues can be used to cyclize or attach two or more peptides together. This can be beneficial to constrain peptides into particular conformations. (Rizo and Gierasch Ann. Rev. Biochem. 61:387 (1992), incorporated herein by reference).

b) Pharmaceutically Acceptable Carriers

172. Disclosed herein are methods for inhibiting a receptor comprising incubating the receptor with a drug. Examples of such drugs discussed herein are purvalanol, imatinib, and SB203580. These drugs can be administered to treat a variety of diseases and disorders.

173. Suitable carriers and their formulations of the drugs disclosed herein are described in *Remington: The Science and Practice of Pharmacy* (19th ed.) ed. A.R. Gennaro, Mack Publishing Company, Easton, PA 1995. Typically, an appropriate amount of a pharmaceutically-acceptable salt is used in the formulation to render the formulation isotonic. Examples of the pharmaceutically-acceptable carrier include, but are not limited to, saline, Ringer's solution and dextrose solution. The pH of the solution is preferably

from about 5 to about 8, and more preferably from about 7 to about 7.5. Further carriers include sustained release preparations such as semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, e.g., films, liposomes or microparticles. It will be apparent to those persons skilled in the art that certain carriers may be more preferable depending upon, for instance, the route of administration and concentration of composition being administered.

174. Pharmaceutical carriers are known to those skilled in the art. These most typically would be standard carriers for administration of drugs to humans, including solutions such as sterile water, saline, and buffered solutions at physiological pH. The compositions can be administered intramuscularly or subcutaneously. Other compounds will be administered according to standard procedures used by those skilled in the art.

175. Pharmaceutical compositions may include carriers, thickeners, diluents, buffers, preservatives, surface active agents and the like in addition to the molecule of choice. Pharmaceutical compositions may also include one or more active ingredients such as antimicrobial agents, antiinflammatory agents, anesthetics, and the like.

176. The pharmaceutical composition may be administered in a number of ways depending on whether local or systemic treatment is desired, and on the area to be treated. Administration may be topically (including ophthalmically, vaginally, rectally, intranasally), orally, by inhalation, or parenterally, for example by intravenous drip, subcutaneous, intraperitoneal or intramuscular injection. The disclosed antibodies can be administered intravenously, intraperitoneally, intramuscularly, subcutaneously, intracavity, or transdermally.

177. Preparations for parenteral administration include sterile aqueous or non-aqueous solutions, suspensions, and emulsions. Examples of non-aqueous solvents are propylene glycol, polyethylene glycol, vegetable oils such as olive oil, and injectable organic esters such as ethyl oleate. Aqueous carriers include water, alcoholic/aqueous solutions, emulsions or suspensions, including saline and buffered media. Parenteral vehicles include sodium chloride solution, Ringer's dextrose, dextrose and sodium chloride, lactated Ringer's, or fixed oils. Intravenous vehicles include fluid and nutrient replenishers, electrolyte replenishers (such as those based on Ringer's dextrose), and the like. Preservatives and other additives may also be present such as, for example, antimicrobials, anti-oxidants, chelating agents, and inert gases and the like.

178. Formulations for topical administration may include ointments, lotions, creams, gels, drops, suppositories, sprays, liquids and powders. Conventional pharmaceutical carriers, aqueous, powder or oily bases, thickeners and the like may be necessary or desirable.

5 179. Compositions for oral administration include powders or granules, suspensions or solutions in water or non-aqueous media, capsules, sachets, or tablets. Thickeners, flavorings, diluents, emulsifiers, dispersing aids or binders may be desirable..

10 180. Some of the compositions may potentially be administered as a pharmaceutically acceptable acid- or base- addition salt, formed by reaction with inorganic acids such as hydrochloric acid, hydrobromic acid, perchloric acid, nitric acid, thiocyanic acid, sulfuric acid, and phosphoric acid, and organic acids such as formic acid, acetic acid, propionic acid, glycolic acid, lactic acid, pyruvic acid, oxalic acid, malonic acid, succinic acid, maleic acid, and fumaric acid, or by reaction with an inorganic base such as sodium hydroxide, ammonium hydroxide, potassium hydroxide, and organic
15 bases such as mono-, di-, trialkyl and aryl amines and substituted ethanolamines.

c) Therapeutic Uses

181. Effective dosages and schedules for administering the compositions may be determined empirically, and making such determinations is within the skill in the art. The dosage ranges for the administration of the compositions are those large enough to
20 produce the desired effect in which the symptoms disorder are effected. The dosage should not be so large as to cause adverse side effects, such as unwanted cross-reactions, anaphylactic reactions, and the like. Generally, the dosage will vary with the age, condition, sex and extent of the disease in the patient, route of administration, or whether other drugs are included in the regimen, and can be determined by one of skill in the art.
25 The dosage can be adjusted by the individual physician in the event of any counterindications. Dosage can vary, and can be administered in one or more dose administrations daily, for one or several days. Guidance can be found in the literature for appropriate dosages for given classes of pharmaceutical products. For example, guidance in selecting appropriate doses for antibodies can be found in the literature on therapeutic
30 uses of antibodies, e.g., Handbook of Monoclonal Antibodies, Ferrone et al., eds., Nokes Publications, Park Ridge, N.J., (1985) ch. 22 and pp. 303-357; Smith et al., Antibodies in Human Diagnosis and Therapy, Haber et al., eds., Raven Press, New York (1977) pp. 365-389. A typical daily dosage of the antibody used alone might range from about 1

µg/kg to up to 100 mg/kg of body weight or more per day, depending on the factors mentioned above.

182. Following administration of a disclosed composition, such as an antibody, for treating, inhibiting, or preventing a disease, the efficacy of the therapeutic antibody can be assessed in various ways well known to the skilled practitioner. For instance, one of ordinary skill in the art will understand that a composition, such as an antibody, disclosed herein is efficacious in treating or inhibiting a disease or disorder in a subject.

183. Other molecules that interact with targets to inhibit various interactions which do not have a specific pharmaceutical function, but which may be used for tracking changes within cellular chromosomes or for the delivery of diagnostic tools for example can be delivered in ways similar to those described for the pharmaceutical products.

184. The disclosed compositions and methods can also be used for example as tools to isolate and test new drug candidates for a variety of diseases and conditions.

E. Methods of making the compositions

185. The compositions disclosed herein and the compositions necessary to perform the disclosed methods can be made using any method known to those of skill in the art for that particular reagent or compound unless otherwise specifically noted.

1. Peptide synthesis

186. One method of producing the molecules, such as proteins, disclosed herein is to link two or more peptides or polypeptides together by protein chemistry techniques. For example, peptides or polypeptides can be chemically synthesized using currently available laboratory equipment using either Fmoc (9-fluorenylmethyloxycarbonyl) or Boc (*tert*-butyloxycarbonyl) chemistry. (Applied Biosystems, Inc., Foster City, CA). One skilled in the art can readily appreciate that a peptide or polypeptide corresponding to the disclosed proteins, for example, can be synthesized by standard chemical reactions. For example, a peptide or polypeptide can be synthesized and not cleaved from its synthesis resin whereas the other fragment of a peptide or protein can be synthesized and subsequently cleaved from the resin, thereby exposing a terminal group which is functionally blocked on the other fragment. By peptide condensation reactions, these two fragments can be covalently joined via a peptide bond at their carboxyl and amino termini, respectively, to form an antibody, or fragment thereof. (Grant GA (1992) Synthetic Peptides: A User Guide. W.H. Freeman

and Co., N.Y. (1992); Bodansky M and Trost B., Ed. (1993) Principles of Peptide Synthesis. Springer-Verlag Inc., NY (which is herein incorporated by reference at least for material related to peptide synthesis). Alternatively, the peptide or polypeptide is independently synthesized *in vivo* as described herein. Once isolated, these independent peptides or polypeptides may be linked to form a peptide or fragment thereof via similar peptide condensation reactions.

187. For example, enzymatic ligation of cloned or synthetic peptide segments allow relatively short peptide fragments to be joined to produce larger peptide fragments, polypeptides or whole protein domains (Abrahmsen L et al., Biochemistry, 30:4151 (1991)). Alternatively, native chemical ligation of synthetic peptides can be utilized to synthetically construct large peptides or polypeptides from shorter peptide fragments. This method consists of a two step chemical reaction (Dawson et al. Synthesis of Proteins by Native Chemical Ligation. Science, 266:776-779 (1994)). The first step is the chemoselective reaction of an unprotected synthetic peptide--thioester with another unprotected peptide segment containing an amino-terminal Cys residue to give a thioester-linked intermediate as the initial covalent product. Without a change in the reaction conditions, this intermediate undergoes spontaneous, rapid intramolecular reaction to form a native peptide bond at the ligation site (Baggiolini M et al. (1992) FEBS Lett. 307:97-101; Clark-Lewis I et al., J.Biol.Chem., 269:16075 (1994); Clark-Lewis I et al., Biochemistry, 30:3128 (1991); Rajarathnam K et al., Biochemistry 33:6623-30 (1994)).

188. Alternatively, unprotected peptide segments are chemically linked where the bond formed between the peptide segments as a result of the chemical ligation is an unnatural (non-peptide) bond (Schnolzer, M et al. Science, 256:221 (1992)). This technique has been used to synthesize analogs of protein domains as well as large amounts of relatively pure proteins with full biological activity (deLisle Milton RC et al., Techniques in Protein Chemistry IV. Academic Press, New York, pp. 257-267 (1992)).

F. Examples

189. The following examples are put forth so as to provide those of ordinary skill in the art with a complete disclosure and description of how the compounds, compositions, articles, devices and/or methods claimed herein are made and evaluated, and are intended to be purely exemplary and are not intended to limit the disclosure. Efforts have been made to ensure accuracy with respect to numbers (e.g., amounts,

temperature, etc.), but some errors and deviations should be accounted for. Unless indicated otherwise, parts are parts by weight, temperature is in °C or is at ambient temperature, and pressure is at or near atmospheric.

1. Example 1

a) Choice of drugs to study

190. The computational methods developed here accurately quantify the relative binding energetics of a drug of interest with many homologous receptors. To test the methods, seven different small-molecule inhibitors were selected for study, subject to the dual requirements of having high resolution structures in complex with at least one protein kinase target, and having experimental data for the drug's activity against ~20 protein kinases. Kinase inhibitors provide an excellent test system since they target members of a very large family of closely related sequences (6), are attractive therapeutic agents (7), yield high resolution crystal structures in complex with their targets (8), and are often experimentally assayed against panels of several different potential kinase targets (9), thus providing a body of data that any predictive method must be able to reproduce. In the present work, it is shown that it is possible to compute relative binding energies of drugs in complex with ~20 protein kinases that closely mirror experimental inhibition data. Three of the seven drugs (SB 203580, purvalanol B, and imatinib) have been screened against 493 human protein kinases; each of these 'kinome'-wide screens was completed on modest computational hardware in a single day.

191. The drugs chosen, followed in parentheses by the PDB codes of their kinase co-crystal structures, were: SB 203580 (1PME; 10), H89 (1YDT; 11), purvalanol B (1CKP; 12), hymenialdisine (1DM2; 13), imatinib (1OPJ; 14), indirubin-3'-monoxime (1E9H; 15), and quercetin (2HCK; 16). To prepare each drug structure for calculations, the PRODRG webserver (17) was used to build in atoms that were unresolved or absent from the crystal structures.

b) Construction of protein kinase models

192. The computation of a drug's affinity for a kinase of interest requires an accurate structural model of the protein. To construct the latter, a conservative approach was adopted: for each drug, all kinases to be tested were modeled with backbone conformations identical to those found in the drug-receptor crystal structure. With this assumption all that is required to construct structural models is a reliable sequence alignment. To this end, the Kinase Sequence Database (KSD) was used, a curated

alignment of the ATP- catalytic domains of over 7000 kinases constructed by Shokat and co-workers (21). The alignment of the sequences in this database with the sequence of each drug-bound crystal structure kinase was accomplished using the 'profile alignment' function of the multiple sequence alignment program CLUSTALW (22). Using the resulting alignment, a structure file for each kinase to be tested was created; no attempt was made to model insertions relative to the crystal structure, and residues outside of the catalytic (ATP-binding) domain were omitted. To complete preparation of the structures, sidechains were added using the rotamer-modeling program SCWRL (23), and hydrogens were added using the hydrogen bond-optimization module of the modeling program WHATIF (24).

193. FIG. 17 shows a comparison of the inhibitor-kinase contacts made by SB203580 and the crystal structures 1PME (mutant Erk2) and 1A9U (p38ALPHA). Leu103/104, Lys 53/54, Thr 105/06, Ala 51/52, and Met 108/109 interact with the inhibitor in both structures. This figure was generated by the program LIGPLOT (Wallace, A.C.; Laskowski, R. A.; Thornton, J.M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. Prot. Eng. 1995, 8, 127-134).

194. FIG. 18 shows a superimposed view of SB203580 in the binding sites of the crystal structures 1PME and 1A9U. Portions of the inhibitor making contacts important for affinity overlay well. The structure of the kinase in 1PME is shown in red and its bound SB203580 is shown in cyan; the structure of the kinase in 1A9U is shown in yellow and its bound inhibitor is shown in purple. The residues Lys 53 and MET 109 are shown making hydrogen-bonds to the inhibitor in dashed green lines (1A9U numbering). Residues ILE 31 and GLY 32 in the 1PME structure have been removed from the figure for clarity. An all-atom alignment was performed using residues 61-79, 104-116 (1A9U numbering) and SB203580. The superimposition and figure were prepared using Swiss-Pdb Viewer: An environment for comparative protein modeling. Electrophoresis 1997, 18, 2714-2723).

c) Computation of drug-receptor binding energies

195. For the purposes of calculating the binding energy of a drug with a potential receptor, a new software program (SCR) was written in FORTRAN90. The program incorporates flexibility in the protein residues that are close to the drug through the use of rotamer libraries (5) that are sampled by Monte Carlo methods (25). The rotamer libraries used in the calculations reported here were the highest resolution sets

developed by Xiang and Honig (26), providing good coverage of conformational possibilities; the library for arginine for example contains 1948 different rotamers. For each drug-receptor combination studied, flexibility was modeled for all residues that had at least one atom located within 5.0Å of any drug atom in the initial SCWRL- built model (see above); typical calculations involved ~20 moving residues. A structural picture of the rotamers sampled in a typical binding site is shown in Figure 3.

196. The sampling of sidechain positions and the computation of the binding thermodynamics were accomplished using a simple, empirical function that models the energy of the drug- receptor system as a sum of electrostatic and van der Waals interactions between all pairs of atoms. Electrostatic interactions (in kcal/mol) were computed using a basic Coulombic model:

$$E_{\text{elec}} = 332.08 q_1 q_2 / \epsilon r$$

where q_1 and q_2 are the partial atomic charges on the interacting atoms (in proton units), r is the distance between the two atoms (in Ångstroms), and ϵ is the relative dielectric constant, assumed here to be 78 (i.e. that of water). For the protein atoms, partial atomic charges were taken from the PARSE parameter set (27); for the drugs, partial charges were assigned on the basis of analogy with similar functional groups present in the PARSE parameter set, e.g. for carbonyl groups, charges of +0.5e and -0.5e were assigned to the C and O atoms respectively.

197. Additional contributions to the computed energy were made by van der Waals interactions, with a deliberately simple approach again being adopted. For interacting pairs in which both atoms were carbon or sulfur, a Lennard-Jones 12-6 interaction was used:

$$E_{\text{vdw}} = \epsilon \{ \sigma_{\text{att}}^{12} / r^{12} - \sigma_{\text{att}}^6 / r^6 \}$$

For all other interacting pairs, a purely repulsive $1/r^{12}$ term was calculated:

$$E_{\text{vdw}} = \epsilon \{ \sigma_{\text{rep}}^{12} / r^{12} \}$$

198. In these expressions σ_{att} and σ_{rep} are constants (in the former case corresponding to the distance at which the Lennard-Jones interaction changes from being attractive to repulsive), ϵ has the units of energy (kcal/mol) and r is the distance between the two atoms. A single σ_{rep} value was used for all interactions between non-hydrogen atoms, and (based on early calculations) this was always 0.75Å shorter than the single σ_{att}

value used for all C–C, C–S and S–S interactions. H–H interactions were assigned a σ_{rep} of 0.5 Å; for mixed interactions between hydrogen and non-hydrogen atoms, we calculated σ_{rep} as the geometric mean: $\sigma = \sqrt{\sigma_{\text{rep},i} \sigma_{\text{rep},j}}$. As a further simplifying assumption, a single value of ϵ (0.2 kcal/mol) was assigned to all pairwise interactions. The rationale for the above definitions is that they provide a straightforward but effective way of accounting for hydrophobic interactions: contacts between carbons and sulfurs (all of which are here considered hydrophobic) are energetically rewarded, whereas all other contacts are considered to be energetically neutral (the $1/r^{12}$ term being used only to prevent overlap of atoms).

199. With the energy function defined above, binding energies for each drug-receptor combination were calculated in the following way. Monte Carlo sampling of the flexible sidechains in the receptor was conducted in the presence and absence of the drug; in the former simulations the drug was held fixed in the position it adopted in the crystal structure. Trial moves were made by first randomly selecting one of the moveable sidechains, and then randomly selecting a new rotamer for the chosen sidechain. Following each trial move and the computation of the new energy, a Metropolis test (28) was applied (assuming a temperature of 298K) to determine whether to accept the new conformation. For both the drug-present and drug-absent situations, 10 independent simulations were performed, each consisting of one million ‘equilibration’ MC steps, followed by 5 million ‘production’ MC steps; during each production phase the average energy of the moving residues was accumulated. The drug-receptor binding energy was then computed as the difference between the best average energy found in the 10 simulations without the drug, and the best average energy found in the 10 simulations with the drug present. Use of 10 independent simulations was made to avoid complications occurring when individual MC simulations became locked in local energy minima; such situations were rare, and the differences among the 10 computed values were usually within 0.01-0.03 kcal/mol.

d) Quantifying predictions of receptor selectivity

200. To validate the predictions made by SCR and to pave the way for its use in making large-scale predictions for the entire human ‘kinome’, one of the first goals of the present work was to calculate the binding energy of each of the seven selected drugs with ~20 protein kinases and to compare the computed relative binding energies with

experimental results. Experimental inhibition data for the drugs was obtained from several sources. For SB 203580, H89, indirubin-3'-monoxime and quercetin, data were taken from studies by Cohen's group (9, 18); these data consist of the percent activity of >20 protein kinases following addition of a fixed concentration (usually 10 μ M) of the drug. Inhibition data for the compounds purvalanol B, hymenialdisine and imatinib were obtained in the form of IC₅₀ data primarily from references 12, 13, and 19 respectively; for imatinib three additional targets (cKit, PDGFR α , and PDGFR β) were identified in ref. 20.

201. If direct experimental *binding* data were available, it would be possible to assess the computational predictions by linear regression of the computed and experimental energies. However, since the available data (e.g. IC₅₀ values) provide only indirect estimates of relative binding affinities, we have used a simple classification-based scheme to quantify the accuracy of the computed results: kinases were classified as either being 'targets' or 'non-targets' of a particular drug according to their degree of inhibition observed experimentally. In the case of those drugs for which IC₅₀ data have been reported, kinases with IC₅₀ values <100 nM were categorized as 'targets'; in the case of drugs for which percent kinase activity has been reported (refs. 9 and 26), only those kinases with activities <50% were categorized as 'targets'. All kinases not meeting these criteria were categorized as 'non-targets' of the drug.

202. A similar binary classification scheme was also applied to the computed binding energies: a cut-off was chosen (see below), as described below, such that kinases with binding energies equal to or more favorable than the cut-off were classified as 'computed targets', with all other kinases being 'computed non-targets'. Once kinases were classified as 'targets' or 'non-targets' according to both the computed and experimental data, the measures of predictive value theory were used to quantify the degree of agreement between the two sets of classifications. Although a variety of measures could in principle be used (specificity, sensitivity etc.), we chose to quantify the success of the computations were quantified by calculating the classification *efficiency*, defined as:

$$\text{Efficiency} = \frac{\{\# \text{truepositives} + \# \text{truenegatives}\} \times 100\%}{\{\# \text{truepositives} + \# \text{falsepositives} + \# \text{truenegatives} + \# \text{falsenegatives}\}} \quad (1)$$

where 'true positive' denotes a kinase determined to be a 'target' of the drug both computationally and experimentally, 'false positive' a kinase computed to be a 'target'

but designated a "non-target" experimentally, 'false negative' a kinase computed to be a 'non-target' when designated a 'target' experimentally, and 'true negative' a kinase determined to be a 'non-target' both computationally and experimentally.

203. Having a defined, quantitative measure of the degree of correspondence between computed and experimental 'target'/'non-target' classifications of kinases provides a route to optimizing SCR so that its results more closely describe reality. Only very minor adjustments were made; in fact, for each drug studied here, we have adjusted only a single parameter parameter was adusted: the value of σ_{rep} (which in turn defines σ_{att} ; see above). For six of the seven drugs, an optimal value of σ_{rep} was found simply by varying it in 0.1 Å increments within the range 2.6 Å to 3.4 Å and finding the value that maximized the classification efficiency; for the particular case of hymenialdisine, optimal results were obtained with a σ_{rep} of 2.4 Å. The extent to which adjustment of this single parameter maximizes the classification efficiency for the full set of ~20 kinases for each drug provides a key first indication of the potential utility of the computational method; in particular, it demonstrates how well the method operates when 'trained' on a relatively large set of data. (Figure 8)

204. Binding energies computed for a prototypical panel of six kinases tested against a hypothetical drug are shown. The three kinases that are true experimental 'targets' (A, B and C) are shaded; those that are experimental 'non-targets' (D, E and F) are unshaded. In each column, kinases are listed in order of their *computed* binding energies. A binding energy cutoff (indicated by the bold line) separates those kinases that are computed to be 'targets' from 'non-targets': those kinases lying above the cutoff in the Figure 8 are computed 'targets.' In the above example, the cutoff is set equal to the least favorable binding energy of the experimental 'targets' (see text). When the calculations are performed with the atomic van der Waals radius (σ_{rep}) set to 2.6 Å, the cutoff is such that *all* kinases in the panel would be computed 'targets'; given that this would result in three false positives, this would clearly be a less than ideal result. When the radius is increased to 2.7 Å, it is possible to set the cutoff so that only one false positive (kinase D) is obtained; when a radius of 2.8 Å is used, a perfect result can be achieved with no false positives. TP denotes true positive, and TN denotes true negative.

205. FIG. 19 shows an ordered list of the computed binding energies obtained with SCR for five inhibitors for which the calculations were successful. Results shown are those obtained when σ_{rep} was set to 2.8 Å and σ_{att} to 3.55 Å in all calculations.

Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes. The precision of the computed binding energies is ~0.03 kcal/mol.

206. FIG. 20 shows an ordered list of the computed binding energies obtained from docking calculations with AutoDock. Energies shown are those of the docking pose with the most favorable binding energy. Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes.

207. For the inhibitors that have experimental data in the form of percentage activity, FIG. 21 shows only the kinases that have activities of <10% to be targets, rather than the 50% cutoff used in the other tables. The only inhibitor panel to be seriously affected is H89, whose interaction energy with S6K1 is significantly lower than those with its other targets. S6K1 could be considered a “false negative” in this case. Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes (percent activity <10%).

208. FIG. 22 shows an ordered list of the computed binding energies obtained with SCR for five inhibitors for which the calculations were successful. The parameterized values of σ_{rep} used to obtain these results for each inhibitor/kinase panel were SB203580 (1PME & 1A9U): 2.8 Å; purvalanol B: 2.6 Å; imatinib: 2.9 Å; H89: 2.9 Å; and hymenialdisine: 2.4 Å. Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes. The precision of the computed binding energies is ~0.03 kcal/mol.

209. FIG. 24 shows an ordered list of the computed binding energies obtained using AutoDock and the fixed-inhibitor assumption. Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes.

210. FIG. 25 shows an ordered list of the computed binding energies obtained from docking calculations conducted with AutoDock. Energies shown are those of the docking pose with the RMSD closest to the crystal structure ligand. Kinases determined to be “experimental targets” of the inhibitor are indicated by shaded boxes.

211. FIG. 26 shows an ordered list of the computed binding energies obtained by applying AutoDock’s energy function to complexes that were first energy-minimized by GROMACS.

212. FIG. 27 shows an ordered list of the computed binding energies obtained with SCR.

e) Validation of predictive ability

213. In order to provide a direct route to assessing the likely *predictive* ability of the computational method, a training/testing procedure was devised and carried out for the three drugs that showed the most selective experimental inhibition profiles (SB 203580, purvalanol B, and imatinib). For each drug, the ~20 kinases that have been experimentally studied were randomly divided into two sets (a 'training' set and a 'testing' set), subject only to the requirement that both sets contained the same numbers of 'targets' (one or two) and 'non-targets' (between seven and ten). For the training set, a σ_{rep} value that maximized the classification efficiency was found in the same way as described above for the full set of ~20 kinases; note however, that in this new scenario the method was 'trained' with only *half* of the available experimental data instead of the full set used previously. The same σ_{rep} value found for the training set, together with its corresponding cutoff energy, was then applied *unaltered* to the kinases in the testing set and the classification efficiency in the testing set determined. This process of randomly dividing the kinases into training and testing sets, determining an optimal σ_{rep} value from a training set and quantifying classification efficiencies for a testing set using identical conditions, was repeated a total of 1000 times for each drug in order to obtain a statistically meaningful measure of the classification efficiencies that might be expected in truly predictive scenarios. In order to determine whether the results obtained with the method exceed those expected on the basis of chance, the above sampling procedure was conducted a second time, but with the labels 'target' and 'non-target' being randomly reassigned among the kinases within both the training and testing sets in each of the 1000 randomly drawn samples (while keeping the number of targets and non-targets in each panel the same as in the non-random scenario). Histograms of the computed classification efficiencies in the testing sets were constructed for each drug with both the true experimental 'target'/'non-target' classifications and the randomly reassigned classifications (see Results).

214. FIG. 23 shows a summary of training/testing results for various screening protocols applied to five inhibitors and their respective kinase panels. Percent efficiencies are shown along with standard deviations for 1000 training/testing iterations. SCR refers to Side Chain Rotamer program; AD FI refers to AutoDock calculations within the fixed-inhibitor assumption; AD GA (BE) refers to AutoDock genetic algorithm dockings in which the docked ligand with the best energy is selected; AD GA (BR) refers to docking in which the docked ligand that has the lowest RMSD to the crystal structure is selected;

GMC/AD FI refers to the case where the models are energy-minimized with GROMACS prior to the AutoDock fixed-inhibitor calculation, and Random shows the efficiency that would be expected if a method had no predictive ability. With the exception of the hymenialdisine panel, keeping the ligand in the same orientation for all members of a panel (as is done in the SCR and AD FI cases) produces better efficiencies than allowing the ligand to assume different orientations in each receptor of a given panel.

f) Defining a binding energy cutoff

215. Central to the method's use in each of the above tests is the definition of the binding energy cutoff that separates the computed 'targets' from 'non-targets.' In the initial set of calculations aimed at investigating how well the method classifies kinases when trained on *all* the available experimental data for a given drug (see above), the cut-off was simply set equal to the weakest computed binding energy of the known experimental 'targets'. By definition, this approach ensures that all of the known 'experimental targets' of the drug are also 'computed targets', so the real test of the method in such a scenario is the extent to which it also incorrectly identifies 'experimental non-targets' as 'computed targets' (i.e. the rate at which it predicts 'false positives').

216. For the training/testing calculations aimed at demonstrating how well the method can predict 'targets' and 'non-targets' when trained on only half of the available data, it became clear that in many cases assigning the binding energy cut-off to be equal to the weakest binding energy of the known targets in the training set was too restrictive a criterion. In particular, experience showed that such a criterion can fail to identify true 'targets' present in the testing set because their computed binding energies were slightly less favorable than the computed binding energies of the true 'targets' present in the training set. To eliminate this problem – which results in an unnecessarily high rate of false negatives in the testing set – the definition of the binding energy cutoff was modified to be equal to the weakest binding energy of the known targets in the training set *multiplied by a scaling factor* between 0.80 and 1.00. This has the effect of making the binding energy cutoff less negative, so that more receptors in the testing set are classified as computed 'targets'; this in turn means that the number of false negatives is reduced, while the number of false positives increases. To thoroughly test the use of different scaling factors, the full set of 1000-sample training/testing calculations was repeated with each of the following factors being applied: 0.80, 0.85, 0.90, 0.95 and 1.00.

g) Large scale applications to human protein kinases

217. For three of the drugs studied here, the computations were extended beyond the ~20 kinases for which experimental data was available, to the full array of human protein kinases. As before, the creation of structural models required a suitable sequence alignment. To this end, 637 human kinase sequences were taken from the comprehensive study performed by Manning et al. (6); sequences with atypical kinase domains and non-coding pseudogenes were removed from this list to give a total of 493 kinases to be screened against the drugs. 363 of these sequences could be found immediately in the Shokat group's KSD alignment; the remaining 130 sequences were added to the alignment using CLUSTALW as described above. Binding energy computations for these 493 kinases were conducted in the same way as before; a complete set of calculations for one drug required about 24hours of CPU time on a 16-node 2.6GHz PC cluster.

h) Results

218. The goal of the methodology reported here is the accurate identification of those receptors that are targets of a drug based solely on relative binding energies computed from atomistic MC simulations. As a first test of the method, its ability to correctly discriminate between ~20 'targets' and 'non-targets' of seven different kinase inhibitors was investigated (see Methods). The extent to which accurate discrimination could be achieved is indicated by the computed binding energies listed in Table 3 for each drug-kinase combination tested. For each drug, the numbers listed are those producing the best agreement with experiment, as quantified by the 'classification efficiency' (see Methods). Importantly, calibration of the method was achieved (separately for each drug) by adjusting only a *single* parameter in the energy function: the van der Waals radius (σ_{rep}) – identical for all non-hydrogen atom types. The dependence of the classification efficiency on the range of σ_{rep} values investigated is shown for all seven drugs in Supporting Information.

219. For five of the drugs (SB203580, purvalanol B, imatinib, H89 and hymenialdisine), the agreement is clearly excellent: when the kinases are listed in order of decreasing affinity for each drug, the known experimental targets (shaded boxes) all cluster at the top of each list (Table 3). For two of these five drugs, SB 203580 and purvalanol B, the calculations correctly identify all experimental 'targets' as having

computed binding energies that are clearly more favorable than all ‘non-targets’. This general trend is maintained with imatinib, H89 and hymenialdisine, The agreement is very good, and if the weakest binding energy of the known ‘targets’ of each drug as a *binding energy cutoff* for discriminating between ‘targets’ and ‘non-targets’, a total of 27 true ‘targets’ can be correctly identified at the expense of only 8 experimental ‘non-targets’ being incorrectly computed to be ‘targets’. A summary of the accuracy of the ‘target’/‘non-target’ classifications is provided in Table 4.

220. For the five drugs for which the calculations were successful, it is clear that ‘targets’ and ‘non-targets’ can be discriminated with a high degree of confidence, and that this appears to be true for the more selective inhibitors SB 203580, purvalanol B and imatinib (Table 3). Crucially, the kinases that are correctly computed to be ‘targets’ of a particular drug are often only distantly related to each other in terms of sequence, and their common status as ‘targets’ is therefore not a result that could be reproduced by a trivial comparison of their sequences. To illustrate this, phylogenetic trees constructed from the ‘sequence’ of residues in contact with each drug are shown in Figure 4 for the five drugs for which the computations were successful; similar results were obtained when trees were constructed from *all* residues in the catalytic domain .

221. Although there is an expected overall tendency for the ‘experimental targets’ of each drug to cluster in the trees, there are clear and interesting exceptions, and a particularly striking observation is that two of the ‘targets’ of SB 203580 (p38 α and p38 β) are phylogenetically very remote from the drug’s third ‘target’, LCK (6). A closer examination of the computed results for these kinases reveals why they are successfully identified as ‘targets’. In the computations a significant favorable contribution to drug-binding to the p38 kinases is made by the sidechain of Leu144 (numbered according to the KSD entry for p38 α). This contribution is absent in LCK because of a Leu $\rightarrow$ Ala substitution at that position, and if this were the only difference between the two kinases, a decreased affinity for the drug would be computed. However, a simultaneous Ala $\rightarrow$ Leu substitution at the nearby position 137 in LCK compensates almost perfectly for the lost interaction such that the binding affinities to the p38 and LCK kinases are all computed to be more or less identical. When phylogenetic trees were constructed for the problematic drugs quercetin and indirubin-3’-monoxime, the kinases identified experimentally as targets did not show any obvious clustering at all.

222. A scaling factor was applied to the binding energy cutoff determined from the training set results in order to limit the number of false negatives obtained in the testing set calculations. Based on a more complete set of calculations, a scaling factor of 0.90 was determined to be most appropriate, which in a practical situation means that if a binding energy cutoff of -10.00 kcal/mol is found to be sufficient to identify all true 'targets' in a training set, then a more reasonable binding energy cutoff to use in a predictive, testing scenario can be $-10.00 \times 0.90 = -9.00$ kcal/mol. This would in effect allow for the possibility that other kinases can exist outside of the training set that, while binding the drug somewhat weaker, nevertheless still bind it sufficiently strongly to be considered 'targets'.

223. The distributions of calculated classification efficiencies for the testing sets, obtained from 1000 training/testing procedures with this scaling factor of 0.90 applied, are shown for each of the three drugs in Figure 9. Also shown are distributions of calculated classification efficiencies obtained when the 'target'/'non-target' classifications are reassigned randomly among the kinases rather than being taken from experiment. For all three drugs, the predictions obtained with the computational method are clearly far superior to those expected on the basis of chance: the non-parametric Mann-Whitney U test two-tailed p value was <0.0001 for a 99.0% confidence interval, establishing a highly significant difference between the medians of the model-computed and randomized distributions.

224. FIG. 14 shows the distribution of classification efficiencies of the testing sets obtained from SCR calculations. SCR's computed efficiencies are shown as dark bars; those obtained from randomized trials as white bars. FIG. 15 shows the distribution of testing set classification efficiencies obtained from AutoDock calculations with the fixed inhibitor assumption. AutoDock's computed results are shown as dark bars, those obtained from randomized trials as white bars.

225. The demonstration of a significant predictive power to the computations provided sufficient confidence to attempt the ultimate goal of this work, which was to make reasonable predictions of the drugs' activity against a far larger sample of potential targets. To this end, binding energy calculations were carried out for SB 203580, purvalanol B and imatinib, each with a more or less full complement of 493 human protein kinases (see Methods). A histogram of the computed binding energies obtained from the large-scale screen of imatinib is shown in Figure 2; the distribution is

approximately normal, with a pronounced skew to the right due to kinases computed to have very positive binding energies (because of strong steric clashes with the drug). From the earlier calculations of imatinib binding to ~20 kinases, a binding energy cutoff of -16.78 kcal/mol was identified (Table 3); applying a scaling factor of 0.90 converts this to a cutoff of -15.10 kcal/mol that might be reasonable to use in a predictive setting (see above). Using this cutoff, 32 of the 493 kinases (6.5 %) were predicted to be 'targets' of imatinib. Similar histograms of computed binding energies were obtained from computational screens of SB 203580 and purvalanol B; these screens produced a total of 55 and 111 predicted 'targets' respectively for the two drugs.

5

226. **Table 3.** An ordered list of the computed binding energies obtained for each of the seven drugs with each kinase for which experimental data are available.

SB 203580	Purvalanol B	Imatinib	H89	Hymenialdisine	Indirubin	Quercetin
Kinase ΔE	Kinase ΔE	Kinase ΔE	Kinase ΔE	Kinase ΔE	Kinase ΔE	Kinase ΔE
LCK -11.82	CDK2 -8.83	ABL -17.91	ROCK2 -13.99	CK1 α -6.90	GSK3 β -10.35	PKB α -9.35
p38 α -11.57	cdc2 -8.77	SRC -17.90	CHK1 -13.00	ARAF -6.57	PHK γ 1 -8.98	PKA -9.21
p38 β -11.41	CDK5 -8.42	LYN -17.90	LCK -12.80	PKA -5.92	PKC α -8.88	PRAK -9.17
Erk2 -9.24	DYRK1A -8.41	PDGFR α -17.07	MSK1 -12.64	cdc2 -5.91	MKK1 -8.53	SGK -8.98
PRAK -8.54	ABL -8.08	PDGFR β -16.97	PKA -12.38	CDK3 -5.89	AMPK α 1 -8.40	MKK1 -8.81
MKK1b -8.28	PKC ζ -7.66	KIT -16.78	MKK1 -12.36	CDK2 -5.86	S6K1 -8.36	GSK3 β -8.74
S6K1 -8.26	GSK3 β -7.64	FGFR1 -14.80	AMPK α 1 -11.57	PKC ζ -5.81	CHK1 -7.97	AMPK α 1 -8.19
PKB α -7.96	cdc28 -7.42	PKC α -9.44	SGK -11.49	GSK3 β -5.69	MSK1 -7.82	MSK1 -8.09
PKA -7.70	CDK4 -7.39	PKC β -8.86	GSK3 β -11.24	CDK5 -5.56	LCK -7.78	MKK2 -8.00
ROCK2 -7.59	PKC ϵ -7.35	PKC η -8.43	PKB α -11.11	ASK γ -5.54	MKK1b -7.58	PKC α -7.81
MKK2 -6.97	PKA -7.33	PKC ϵ -5.17	S6K1 -11.06	PKC δ -5.47	PKA -7.11	CHK1 -7.69
PHK γ 1 -6.57	PKC δ -7.20	EGFR -2.77	MKK1b -10.74	PKC ϵ -5.45	SGK -7.10	LCK -7.50
AMPK α 1 -4.79	INSR -7.00	PKA -1.29	PKC α -10.62	PKC η -5.42	ROCK2 -7.08	MKK1b -7.35
PKC α -4.50	PKC η -6.96	PKC γ -0.40	Erk2 -10.60	CDK4 -5.33	Erk2 -6.65	ROCK2 -7.05
GSK3 β -4.21	Erk1 -6.78	PHK γ 1 -11.68	PRAK -9.85	PKC γ -5.31	S6K1 -6.53	S6K1 -6.73
CHK1 -3.99	PKC β -6.66	cdc2 -13.87	MKK2 -9.05	Erk1 -5.27	p38 δ -6.36	Erk2 -5.76
JNK1 -3.22	PKC α -6.65	CK2 α 1 -18.69	p38 γ 4.25	CDK6 -5.25	PKB α -6.27	p38 γ -5.44
MKK1 -3.20	CK1 α -6.53	PKC ζ 277.04	p38 β 4.93	Erk2 -5.23	MKK2 -5.93	p38 δ -5.38
p38 δ -2.90	ARAF -6.47	CK1 α 470.10	p38 α 5.96	MKK1 -5.16	p38 α -5.41	JNK1 -4.82
p38 γ -2.66	PKC γ -6.18		p38 δ 6.13	PKC β -5.16	CK2 α 1 -5.33	p38 α -4.67
SGK 1.76	JNK1 -5.61		JNK1 12.08	PKC α -5.15	PRAK -5.16	p38 β -4.28
MSK1 2.04	PKG1 -5.22		PHK γ 1 51.85	SRC -5.05	p38 β -5.00	PDK1 -1.86
PDK1 28.42	CK2 α 1 -4.92		CK2 α 1 71.70	PKG1 -5.03	PDK1 -4.26	CK2 α 1 -13.14
CK2 α 1 38.97			PDK1 1053.26	ABL -5.01	JNK1 -2.17	PHK γ 1 -54.13
			Eg2	-4.99		
			INSR	-4.84		
			CK2 α 1	-3.73		
			JNK1	-3.30		

227. **Table 4.** A summary of the accuracies of the predictions tabulated in Table 1

	SB.	P.B	Ima.	H89
Tested	24	23	19	24
TP	3	4	4	9
FP	0	0	2	3
TN	21	19	13	12
FN	0	0	0	0
Eff.%	100	100	89.5	87.5
	Hym.	Ind.	Que.	
Tested	28	24	24	
TP	7	13	8	
FP	3	11	16	
TN	18	0	0	
FN	0	0	0	
Eff.	89.3	54.1	33.3	

228. Importantly, the kinases that were correctly identified as targets by the binding energy computations were often only distantly related to each other in terms of sequence. Phylogenetic trees constructed from the 'sequence' of residues in contact with each drug are shown in Figure 4 for the 5 drugs for which the computations were successful; similar results were obtained when trees were constructed from all residues in the catalytic domain. Although there is an expected overall tendency for the targets of each drug to cluster in the trees, there are clear exceptions: the two p38 targets of SB 203580 for example are phylogenetically remote from the drug's third target, LCK. Interestingly, a significant favorable contribution to drug-binding to p38 α is made by the sidechain of Leu144 (KSD numbering); this contribution is lost in LCK because of a Leu to Ala substitution at that position. However, a simultaneous Ala $\leftrightarrow$ Leu substitution at the nearby Ala137 (in p38 α) neatly compensates for the lost interaction such that the binding affinities to p38 α and LCK are more or less identical. The notion that even quite distantly related kinases can be targets of the same drug has been expressed already by Cohen and co-workers (8); it is notable however that this finding is reproduced by the current structure-based computations since it is unlikely to be easily captured with purely sequence-based approaches. Interestingly, when phylogenetic trees were constructed for quercetin and indirubin-3'-monoxime, the kinases identified experimentally as targets did not show any obvious clustering (Figure 5).

229. For the five successful drugs, the ability of the computations to reproduce experimental data for ~20 kinases provided sufficient confidence to make predictions of drug activity against a much larger sample of potential targets. To this end, binding energy calculations were carried out for three of the drugs (SB 203580, purvalanol B and imatinib) with 493 human kinases. Figure 2 shows a histogram of the computed binding energies obtained from the large- scale screen of imatinib; the distribution is approximately normal, with a pronounced skew to the right due to kinases computed to have very positive binding energies (because of strong steric clashes with the drug). In total, 22 of the 493 kinases (4.5%) are predicted to be targets of imatinib; the full list of such targets, which contains several kinases that are of potential therapeutic interest is given in Table 5. Similar histograms of computed binding energies were obtained from computational screens of purvalanol B and SB 203580 (Figure 6); these screens produced a total of 33 and 36 predicted targets respectively for the two drugs (Tables 6 & 7).

230. Table 5. An ordered list of human protein kinases predicted to be targets of imatinib. Known targets of the drug are indicated by shaded boxes. An additional known target (ARG) that falls just outside of the cutoff is indicated at the bottom of the list.

5

Kinase	Acc.	Energy
FMS	SK094	-22.20
TEC	4507429	-20.30
MYT1	2460023	-19.56
IKKb	3213217	-19.52
RiPK2	4506537	-18.66
RET	5419753	-18.64
YES	125870	-18.38
BMX	4502435	-18.36
CSK	729887	-18.32
HCK	2194103	-18.14
FRK	4503787	-18.10
BLK	914204	-18.04
FGR	125358	-17.96
ABL	625609	-17.91
SRC	125711	-17.90
LCK	8569365	-17.90
LYN	125480	-17.90
ACK	423137	-17.80
PDGFR α	5453870	-17.07
IKKa	4502843	-17.04
PDGFR β	4505683	-16.97
KIT	1817734	-16.78
FGFR2	533220	-16.66
HER4/ErbB4	422854	-16.10
FYN	6978863	-15.96
FLT1	4503749	-15.92
SYK	2136036	-15.88
FGFR4	120051	-15.72
FAK	3183518	-15.70
FLT4	388522	-15.62
Wnk2	SK016	-15.42
Wnk3	SK641	-15.24
...		
ARG (rank #39)	1168268	-14.32

231. Table 6. Same as the previous table, except results are for purvalanol B.

Kinase	Acc.	Energy	Kinase	Acc.	Energy	Kinase	Acc.	Energy
MAK	7019738	-9.02	(cont.)			(cont.)		
FLT4	388522	-8.98	DYRK1B	6753698	-8.29	FGFR1	220738	-7.81
MUSK	5031927	-8.95	CDK9	4502747	-8.28	DYRK4	SK116	-7.80
CDK3	231726	-8.88	CDKL3	7108631	-8.24	EphB3	542989	-7.80
KDR	2655412	-8.86	AurB	5688866	-8.23	TIE2	464868	-7.79
PCTAIRE2	4505649	-8.83	JAK2	3236322	-8.22	CDK6	266423	-7.78
CDK2	6733417	-8.83	TIE1	512414	-8.21	Fused	6331315	-7.77
PCTAIRE1	266425	-8.82	AurC	3127068	-8.21	PKACg	3115220	-7.76
CDC2	115922	-8.77	MSK1	3411157	-8.17	NEK9	4885696	-7.75
FLT3	1362915	-8.76	PEK	7341091	-8.15	SRPK1	3135976	-7.74
CDKL1	SK203	-8.76	MER	5453738	-8.11	TYK2	267184	-7.74
Erk3	1294779	-8.73	PFTAIRE2	SK462	-8.11	RSK1	401772	-7.72
ICK	SK497	-8.72	PIM2	6470337	-8.11	RSK3	6677811	-7.72
CDK7	4502743	-8.68	SGK	5442269	-8.10	HCK	2194103	-7.72
TRKA	339918	-8.65	ABL	625609	-8.08	RSK2	1730070	-7.71
PCTAIRE3	547804	-8.64	Wee1	499072	-8.08	RSK4	6467562	-7.69
CDC7	2102637	-8.63	PFTAIRE1	4240157	-8.08	EphA6	SK646	-7.68
Erk4	4506089	-8.60	LMR2	7662476	-8.07	PKC ζ	6166241	-7.66
GCN2	7243057	-8.59	CDKL2	4505569	-8.07	CHED	4502711	-7.65
ROR1	346351	-8.57	Wee1B	4153873	-8.07	GSK3B	423729	-7.64
NEK3	SK252	-8.55	PAK5	6331022	-8.03	DMPK2	SK112	-7.64
FLT1	4503749	-8.55	CLK3	3913243	-8.03	JAK3	508731	-7.64
NEK6	6009759	-8.53	TLK1	6063017	-8.02	MRCKb	5006445	-7.64
PDGFRa	5453870	-8.50	TLK2	6063019	-8.02	PYK2	1165219	-7.63
FGFR2	533220	-8.49	PAK4	6329959	-8.01	ITK	5031811	-7.61
CASK	3087818	-8.49	EphA1	339717	-7.99	IRAK1	1302664	-7.61
ROR2	4758842	-8.49	EphA7	755568	-7.97	PKC ι	6679349	-7.61
Erk7	SK465	-8.48	JAK1	125060	-7.94	MRCKa	1695873	-7.60
NEK7	SK421	-8.45	MSK2	4506735	-7.90	MLK1	462606	-7.60
CCRK	7106269	-8.44	DDR1	4503451	-7.89	MAP2K5	1616781	-7.60
TRKB	530791	-8.42	KIT	1817734	-7.89	HRI	6580979	-7.59
CDK5	572619	-8.42	CDK11	SK443	-7.89	EphA10	SK627	-7.59
DYRK1A	3219996	-8.41	CDK8	4502745	-7.88	DMPK1	976147	-7.59
TRKC	442390	-8.34	FGFR3	182569	-7.87	CDKL4	SK466	-7.58
MPSK1	3646280	-8.33	PKC τ	5453976	-7.85	YES	125870	-7.58
AurA	4507275	-8.31	DDR2	5453814	-7.85	EphB6	4758292	-7.58
MAP3K4	1504010	-8.30	SRPK2	3406051	-7.84	SYK	2136036	-7.57
RET	5419753	-8.29	PDGFRb	4505683	-7.83			

232. **Table 7.** An ordered list of the most favorable computed binding energies obtained from a screen of SB 203580 with 493 human protein kinases.

Kinase	Acc.	Energy	Kinase	Acc.	Energy
EphA1	339717	-13.64	(cont.)		
EphB3	542989	-13.44	p38 α	6733369	-11.57
EphB1	4104413	-13.04	EphA5	1177466	-11.56
EphB4	495473	-12.90	SRM	SK425	-11.52
RIPK3	5803143	-12.82	YANK3	SK469	-11.52
EphB2	1706665	-12.54	YANK1	SK624	-11.48
DDR1	4503451	-12.44	SRC	125711	-11.48
FRK	4503787	-12.40	FYN	6978863	-11.42
DDR2	5453814	-12.40	p38 β	2499603	-11.41
EphA8	7263928	-12.34	RiPK2	4506537	-11.34
PDGFR α	5453870	-12.28	MLK4	SK691	-11.22
YES	125870	-12.26	EphA2	125333	-11.20
BRK	5174647	-12.22	EphB6	4758292	-11.14
BLK	914204	-12.10	RSK4	zSK518b	-11.02
MAP2K5	1616781	-12.10	GAK	4885251	-10.96
QIK	SK513	-12.10	RET	5419753	-10.94
LYN	125480	-12.08	RSK3	zSK338b	-10.80
QSK	4589642	-12.06	TGF β R2	1827475	-10.76
FGR	125358	-12.02	BRAF	7304933	-10.68
EphA6	SK646	-12.00	CSK	729887	-10.62
HCK	2194103	-11.94	ACK	423137	-10.60
PDGFR β	4505683	-11.92	RAF1	525211	-10.58
LCK	8569365	-11.82	CaMKK1	SK697	-10.40
YANK2	7160989	-11.78	HER4/ErbB4	422854	-10.38
EphA3	125387	-11.78	BTk	575890	-10.32
SIK	SK604	-11.74	KDR	2655412	-10.28
EphA4	2833208	-11.72	FLT4	388522	-10.28
MOK	5139689	-11.62	KIT	1817734	-10.26

233. Known targets of the drug are indicated by shaded boxes. Acc. is the KSD accession number, or in the cases where kinase sequences were obtained from our own CLUSTALW alignment (see Methods), the accession number used by Manning et al. (denoted by the prefix "SK").

234. Table 8. Preliminary results for the application of the method to protein phosphatases. Four protein phosphatases were screened for activity with the inhibitor okadaic acid. Although the method failed to identify the phosphatases inhibited by okadaic acid when the simulations were run with a dielectric value of 78, PP1 and PP2A were correctly identified as strongly binding this inhibitor when the dielectric value was adjusted to 4.

Dielectric 78		
Phosphatase	Energy	IC50 (nM)
PP2B	-6.453	5000
PP1	-5.952	42
PP2A	-5.209	0.51
PP2C	-0.353	>>10,000

Dielectric 4		
Phosphatase	Energy	IC50 (nM)
PP1	-12.453	42
PP2A	-8.778	0.51
PP2C	-2.226	>>10,000
PP2B	-2.182	5000

i) Discussion

235. Validating the methods used herein, three new kinase targets of the drug: RiPK2 (RICK), GAK, and CK1 δ were discovered (4). In our computational screen of the same drug, RiPK2 and GAK ranked #37 and #42 respectively in the list of 493 computed binding energies (see Table 7), with energies that placed them inside the scaled binding energy cutoff (-10.27 kcal.mol) identified in the ~20 kinase calculations (Table 3). These two kinases can also be considered to be predicted targets of the drug.

236. Regarding the disagreement between prediction and experiment for quercetin and indirubin-3'-monoxime, the explanation for these discrepancies comes from work by Shoichet and co-workers (36) showing that certain kinase inhibitors can aggregate at concentrations used in experimental inhibition studies: the aggregates thus formed can sequester the kinase leading to an artifactual decrease in kinase activity. Crucially, both quercetin and indirubin exhibit this behavior; of the other drugs studied here only SB 203580 has so far been tested in this way and it was shown not to undergo significant aggregation (36). This suggests that the two apparent failures of the computational method are actually due to problems with the experimental data; support

for this idea comes from the seemingly random distribution of 'inhibited' kinases among the branches of the computed phylogenetic trees (Figure 5). These can be resolved by the use of computational estimates of binding affinities to identify suspect experimental data.

G. References

1. Hardman, J. G., Limbird, L. E., & Gilman, A. G. *The Pharmacological Basis of Therapeutics* McGraw-Hill, New York, NY (2001).
2. Tuveson, D. A., Willis, N. A., Jacks, T., Griffin, J. D., Singer, S., Fletcher, C. D. M., Fletcher, J. A., & Demetri, G. D. (2001) *Oncogene* 20, 5054-5058.
- 10 3. Zhu, H., & Snyder, M. (2003) *Curr. Opin. Chem. Biol.* 7, 55-63.
4. Godl, K., Wissing, J., Kurtenbach, A., Habenberger, P., Blencke, S., Gutbrod, H., Salassidis, K., Stein-Gerlach, M., Missio, A., Cotton, M., & Daub, H. (2003) *Proc. Natl. Acad. Sci. USA* 100, 15434-15439.
5. Dunbrack, R. L. (2002) *Curr. Opin. Struct. Biol.* 12, 431-440.
- 15 6. Manning, G., Whyte, D. B., Martinez, R., Hunter, T., & Sudarsanam, S. (2002) *Science* 298, 1912-1934.
7. Cohen, P. (2002) *Nat. Rev. Drug Discov.* 1, 309-315.
8. Woolfrey, J. R., Weston, G. S. (2002) *Curr. Pharm. Design* 8, 1527-1545.
9. Davies, S. P., Reddy, H., Caivano, M., & Cohen, P. (2000) *Biochem. J.* 351, 95-105.
- 20 10. Fox, T., Coll, J. T., Xie, X., Ford, P. J., Germann, U. A., Porter, M. D., Pazhanisamy, S., Fleming, M. A., Galullo, V., Su, M. S. S., & Wilson, K. P. (1998) *Protein Sci.* 7, 2249-2255.
11. Engh, R. A., Girod, A., Kinzel, V., Huber, R., & Bossemeyer, D. (1996) *J. Biol. Chem.* 271, 26157-26164.
- 25 12. Gray, N. S., Wodicka, L., Thunnissen, A. -M. W. H., Norman, T. C., Kwon, S., Espinoza, F. H., Morgan, D. O., Barnes, G., LeClerc, S., Meijer, L., Kim, S. -H., Lockhart, D. J., & Schultz, P. G. (1998) *Science* 281, 533-538.
13. Meijer, L., Thunnissen, A. -M. W. H., White, A. W., Garnier, M., Nikolic, M., Tsai, L. -H., Walter, J., Cleverley, K. E., Salinas, P. C., Wu, Y. -Z., Biernat, J., Mandelkow, E. -M.,
- 30 Kim, S. -H., & Pettit, G. R. (2000) *Chem. Biol.* 7, 51-63.
14. Nagar, B., Hantschel, O., Young, M. A., Scheffzek, K., Veach, D., Bornmann, V., Clarkson, B., Superti-Furga, G., & Kuriyan, J. (2003) *Cell* 112, 859-871.
15. Davies, T. G., Tunnah, P., Meijer, L., Marko, D., Eisenbrand, G., Endicott, J. A., & Noble, M. E. M. (2001) *Structure* 9, 389-397.
- 35 16. Sicheri, F., Moarefi, I., & Kuriyan, J. (1997) *Nature* 385, 602-609.
17. van Aalten, D. M. F., Bywater, R., Findlay, J. B. C., Hendlich, M., Hooft, R. W. W., & Vriend, G. (1996) *J. Comput. Aid. Mol. Design* 10, 255-262.
18. Bain, J., McLauchlan, H., Elliott, M., & Cohen, P. (2003) *Biochem. J.* 371, 199-204.
19. Druker, B. J., Tamura, S., Buchdunger, E., Ohno, S., Segal, G. M., Fanning, S.,
- 40 Zimmermann, J., & Lydon, N. B. (1996) *Nat. Med.* 2, 561-566.
20. Roskoski Jr., R. (2003) *Biochem. Biophys. Res. Commun.* 309, 709-717.
21. Buzko, O., & Shokat, K. M. (2002) *Bioinformatics* 18, 1274-1275.
22. Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994) *Nucleic Acids Res.* 22, 4673-4680.
- 45 23. Canutescu, A. A., Shelenkov, A. A., & Dunbrack, R. L. (2003) *Protein Sci.* 12, 2001-2014.
24. Vriend, G. (1990) *J. Mol. Graph.* 8, 52-52.

25. Allen, M. P., & Tildesley, D. J. *Computer Simulation of Liquids* Clarendon Press, Oxford Science Publications, Oxford, UK (1987).
26. Xiang, Z. X., & Honig, B. (2001) *J. Mol. Biol.* 311, 421-430.
27. Sitkoff, D., Sharp, K. A., & Honig, B. (1994) *J. Phys. Chem.* 98, 1978-1988.
- 5 28. Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953) *J. Chem. Phys.* 21, 1087-1093.
29. Wong, C. F., Hünenberger, P. H., Akamine, P., Narayana, N., Diller, T., McCammon, J. A., Taylor, S. S., & Xuong, N. -H. (2002) *J. Med. Chem.* 44, 1530-1539.
30. Shoichet, B. K., Leach, A. R., & Kuntz, I. D. (1999) *Proteins* 34, 4-16.
- 10 31. Chen, Y. Z., & Zhi, D. G. (2001) *Proteins* 43, 217-226.
32. Rockey, W. M., & Elcock, A. H. (2002) *Proteins* 48, 664-671.
33. Sims, P. A., Wong, C. F., & McCammon, J. A. (2003) *J. Med. Chem.* 46, 3314-3325.
34. De Moliner, E., Brown, N. R., & Johnson, L. N. (2003) *Eur. J. Biochem.* 270, 3174-3181.
- 15 35. Schindler, T., Bornmann, W., Pellicena, P., Miller, W. T., Clarkson, B., & Kuriyan, J. (2000) *Science* 289, 1938-1942.
36. McGovern, S. L., & Shoichet, B. K. (2003) *J. Med. Chem.* 46, 1478-1483.
37. Honkanen, R.E., Zwiller, J., Moore, R.E., Daily, S.L., Khatra, B.S., Dukelow, M., & Boynton, A.L. (1990) *J. Biol. Chem.* 265, 19401-19404.
- 20 38. Bialojan, C. & Takai, A. (1988) *Biochem. J.* 256, 283-290.

V. CLAIMS

What is claimed is:

1. A method of identifying a target for a molecule comprising the steps: a) modeling the molecule in complex with a known target for the molecule, b) obtaining potential target molecules by selecting potential target molecules with a defined homology to the known target, c) determining the binding affinity of a potential target with the molecule by modeling the potential target with the molecule, wherein side chain rotamers are sampled from a rotamer library during homology modeling.
2. A method of identifying a target for a molecule comprising the steps: a) obtaining a structural model of the molecule and a known target, wherein the known target comprises a known target-molecule binding domain, b) obtaining a potential target by identifying potential targets having a defined homology with the known target, c) performing homology modeling with the identified potential target, wherein during the homology modeling backbone conformations are held identical to the known target, wherein sidechains are sampled from a library of rotamers, and d) calculating a binding energy of the molecule and the identified potential target.
3. The method of claim 2, wherein the homology modeling is performed with only those positions in the potential target that have a cognate position in the known target.
4. The method of claim 2, wherein all residues outside of the known target-molecule binding domain are left out of the modeling process.
5. The method of claim 2, wherein the binding energy is calculated such that positions of the residues close to the molecule in the homology model are sampled using a rotamer library.
6. The method of claim 5, wherein the sampling occurs using Monte Carlo methods.
7. The method of claim 5, wherein the residues within 5 Å of the molecule are sampled.
8. The method of claim 2, wherein the binding energy is calculated by obtaining the total energy of the potential target and molecule system, and wherein the total energy is the sum of the electrostatic and van der Waals forces.
9. The method of claim 8, wherein the electrostatic interaction is calculated by a Debye-Huckel model.
10. The method of claim 9, wherein the Debye Huckel model is calculated using the formula $E_{elec} = 332.08 q_1 q_2 \exp(-\kappa r) / \epsilon r$.
11. The method of claim 8, wherein the van der Waals forces are calculated using the formula $E_{vdw} = C \{ \sigma_{att}^{12} / r^{12} - \sigma_{att}^6 / r^6 \}$.
12. The method of claim 8, wherein the van der Waals forces are calculated using the formula $E_{vdw} = C \{ \sigma_{rep}^{12} / r^{12} \}$.
13. The method of claim 8, wherein the binding energy is calculated by obtaining the difference between the energy of the potential target after optimization and the energy of the potential target complexed

with the molecule after optimization.

14. The method of claim 8, further comprising selecting a potential target as a candidate target when the binding energy of the potential target complexed with the target is equal to or less than the binding energy of the known target complexed with the molecule.

5 15. A method of identifying a desired protein-molecule interaction comprising: (a) determining structural information for a protein known to interact with the molecule of interest; (b) identifying which residues of the protein of step a) interact with the molecule; (c) comparing the residues identified in step b) with a database of proteins; (d) selecting proteins having an area of similarity to the residues identified in step b); (e) calculating interaction energies between the proteins of step d) and the molecule of interest; and (f) determining which proteins are capable of interacting in a desired fashion with the molecule of interest.

10 16. The method of claim 15, wherein the interaction energies are calculated using any one or more of the following: sidechain conformations, electrostatic forces, and van der Waals forces.

17. The method of claim 15, wherein in step c) all possible sidechain conformations are calculated.

15 18. The method of claim 15, wherein the protein of step a) is a receptor.

19. The method of claim 15, wherein the proteins of step b) are receptors.

20. The method of claim 15, wherein the molecule is a drug.

21. The method of claim 15, wherein the molecule-protein interaction comprises hydrogen bonding.

22. The method of claim 15, wherein structural information of the protein of step a) is known.

20 23. The method of claim 15, wherein the structural information of the protein of step a) was obtained from a crystal structure.

24. The method of claim 16, wherein the structural information of the protein of step a) was obtained from a solution structure.

25 25. The method of claim 15, wherein in step b) the residues are compared only for a given region of the proteins.

26. The method of claim 25, wherein the given region is a drug binding site.

27. The method of claim 16, wherein calculating the interaction energy includes calculating electrostatic interaction.

28. The method of claim 27, wherein the electrostatic interaction is calculated by Debye-Huckel model.

30 29. The method of claim 28, wherein the Debye Huckel model is calculated by using the formula $E_{elec}=332.08 q_1 q_2 \exp(-\kappa r)/\epsilon r$.

30. The method of claim 16, wherein calculating the interaction energy includes calculating van der

Waals forces.

31. The method of claim 30, wherein for interacting pairs in which both atoms are carbon or sulfur, the van der Waals forces are calculated by using the formula $E_{vdw} = \epsilon \{ \sigma_{att}^{12}/r^{12} - \sigma_{att}^6/r^6 \}$.
32. The method of claim 30, wherein for all interacting pairs except those where both atoms are carbon or sulfur, the van der Waals forces are calculated by using the formula $E_{vdw} = \epsilon \{ \sigma_{rep}^{12}/r^{12} \}$.
33. The method of claim 15, wherein a Metropolis test is applied to compute molecule-protein interaction.
34. The method of claim 15, wherein a Monte Carlo function is used for sampling of side chain rotamers when calculating interaction energies.
35. A computer system having a processing means, memory means, and a visual display means, the memory means containing sequence information for a protein known to interact with a molecule of interest, and modules containing information to be compared with the sequence information of the protein known to interact with the molecule of interest, and the processing means being operable to compute molecule-protein binding energy using the method of claim 52.
36. A method of making a pharmaceutical comprising a) modeling the pharmaceutical in complex with a known target for the molecule; b) obtaining potential target molecules by selecting potential target molecules with a defined homology to the known target, c) determining the binding affinity of a potential target with the pharmaceutical by modeling the potential target with the pharmaceutical, wherein a Monte Carlo function is used for sampling of side chain rotamers during homology modeling; d) identifying target molecules of the pharmaceutical; e) synthesizing the pharmaceutical; and f) testing the pharmaceutical for binding to the target molecule.
37. A method of inhibiting a receptor selected from the group consisting of CRK7, SYK, TEC, RET, BMX, ABL, IKKb, TRKA, HER4/ErbB, SgK288, DDR2, TRKB, ARG, TRKC, DDR1, TIE1, FMS, YES, ACK, FGFR2, BLK, FRK, FYN, CSK, ANKRD3, HCK, MST1, SRC, LYN, IKKa, FGR, TXK, NDR2, FLT1, LCK, NDR1, PDGFRa, PDGFRb, FLT4, MST2, and KIT comprising incubating the receptor with the drug imatinib.
38. A method of inhibiting a receptor selected from the group consisting of MUSK, FLT3, CDK2, DYRK1A, FLT4, CDK3, CDC2, KDR, CDKL1, TRKB, CASK, MAK, TRKC, DYRK1B, CDK5, ROR1, PCTAIRE2, TRKA, PCTAIRE1, FLT1, TIE1, RET, CDK7, CDKL3, PDGFRa, ROR2, JAK2, AurA, FGFR2, PCTAIRE3, CDC7, CDKL2, AurC, AurB, GCN2, TLK2, TLK1, CDK9, TIE2, MAP3K4, KIT, MSK1, and PLK2 comprising incubating the receptor with the drug purvanol.
39. A method of inhibiting a receptor selected from the group consisting of FRK, DDR1, BRK, QIK, EphA1, EphB2, DDR2, QSK, EphB3, SRM, EphB1, ACK, EphA6, SIK, MOK, YANK2, EphB4, EphA8, HER4/erbB4, RET, YANK1, YANK3, EphA4, EphA5, GAK, EDRF, PDGFRa, PDGFRb, EphA2, FGR, RIPK3, YES, MLK4, EphA3, LCK, HER2/ErbB2, SRC, BLK, BTK, FYN, LYN, HCK, RIPK2, CSK, ARG, TXK, Domain2_RSK4, p38 α , p38 β , and CaMKK1 comprising incubating the

receptor with the drug SB 203580.

40. The method of claim 37, 38, and 39 further comprising identifying a subject in need of inhibition of the receptor.

41. The method of claim 37, 38, and 39 further comprising identifying a subject having a disease associated with the receptor.

42. A method of characterizing a protein-molecule interaction using the method of claim 1.

43. A method of displaying a representation of a protein-molecule interaction on a computer having a processing means, a memory means, an input means and an output means comprising: receiving the three-dimensional coordinates of atoms of the protein; producing a representation of the protein based upon the received coordinates; and displaying the representation of the protein-molecule interaction on the visual display means, wherein the protein in the protein-molecule interaction comprises a protein which is a homologue of the native target of the molecule.

44. An apparatus comprising:

(a) a system data store capable of storing coordinate sets; and

(b) a system processor in communication with the system data store that carries out the following steps:

(i) modeling a molecule in complex with a known target for the molecule,

(ii) obtaining potential target molecules by selecting potential target molecules with a defined homology to the known target,

(iii) determining the binding affinity of a potential target with the molecule by modeling the potential target with the molecule,

wherein a Monte Carlo function is used for sampling of side chain rotamers during homology modeling.

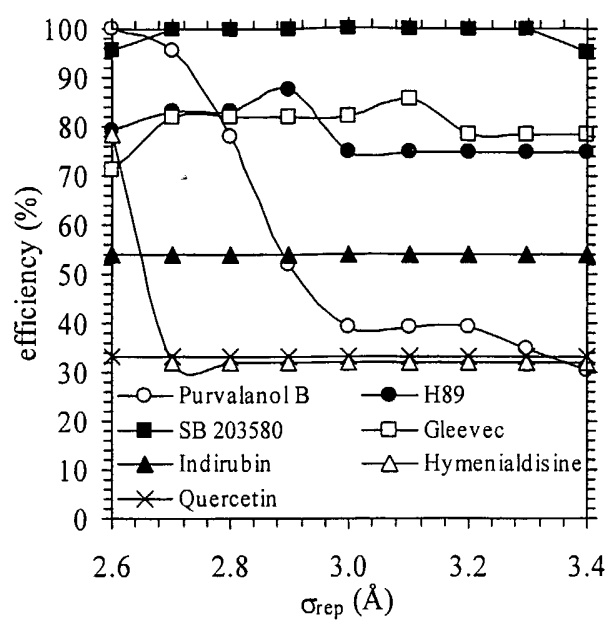


FIG. 1

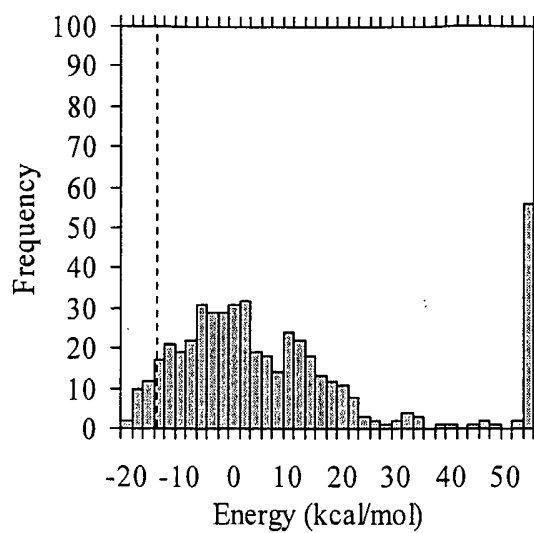


FIG. 2

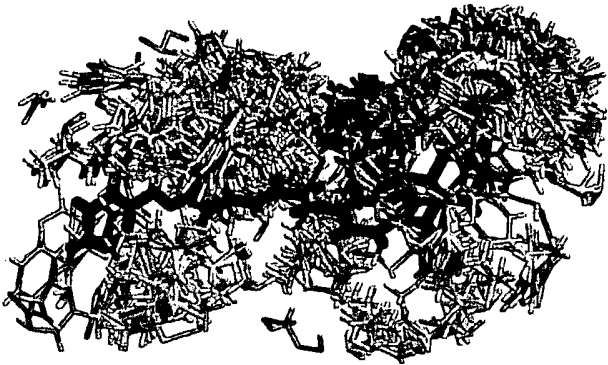


FIG. 3

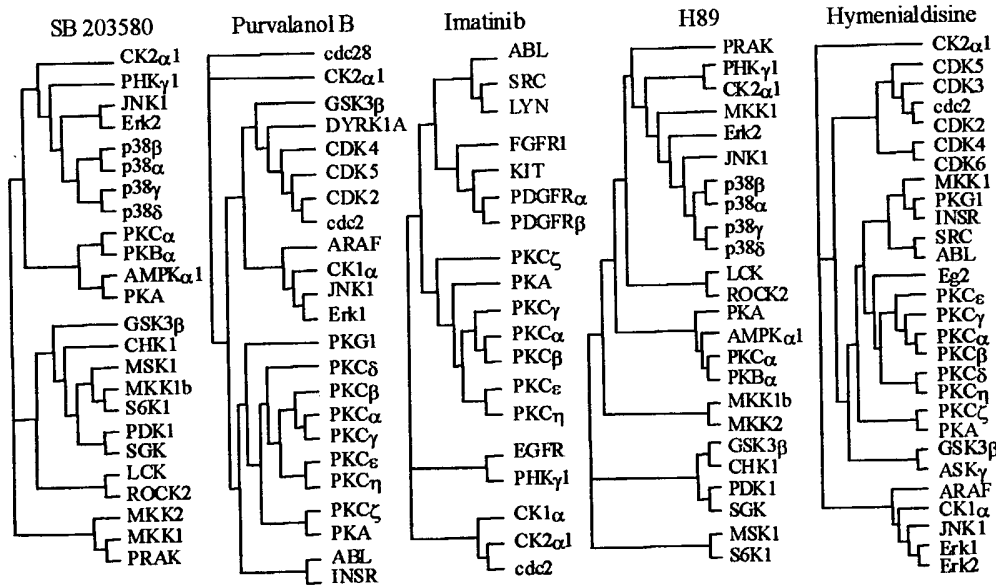
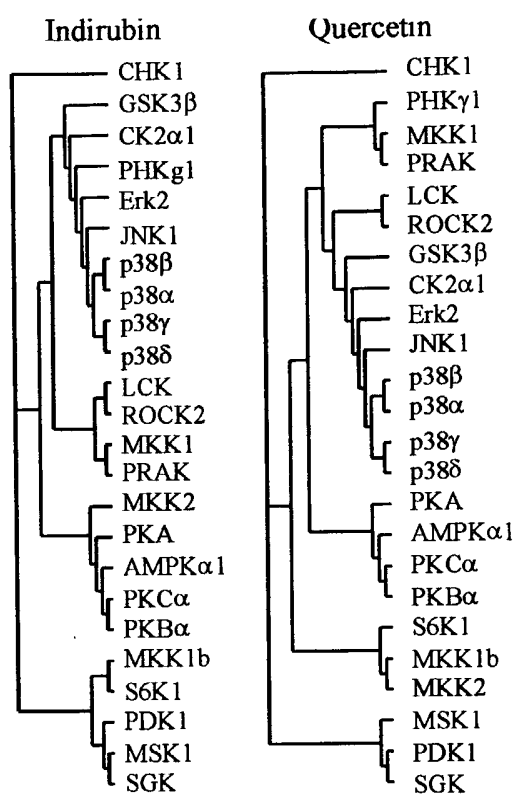
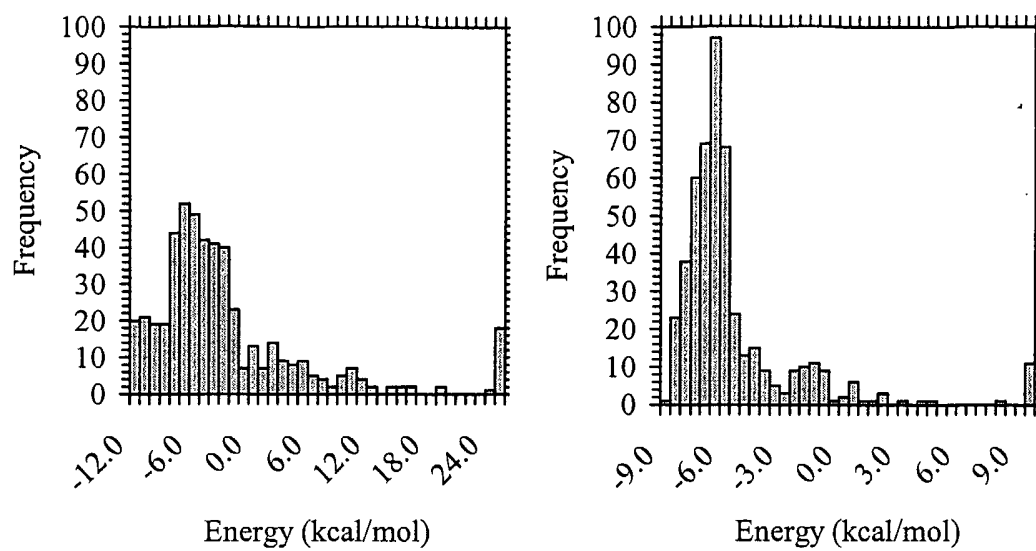


FIG. 4.

**FIG. 5**

**FIG. 6**

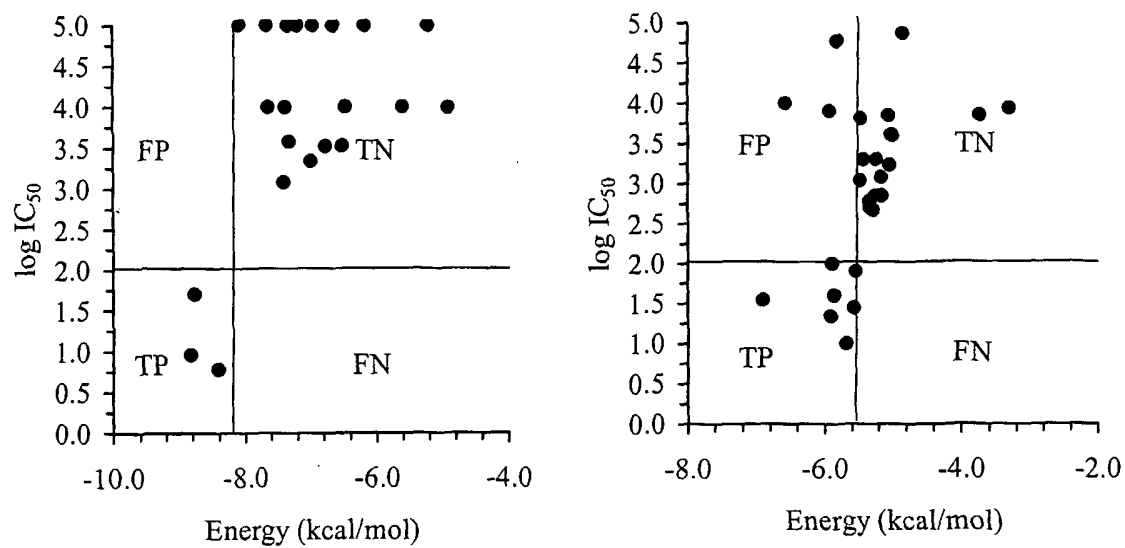


FIG. 7

<i>Kinase</i>	ΔE	<i>Kinase</i>	ΔE	<i>Kinase</i>	ΔE	
Kinase D	-20	Kinase A	-20	Kinase A	-20	TP
Kinase A	-19	Kinase B	-19	Kinase B	-19	
Kinase E	-18	Kinase D	-18	Kinase C	-18	
Kinase B	-17	Kinase C	-17	Kinase D	-17	TN
Kinase F	-16	Kinase E	-16	Kinase E	-16	
Kinase C	-15	Kinase F	-15	Kinase F	-15	
2.6 Å		2.7 Å		2.8 Å		

FIG. 8

8/26

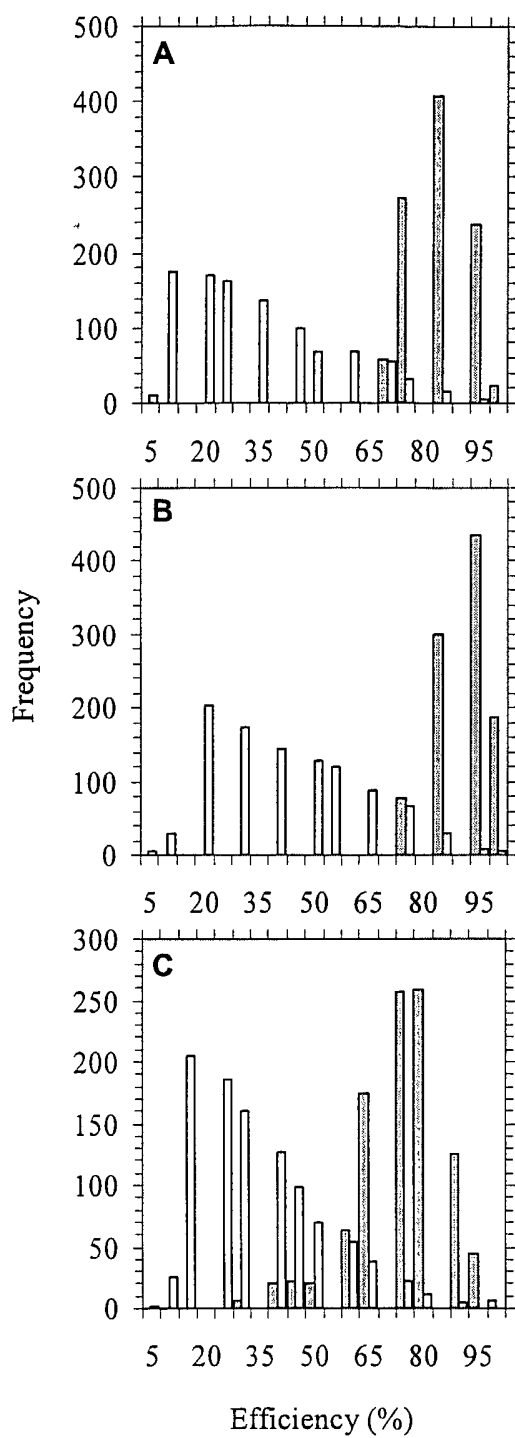
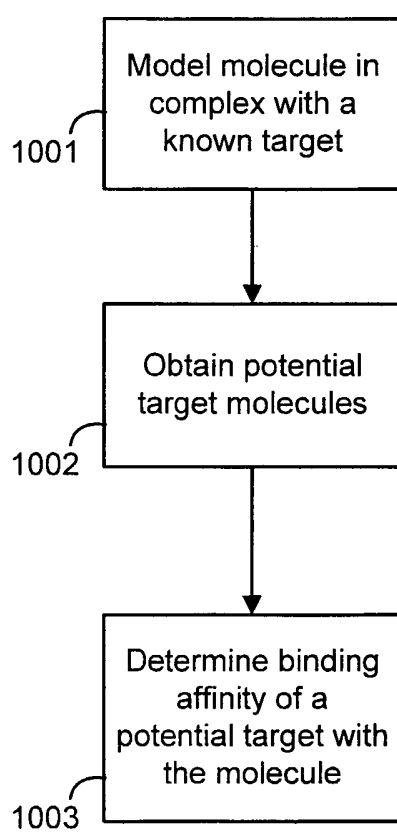


FIG. 9

FIG. 10



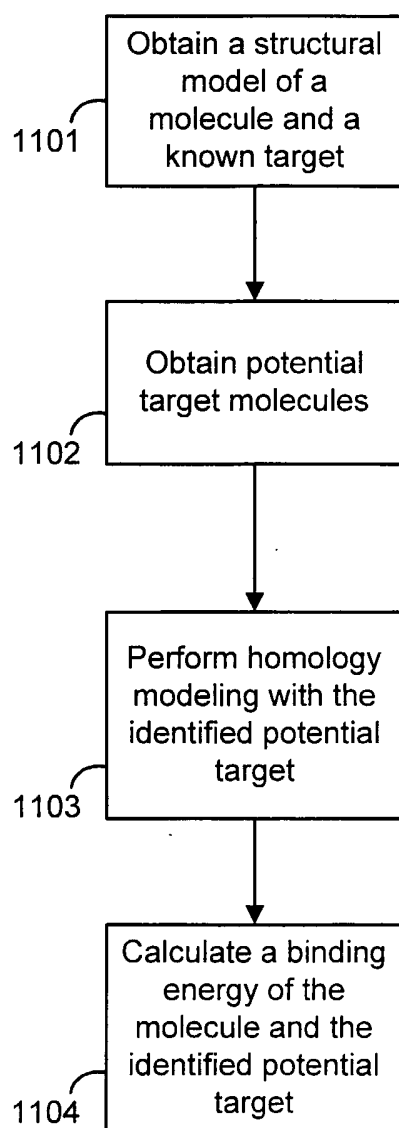


FIG. 11

11/26

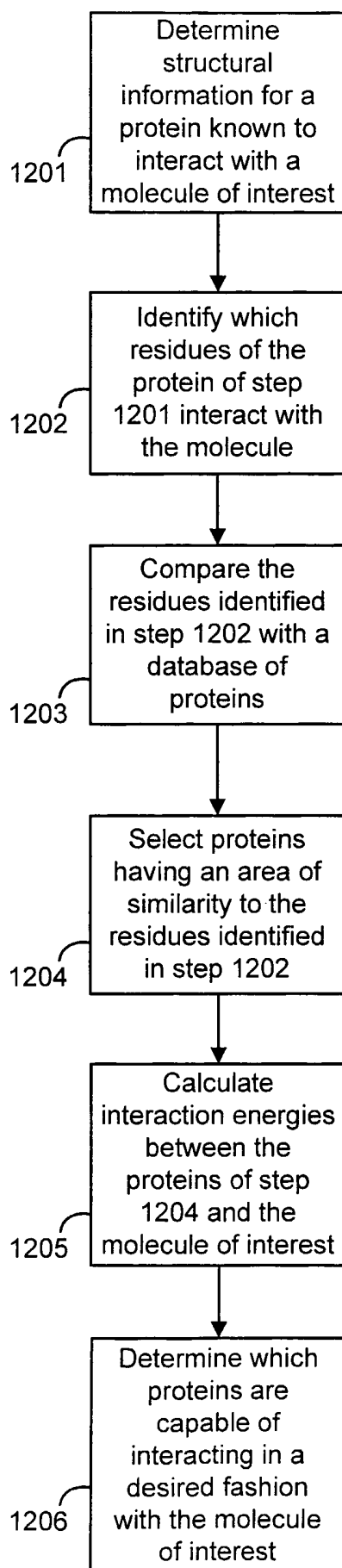


FIG. 12

12/26

FIG. 13

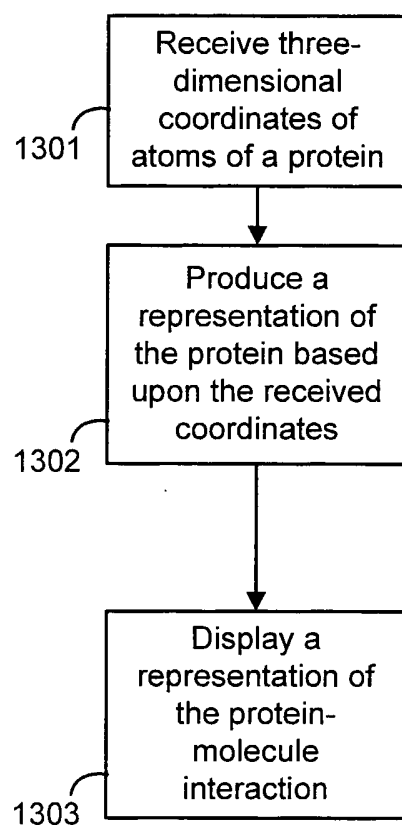
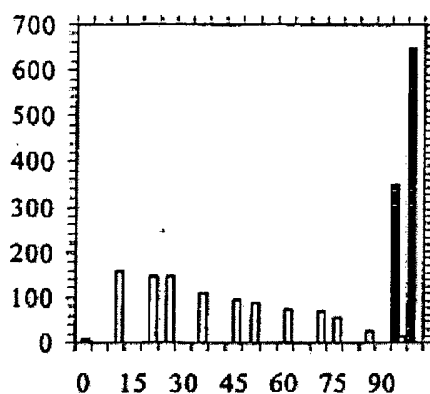
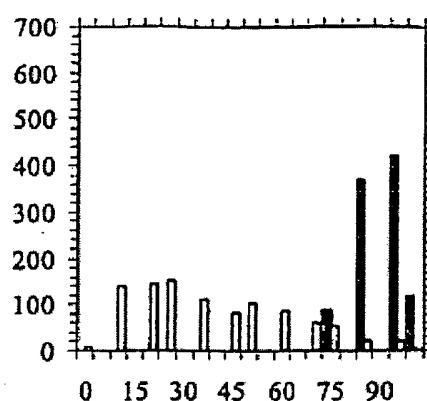
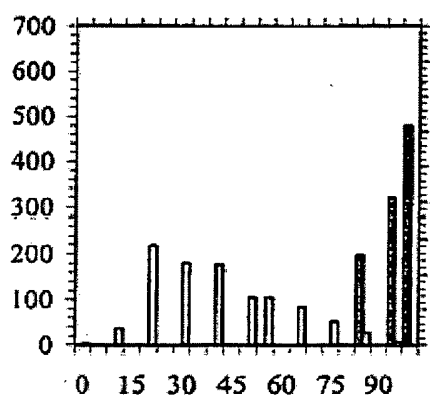


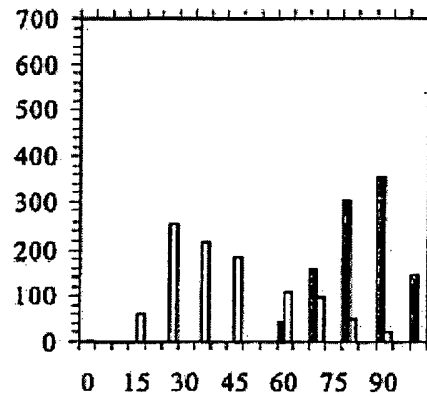
FIG. 14

SB203580^{1PME}SB203580^{1A9U}

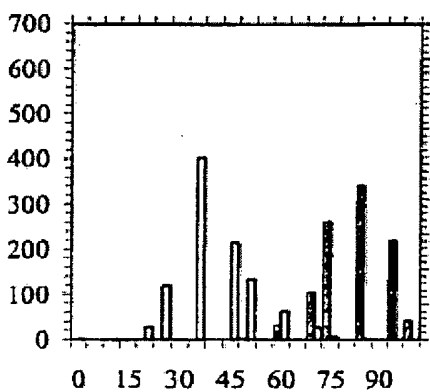
Purvalanol B



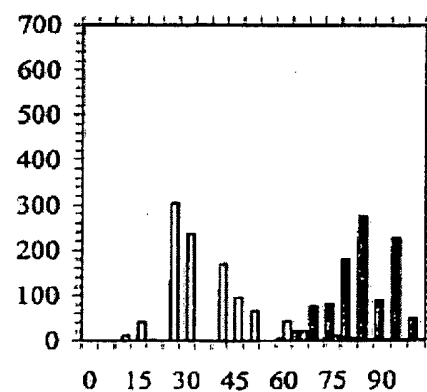
Imatinib



H89

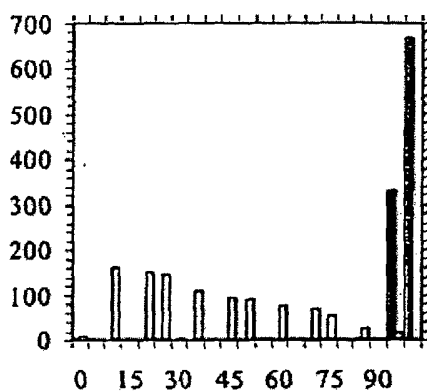
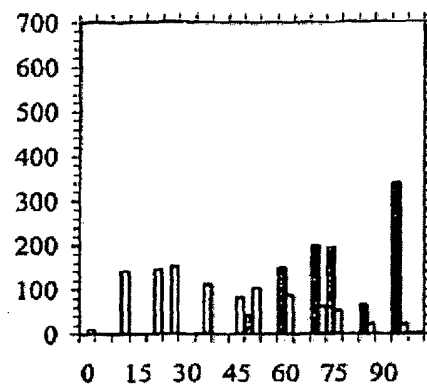


Hymenialdisine

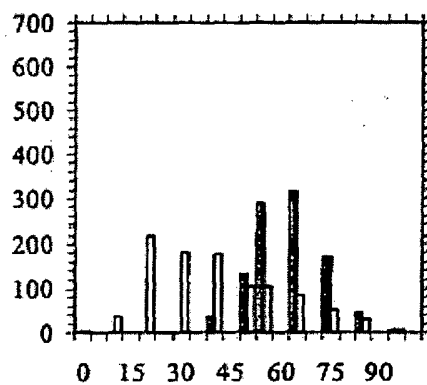


14/26

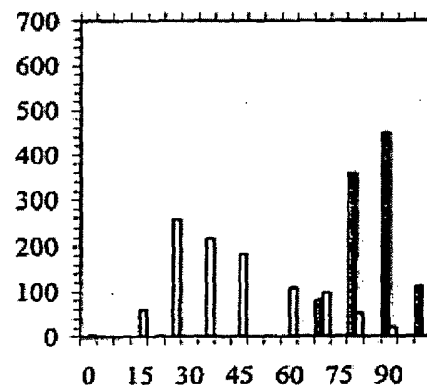
FIG. 15

SB203580^{1PME}SB203580^{1A9U}

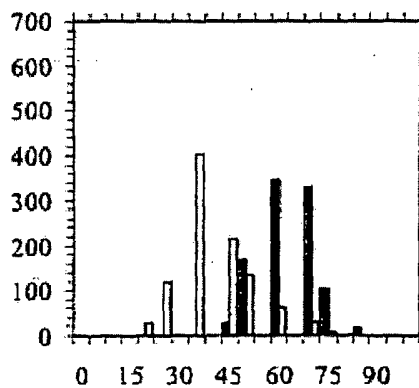
Purvalanol B



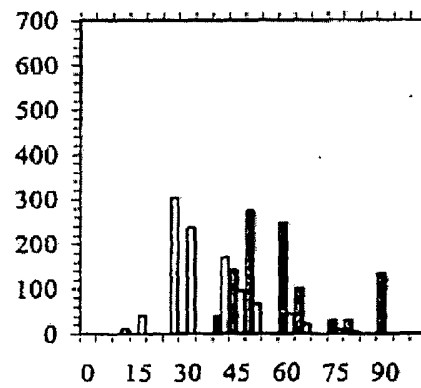
Imatinib

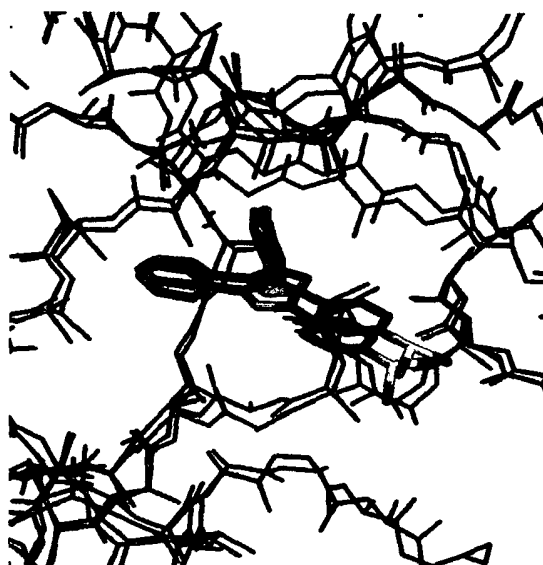


H89



Hymenialdisine

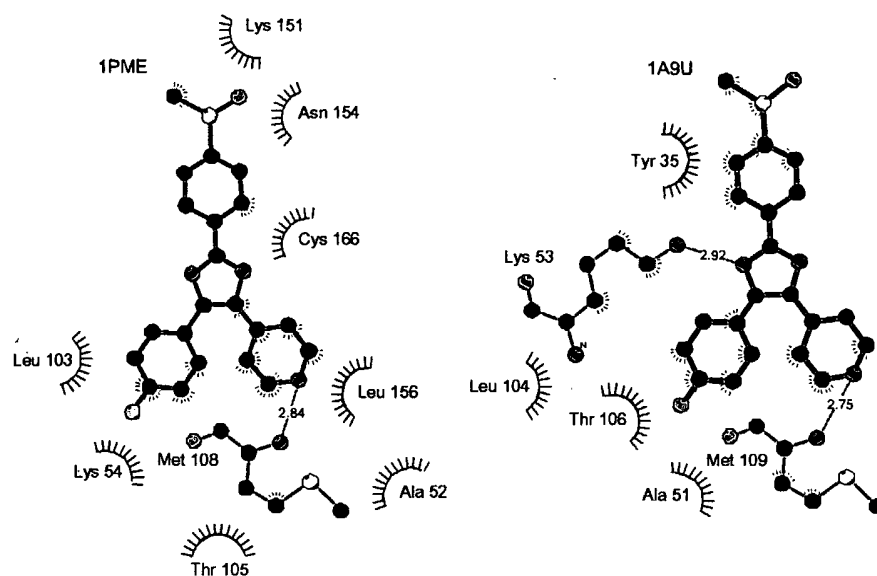




Superimposed views of SB203580 in a model of the p38 α binding site before and after GROMACS energy minimization. Structures prior to minimization are shown in red, those after minimization are shown in green. The kinase sidechains have been removed for clarity.

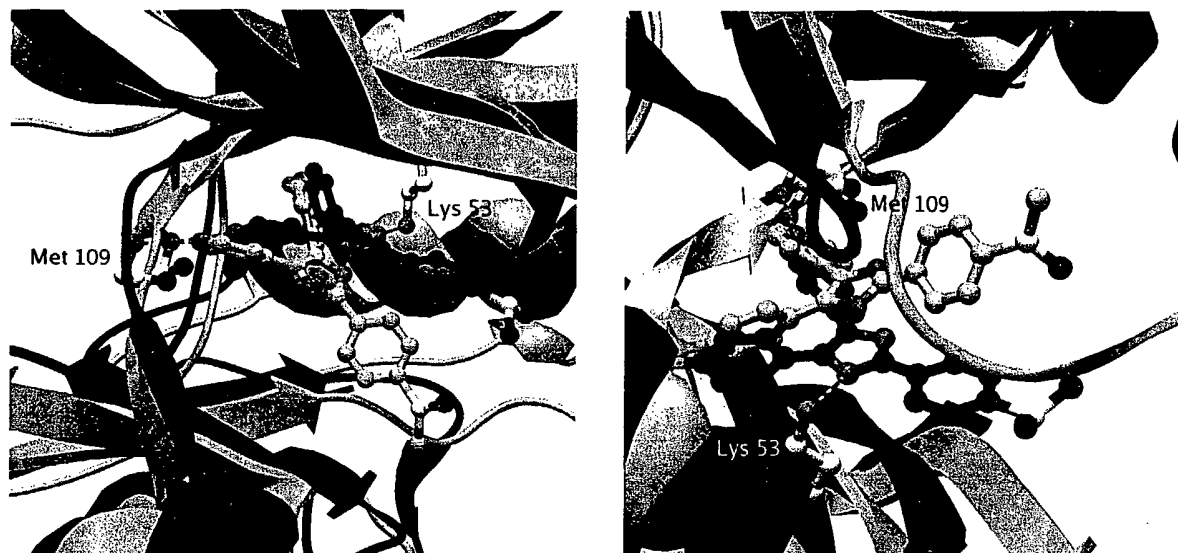
FIG. 16

16/26



A comparison of the inhibitor-kinase contacts made by SB203580 and the crystal structures 1PME (mutant Erk2) and 1A9U (p38 α). Leu103/104, Lys 53/54, Thr 105/06, Ala 51/52, and Met 108/109 interact with the inhibitor in both structures. This figure was generated by the program LIGPLOT (Wallace, A. C.; Laskowski, R. A.; Thornton, J. M. LIGPLOT: A program to generate schematic diagrams of protein-ligand interactions. *Prot. Eng.* 1995, 8, 127-134).

FIG. 17



*A superimposed view of SB203580 in the binding sites of the crystal structures 1PME and 1A9U. Portions of the inhibitor making contacts important for affinity overlay well. The structure of the kinase in 1PME is shown in red and its bound SB203580 is shown in cyan; the structure of the kinase in 1A9U is shown in yellow and its bound inhibitor is shown in purple. The residues Lys 53 and MET 109 are shown making hydrogen-bonds to the inhibitor in dashed green lines (1A9U numbering). Residues ILE 31 and GLY 32 in the 1PME structure have been removed from the figure for clarity. An all-atom alignment was performed using residues 61-79, 104-116 (1A9U numbering) and SB203580. The superimposition and figure were prepared using Swiss-Pdb Viewer (<http://www.expasy.org/spdbv>; Guex, N.; Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling. *Electrophoresis* 1997, 18, 2714-2723).*

FIG. 18

An ordered list of the computed binding energies obtained with SCR for five inhibitors for which the calculations were successful. Results shown are those obtained when σ_{rep} was set to 2.8Å and σ_{att} to 3.55Å in all calculations.

SB203580 ^{1PME}		SB203580 ^{1A9U}		Purvalanol B		Imatinib		H89		Hymenialdisine	
Kinase	ΔE	Kinase	ΔE	Kinase	ΔE			Kinase	ΔE	Kinase	ΔE
LCK	-11.82	Erk2	-14.88	CDK2	-9.87	ABL	-20.33	ROCK2	-13.41	cdc2	-9.01
p38 α	-11.57	ROCK2	-14.64	ABL	-9.39	SRC	-19.97	MSK1	-12.90	CDK2	-9.00
p38 β	-11.41	p38 α	-12.30	cdc2	-9.25	LYN	-19.75	PKA	-12.80	CDK3	-8.99
Erk2	-9.24	p38 β	-12.13	CDK5	-9.15	PDGFR α	-19.65	LCK	-12.55	CDK5	-8.79
PRAK	-8.54	LCK	-11.79	cdc28	-8.33	PDGFR β	-19.48	MKK1	-12.34	CDK4	-8.72
MKK1b	-8.28	PKB α	-10.93	CDK4	-7.96	KIT	-19.41	SGK	-12.19	CDK6	-8.65
S6K1	-8.26	PKC α	-10.62	GSK3 β	-7.81	FGFR1	-18.05	CHK1	-12.18	PKC δ	-6.56
PKB α	-7.96	AMPK α 1	-10.58	PKC δ	-7.78	PKC α	-14.99	PKB α	-12.13	PKC η	-6.49
PKA	-7.70	PKA	-9.54	DYRK1A	-7.56	PKC β	-14.54	GSK3 β	-11.79	GSK3 β	-6.47
ROCK2	-7.59	PHK γ 1	-9.07	PKC α	-7.44	PKC η	-14.24	AMPK α 1	-11.72	ASK γ	-6.45
MKK2	-6.97	MKK1	-8.48	PKC ϵ	-7.42	PKC ϵ	-12.37	PKC α	-11.49	PKC ϵ	-6.44
PHK γ 1	-6.57	PRAK	-7.93	PKC β	-7.42	PKA	-10.95	S6K1	-11.47	INSR	-6.20
AMPK α 1	-4.79	CHK1	-7.38	INSR	-7.39	EGFR	-9.92	MKK1b	-11.23	PKC β	-6.08
PKC α	-4.50	GSK3 β	-6.47	Erk1	-6.94	PKC γ	-5.06	Erk2	-11.07	PKC ζ	-6.02
GSK3 β	-4.21	MKK1b	-5.81	PKC ζ	-6.75	PHK γ 1	-0.88	MKK2	-10.07	PKC α	-5.96
CHK1	-3.99	CK2 α 1	-5.77	PKA	-6.60	cdc2	0.34	PRAK	-9.90	PKA	-5.67
JNK1	-3.22	S6K1	-5.74	PKC η	-6.41	CK2 α 1	7.15	p38 γ	-1.23	ABL	-5.53
MKK1	-3.20	MSK1	-5.66	PKG1	-5.36	PKC ζ	179.15	p38 δ	-0.96	PKC γ	-5.49
p38 δ	-2.90	SGK	-4.72	CK1 α	-5.14	CK1 α	320.03	p38 β	-0.72	SRC	-5.44
p38 γ	-2.66	p38 γ	0.57	PKC γ	-4.94			p38 α	-0.60	Erk2	-4.95
SGK	1.76	MKK2	1.64	JNK1	-4.73			JNK1	3.26	Erk1	-4.90
MSK1	2.04	JNK1	2.52	ARAF	-4.60			PHK γ 1	29.93	PKG1	-4.79
PDK1	28.42	p38 δ	5.46	CK2 α 1	5.15			CK2 α 1	42.05	Eg2	-3.82
CK2 α 1	38.97	PDK1	11.54					PDK1	742.32	MKK1	-2.78
										ARAF	-1.86
										CK1 α	-0.08
										CK2 α 1	15.14
										JNK1	20.07

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes.

The precision of the computed binding energies is ~0.03 kcal/mol.

FIG. 19

19/26

An ordered list of the docking pose with the most favorable binding energy obtained from docking calculations conducted with AutoDock. Energies shown are those of the docking pose with the most favorable binding energy (see Methods).

SB203580 1 ^{AE}			SB203580 1 ^{AE}			Purvalanol B			Imatinib			H89			Hymenialdisine		
Kinase	RMSD	ΔE	Kinase	RMSD	ΔE	Kinase	RMSD	ΔE	Kinase	RMSD	ΔE	Kinase	RMSD	ΔE	Kinase	RMSD	ΔE
MKK1	3.681	-14.63	p38 β	7.417	-14.87	PKC δ	8.437	-24.81	SRC	2.304	-23.8	PKA	0.671	-12.52	CK2 α 1	0.789	-14.09
PRAK	0.973	-13.62	AMPK α 1	7.114	-14.87	PKC β	6.372	-23.93	LYN	2.303	-23.73	SGK1	0.884	-12.41	GSK3 β	1.281	-13.83
S6K1	1.217	-13.61	GSK3 β	5.988	-14.5	PKC α	6.377	-23.75	PKC η	1.395	-23.37	MSK1	0.841	-12.41	ASK1	1.302	-13.78
Erk2	0.879	-13.59	p38 δ	5.998	-14.48	PKC γ	6.452	-23.41	PKC ϵ	1.258	-23.28	MKK1b	0.841	-12.39	CK1 α	0.64	-13.65
MKK2	1.016	-13.57	MSK1	6.385	-14.45	PKC ζ	6.094	-22.56	CK1 α	1.971	-23.17	SGK1	0.811	-12.33	CDK2	0.781	-13.56
PDK1	6.506	-13.5	S6K1	6.146	-14.32	Erk1	6.503	-22.53	PKA	1.064	-23.12	PKB α	0.874	-12.29	cdk2	0.783	-13.55
JNK1	1.011	-13.41	PDK1	7.605	-14.28	ARAF	3.588	-22.53	ABL	1.253	-23.08	p38 γ	1.101	-12.10	CDK3	0.779	-13.54
PKB α	0.984	-13.34	SGK	6.123	-14.14	CK2 α 1	1.633	-22.33	PDGFR α	1.906	-22.97	p38 δ	1.099	-12.09	CDK5	0.785	-13.49
PKA	0.981	-13.31	PKC α	6.247	-14.1	ABL	6.373	-22.27	PDGFR β	1.862	-22.86	PKC α	1.111	-11.98	PKC γ	1.343	-13.46
SGK	1.187	-13.28	PKA	5.977	-14.04	INSR	5.247	-22.15	KIT	1.884	-22.32	ROCK2	1.119	-11.97	CDK4	0.767	-13.44
MSK1	1.078	-13.26	p38 γ	5.697	-14.04	CK1 α	3.618	-22.15	FGFR1	1.932	-22.52	MKK2	0.700	-11.95	CDK6	0.764	-13.43
MKK1b	1.072	-13.25	ROCK2	5.622	-13.86	DYRK1A	5.332	-22.14	EGFR	1.329	-22.49	MKK1	0.518	-11.92	PKA	1.354	-13.3
GSK3 β	1.192	-13.16	CHK1	5.968	-13.76	JNK1	4.145	-22.11	PKC α	1.088	-22.39	LCK	0.947	-11.84	MKK1	1.543	-13.17
CHK1	1.072	-13.15	MKK1b	6.184	-13.75	cdk2	2.639	-21.9	CK2 α 1	2.291	-22.35	p38 β	1.048	-11.83	Eg2	1.294	-13.15
p38 α	3.62	-13.09	p38 α	5.788	-13.73	CDK5	4.687	-21.8	PKC β	1.05	-22.28	CHK1	0.847	-11.83	PKC α	1.34	-13.07
PKC α	0.922	-13.04	PKB α	5.534	-13.68	PKC η	7.232	-21.78	cdk2	0.873	-22.21	AMPK α 1	0.740	-11.75	PKC β	1.34	-13.07
AMPK α 1	0.945	-13.03	JNK1	9.299	-13.56	PKC ϵ	7.5	-21.77	PKC ζ	2.605	-21.62	p38 α	1.048	-11.75	PKC η	7.895	-13.01
ROCK2	1.117	-13	CK2 α 1	5.485	-13.42	PKG1	4.195	-21.47	PKC γ	2.116	-21.19	JNK1	1.100	-11.63	PKG1	7.998	-12.98
p38 β	1.004	-12.92	MKK2	5.761	-13.18	GSK3 β	1.657	-21.44	PHK γ 1	1.141	-20.42	PRAK	0.575	-11.61	Erk1	3.013	-12.87
p38 δ	3.711	-12.92	Erk2	5.34	-13.04	CDK4	4.513	-21.43				PHK γ 1	1.051	-11.52	Erk2	2.987	-12.86
p38 γ	1.868	-12.74	EGK	6.059	-12.63	cdk28	1.682	-21.19				GSK3 β	1.555	-11.27	INSR	8.683	-12.84
LCK	0.88	-12.64	PHK γ 1	6.672	-12.3	CDK2	1.677	-20.99				Erk2	4.612	-11.20	ABL	1.316	-12.83
PHK γ 1	5.193	-12.52	PRAK	1.939	-11.81	PKA	5.031	-20.73				CK2 α 1	0.909	-11.20	SRC	0.799	-12.82
CK2 α 1	5.751	-12.28	MKK1	4.619	-11.48							PDK1	4.727	-10.68	JNK1	1.572	-12.8
															PKC δ	0.791	-12.79
															PKC ζ	1.59	-12.78
															PKC ϵ	0.789	-12.76
															ARAF	4.305	-12.49

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes.

FIG. 20

For the inhibitors that have experimental data in the form of percentage activity (see Methods), we consider here only the kinases that have activities of <10% to be targets, rather than the 50% cutoff used in our other tables. The only inhibitor panel to be seriously affected is H89, whose interaction energy with S6K1 is significantly lower than those with its other targets. S6K1 could be considered a "false negative" in this case.

SB 203580 ^{1PME}		SB 203580 ^{1A9U}		H89	
Kinase	ΔE	Kinase	ΔE	Kinase	ΔE
LCK	-11.82	Erk2	-14.88	ROCK2	-13.41
p38 α	-11.57	ROCK2	-14.64	MSK1	-12.90
p38 β	-11.41	p38 α	-12.30	PKA	-12.80
Erk2	-9.24	p38 β	-12.18	LCK	-12.55
PRAK	-8.54	LCK	-11.79	MKK1	-12.34
MKK1b	-8.28	PKB α	-10.93	SGK	-12.19
S6K1	-8.26	PKC α	-10.62	CHK1	-12.18
PKB α	-7.96	AMPK α 1	-10.58	PKB α	-12.13
PKA	-7.70	PKA	-9.54	GSK3 β	-11.79
ROCK2	-7.59	PHK γ 1	-9.07	AMPK α 1	-11.72
MKK2	-6.97	MKK1	-8.48	PKC α	-11.49
PHK γ 1	-6.57	PRAK	-7.93	S6K1	-11.47
AMPK α 1	-4.79	CHK1	-7.38	MKK1b	-11.23
PKC α	-4.50	GSK3 β	-6.47	Erk2	-11.07
GSK3 β	-4.21	MKK1b	-5.81	MKK2	-10.07
CHK1	-3.99	CK2 α 1	-5.77	PRAK	-9.90
JNK1	-3.22	S6K1	-5.74	p38 γ	-1.23
MKK1	-3.20	MSK1	-5.66	p38 δ	-0.96
p38 δ	-2.90	SGK	-4.72	p38 β	-0.72
p38 γ	-2.66	p38 γ	0.57	p38 α	-0.60
SGK	1.76	MKK2	1.64	JNK1	3.26
MSK1	2.04	JNK1	2.52	PHK γ 1	29.93
PDK1	28.42	p38 δ	5.46	CK2 α 1	42.05
CK2 α 1	38.97	PDK1	11.54	PDK1	742.32

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes (percent activity <10%)

FIG. 21

An ordered list of the computed binding energies obtained with SCR for five inhibitors for which the calculations were successful. The parameterized values of σ_{rep} used to obtain these results for each inhibitor/kinase panel were SB203580 (1PME & 1A9U): 2.8Å; purvalanol B: 2.6Å; imatinib: 2.9Å; H89: 2.9Å; and hymenialdisine: 2.4Å.

SB203580 ^{1PME}		SB203580 ^{1A9U}		Purvalanol B		Imatinib		H89		Hymenialdisine	
Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE
LCK	-11.82	Erk2	-14.88	CDK2	-8.83	ABL	-17.91	ROCK2	-14.09	CK1 α	-6.90
p38 α	-11.57	ROCK2	-14.64	cdc2	-8.77	SRC	-17.90	CHK1	-13.08	ARAF	-6.57
p38 β	-11.41	p38 α	-12.30	CDK5	-8.42	LYN	-17.90	LCK	-12.85	PKA	-5.92
Erk2	-9.24	p38 β	-12.13	DYRK1A	-8.41	PDGFR α	-17.07	MSK1	-12.69	cdc2	-5.91
PRAK	-8.54	LCK	-11.79	ABL	-8.08	PDGFR β	-16.97	PKA	-12.46	CDK3	-5.89
MKK1b	-8.28	PKB α	-10.93	PKC ζ	-7.66	KIT	-16.78	MKK1	-12.31	CDK2	-5.86
S6K1	-8.26	PKC α	-10.62	GSK3 β	-7.64	FGFR1	-14.80	AMPK α 1	-11.69	PKC ζ	-5.81
PKB α	-7.96	AMPK α 1	-10.58	cdc28	-7.42	PKC α	-9.44	SGK	-11.55	GSK3 β	-5.69
PKA	-7.70	PKA	-9.54	CDK4	-7.39	PKC β	-8.86	GSK3 β	-11.29	CDK5	-5.56
ROCK2	-7.59	PHK γ 1	-9.07	PKC ϵ	-7.35	PKC η	-8.43	PKB α	-11.28	ASK γ	-5.54
MKK2	-6.97	MKK1	-8.48	PKA	-7.33	PKC ϵ	-5.17	S6K1	-11.18	PKC δ	-5.47
PHK γ 1	-6.57	PRAK	-7.93	PKC δ	-7.20	EGFR	-2.77	MKK1b	-10.78	PKC ϵ	-5.45
AMPK α	-4.79	CHK1	-7.38	INSR	-7.00	PKA	-1.29	PKC α	-10.73	PKC η	-5.42
PKC α	-4.50	GSK3 β	-6.47	PKC η	-6.96	PKC γ	-0.40	Erk2	-10.67	CDK4	-5.33
GSK3 β	-4.21	MKK1b	-5.81	Erk1	-6.78	PHK γ 1	11.68	PRAK	-9.81	PKC γ	-5.31
CHK1	-3.99	CK2 α 1	-5.77	PKC β	-6.66	cdc2	13.87	MKK2	-9.08	Erk1	-5.27
JNK1	-3.22	S6K1	-5.74	PKC α	-6.65	CK2 α 1	18.69	p38 γ	4.22	CDK6	-5.25
MKK1	-3.20	MSK1	-5.66	CK1 α	-6.53	PKC ζ	277.04	p38 β	4.87	Erk2	-5.23
p38 δ	-2.90	SGK	-4.72	ARAF	-6.47	CK1 α	470.10	p38 α	5.79	MKK1	-5.16
p38 γ	-2.66	p38 γ	0.57	PKC γ	-6.18			p38 δ	5.88	PKC β	-5.16
SGK	1.76	MKK2	1.64	JNK1	-5.61			JNK1	11.82	PKC α	-5.15
MSK1	2.04	JNK1	2.52	PKG1	-5.22			PHK γ 1	51.76	SRC	-5.05
PDK1	28.42	p38 δ	5.46	CK2 α 1	-4.92			CK2 α 1	71.53	PKG1	-5.03
CK2 α 1	38.97	PDK1	-11.54					PDK1	1053.13	ABL	-5.01
										Eg2	-4.99
										INSR	-4.84
										CK2 α 1	-3.73
										JNK1	-3.30

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes. The precision of the computed binding energies is ~ 0.03 kcal/mol.

FIG. 22

22/26

	SB203580 ^{1PME}		SB203580 ^{1A9U}		Purvalanol B	
	<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Test</i>
SCR	100.0 ± 0.0	97.1 ± 4.0	92.3 ± 5.6	88.1 ± 6.8	100.0 ± 0.0	93.9 ± 6.9
SCR (σ_{rep} 2.3)	100.0 ± 0.0	97.1 ± 4.0	92.3 ± 5.6	88.1 ± 6.8	86.9 ± 10.9	80.8 ± 8.2
AD FI	100.0 ± 0.0	97.2 ± 3.9	82.1 ± 14.8	76 ± 13.5	67.1 ± 9.7	60.2 ± 10.9
AD GA (BE)	27.3 ± 7.5	17.1 ± 6.4	34.9 ± 12.3	26 ± 10.3	34.4 ± 15.3	29.9 ± 11.0
AD GA (BR)	32.6 ± 7.7	23.0 ± 8.4	36.8 ± 14.8	28.7 ± 13.1	58.9 ± 19.8	52.6 ± 13.7
GMX/AD FI	-	-	43.2 ± 9.9	34.8 ± 9.2	64.9 ± 12.8	59.3 ± 10.5
Random	44.7 ± 23.1	36.5 ± 22.8	45.3 ± 22.5	37.4 ± 23.1	44.9 ± 22.2	38.4 ± 19.2
	Imatinib		H89		Hymenialdisine	
	<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Test</i>	<i>Train</i>	<i>Test</i>
SCR	89.6 ± 6.8	81.6 ± 11.3	88.8 ± 6.9	81.3 ± 9.5	89.0 ± 6.2	83.2 ± 8.6
SCR (σ_{rep} 2.3)	89.2 ± 6.8	82.3 ± 11.1	85.7 ± 7.3	77.2 ± 9.0	60.2 ± 26.2	50.8 ± 23.9
AD FI	92.3 ± 6.3	84.3 ± 8.8	67.6 ± 13.3	58.4 ± 10.1	65.5 ± 17.2	57.8 ± 14.0
AD GA (BE)	68.0 ± 10.0	61.7 ± 13.8	73.5 ± 11.3	64.2 ± 8.7	96.4 ± 3.6	92.2 ± 6.6
AD GA (BR)	55.6 ± 13.1	49.3 ± 10.8	73.4 ± 11.8	64.8 ± 8.9	96.6 ± 3.6	93.0 ± 5.0
GMX/AD FI	48.5 ± 19.3	42.9 ± 15.0	54.9 ± 16.4	43 ± 12.1	50.5 ± 13.1	40.8 ± 9.5
Random	45.9 ± 22.2	40.2 ± 19.1	51.5 ± 11.3	38.6 ± 11.2	41.5 ± 12.9	32.1 ± 12.8

FIG. 23

23/26

An ordered list of the computed binding energies obtained using AutoDock and the fixed-inhibitor assumption.

SB203580 ^{1PME}		SB203580 ^{1A9U}		Purvalanol B		Imatinib		H89		Hymenialdisine	
Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE
LCK	-11.14	p38 β	-10.28	CK1 α	-18.02	PKA	-20.36	PKA	-12.17	CK1 α	-12.73
p38 β	-11.11	p38 α	-10.26	cdc2	-17.57	ABL	-20.23	PKB α	-11.81	CDK2	-12.44
p38 α	-10.91	p38 δ	-8.89	CDK4	-17.46	PDGFR α	-20.08	MKK1b	-11.77	cdc2	-12.42
MKK1b	-10.58	p38 γ	-8.87	PKC η	-17.15	PDGFR β	-20.04	MSK1	-11.76	CDK3	-12.42
S6K1	-10.56	ROCK2	-8.82	PKC γ	-17.08	LYN	-19.91	S6K1	-11.69	CDK5	-12.38
MSK1	-10.54	AMPK α 1	-8.76	PKC ϵ	-17.03	KIT	-19.88	MKK1	-11.58	CDK6	-12.32
SGK	-10.39	PKA	-8.62	PKC δ	-16.94	SRC	-19.80	MKK2	-11.51	CDK4	-12.32
CHK1	-10.36	PKC α	-8.51	DIAPYK1A	-16.92	EGFR	-19.66	AMPK α 1	-11.26	PKA	-11.96
GSK3 β	-10.25	PKB α	-8.50	PKC α	-16.89	PKC α	-19.51	CHK1	-11.25	SRC	-11.72
Erk2	-10.06	LCK	-7.99	PKC β	-16.87	PKC β	-19.37	LCK	-11.20	ABL	-11.71
PKA	-10.02	MKK2	-6.72	CDK2	-16.75	PKC ϵ	-19.12	PKC α	-11.11	PKC δ	-11.63
PRAK	-9.96	PRAK	-4.94	CDK5	-16.68	PKC η	-19.10	p38 δ	-9.73	PKC ϵ	-11.61
PKB α	-9.94	JNK1	-4.68	cdc28	-16.63	FGFR1	-18.38	p38 γ	-9.72	PKC η	-11.60
MKK2	-9.86	MKK1	-4.22	ABL	-16.62	PKC γ	-18.33	p38 β	-9.62	PKC γ	-11.42
AMPK α 1	-9.65	S6K1	23.29	INSR	-16.53	CK1 α	-12.33	p38 α	-8.76	GSK3 β	-11.33
PKC α	-9.46	MSK1	23.36	PKA	-16.36	cdc2	14.71	SGK	-8.10	PKG1	-11.33
ROCK2	-9.38	GSK3 β	23.45	Erk1	-16.11	CK2 α 1	15.40	Erk2	-7.53	Eg2	-11.18
MKK1	-9.12	SGK	23.45	JNK1	-15.29	PHK γ 1	17.16	GSK3 β	-7.42	PKC β	-11.13
p38 γ	-8.96	MKK1b	23.48	PKG1	-15.13	PKC ζ	34746.38	ROCK2	-7.28	PKC α	-11.12
p38 δ	-8.92	CHK1	23.53	PKC ζ	-12.97			JNK1	-6.36	ASK γ	-11.11
JNK1	-8.72	Erk2	25.31	ARAF	-12.81			PRAK	-5.73	MKK1	-11.05
PDK1	1208.85	PDK1	282.21	GSK3 β	-7.53			PHK γ 1	3.58	INSR	-10.70
PHK γ 1	116105.19	CK2 α 1	49991.23	CK2 α 1	-5.44			CK2 α 1	27.20	CK2 α 1	-9.34
CK2 α 1	116112.83	PHK γ 1	49995.22					PDK1	8577.08	Erk1	6.45
										Erk2	6.45
										JNK1	7.38
										PKC ζ	11.48
										ARAF	11.68

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes.

FIG. 24

SB203580 ^{1PME}				SB203580 ^{1A9U}				Purvalanol B				Imatinib				H89				Hymenialdisine			
Kinase	RMSD	ΔE		Kinase	RMSD	ΔE		Kinase	RMSD	ΔE		Kinase	RMSD	ΔE		Kinase	RMSD	ΔE		Kinase	RMSD	ΔE	
S6K1	0.945	-13.58	AMPK α 1	7.081	-14.68	CK2 α 1	1.611	-22.3	LYN	1.931	-23.28	PKA	0.664	-12.52	CK2 α 1	0.715	-14.07						
Erk2	0.874	-13.58	GSK3 β	5.988	-14.5	DYRK1A	1.72	-21.64	PKC ϵ	1.135	-23.12	S6K1	0.883	-12.41	GSK3 β	1.261	-13.81						
MKK1	1.022	-13.56	p38 δ	5.82	-14.42	ARAF	1.882	-21.53	SRC	1.741	-22.86	MSK1	0.84	-12.41	ASK1	1.286	-13.77						
MKK2	0.905	-13.54	p38 γ	5.697	-14.04	Gdc2	1.307	-21.41	PKA	1.038	-22.81	MKK1b	0.841	-12.39	CK1 α	0.624	-13.63						
PRAK	0.823	-13.53	PKA	5.699	-13.87	CK1 α	1.542	-21.34	PKC η	0.881	-22.72	SGK	0.807	-12.33	Gdc2	0.698	-13.54						
JNK1	0.931	-13.37	ROCK2	5.622	-13.86	GSK3 β	1.355	-21.28	ABL	0.999	-22.71	PKB α	0.867	-12.29	CDK3	0.77	-13.54						
PKB α	0.874	-13.29	CHK1	5.934	-13.75	CDK4	1.635	-21.19	CK1 α	1.606	-22.66		1.099	-12.1	CDK2	0.704	-13.54						
PKA	0.902	-13.28	S6K1	5.916	-13.72	PKC γ	1.576	-21.06	PDGFR β	1.025	-22.47		1.095	-12.09	CDK5	0.689	-13.48						
MSK1	0.973	-13.25	PDK1	6.077	-13.66	PKC η	1.584	-21.04	EGFR	1.188	-22.42	PKC α	1.111	-11.98	PKC γ	1.316	-13.43						
MKK1b	0.907	-13.22	PKC α	5.987	-13.62	PKC ϵ	1.593	-20.94	PDGFR α	0.89	-22.39	ROCK2	1.115	-11.97	CDK4	0.707	-13.43						
SGK	0.951	-13.2	p38 β	5.422	-13.57	CDK2	1.648	-20.91	FGFR1	1.815	-22.28	MKK2	0.693	-11.95	CDK6	0.698	-13.42						
CHK1	0.96	-13.14	PKB α	5.363	-13.56	Gdc28	1.613	-20.89	PKC α	0.772	-22	MKK1	0.516	-11.92	PKA	1.233	-13.29						
GSK3 β	1.084	-13.1	p38 δ	5.677	-13.49	INSR	1.478	-20.86	Gdc2	0.835	-21.97	LCK	0.941	-11.84	MKK1	1.413	-13.16						
PDK1	5.057	-13.07	MSK1	5.68	-13.44	PKC α	1.298	-20.83	KIT	0.794	-21.93	p38 β	1.043	-11.83	Eg2	1.277	-13.08						
PKC α	0.92	-13.02	MKK1b	5.952	-13.25	PKC δ	1.591	-20.82	PKC β	0.835	-21.84	CHK1	0.846	-11.83	PKC α	1.322	-13.05						
ROCK2	0.972	-12.96	SGK	5.387	-13.08	CDK5	1.642	-20.8	CK2 α 1	0.953	-21.78	AMPK α 1	0.727	-11.75	PKC β	1.311	-13.03						
p38 δ	0.961	-12.95	CK2 α 1	5.312	-12.96	PKC β	1.255	-20.8	PKC γ	0.79	-20.98	p38 α	1.048	-11.75	Erk1	2.987	-12.87						
AMPK α 1	0.907	-12.91	Erk2	5.204	-12.61	JNK1	2.308	-20.38	PHK γ 1	1.036	-20.2	JNK1	1.1	-11.63	Erk2	2.984	-12.86						
p38 γ	0.939	-12.78	MKK2	5.242	-12.6	PKA	1.563	-20.23	PKC ζ	1.344	-19.88	PRAK	0.571	-11.61	JNK1	1.572	-12.8						
UGK	0.971	-12.64	JNK1	5.621	-12.18	Erk1	1.665	-20.15				PHK γ 1	1.05	-11.52	SRC	0.652	-12.79						
p38 γ	1.478	-12.63	UGK	5.721	-12.17	PKC ζ	1.551	-20.04				GSK3 β	1.554	-11.27	ABL	0.633	-12.78						
p38 δ	1.643	-12.61	PHK γ 1	5.515	-12.12	ABL	1.665	-19.99				Erk2	4.608	-11.2	PKC δ	0.711	-12.77						
CK2 α 1	1.352	-11.94	PRAK	1.88	-11.7	PKGI	1.506	-18.95				CK2 α 1	0.906	-11.2	PKC ζ	1.585	-12.76						
PHK γ 1	3.527	-11.65	MKK1	4.513	-11.01							PDK1	4.725	-10.68	PKC ϵ	0.693	-12.75						
															PKC η	0.701	-12.74						
															PKGI	0.741	-12.46						
															INSR	1.671	-12.45						
															ARAF	1.636	-12.39						

Kinases determined to be 'experimental targets' of the inhibitor are indicated by shaded boxes.

FIG. 25

An ordered list of computed binding energies obtained by applying AutoDock's energy function to complexes that were first energy-minimized by GROMACS.

SB203580 ^{1A9U}		Purvalanol B		Imatinib		H89		Hymenialdisine	
Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE	Kinase	ΔE
p38 δ	-10.4	PKC ζ	-12.09	PDGFR α	-19.42	PKA	-13.17	CDK5	-12.07
PKB α	-10.2	cdc2	-11.74	PKC ϵ	-18.69	SGK	-13.13	ABL	-11.95
ROCK2	-10.0	GSK3 β	-11.06	PKA	-18.59	CK2 α 1	-13.11	GSK3 β	-11.93
Erk2	-10.0	CDK4	-10.84	KIT	-18.47	CHK1	-12.63	PKC α	-11.67
p38 α	-10.0	PKC ϵ	-10.79	PKC η	-18.4	MKK1	-12.68	MKK1	-11.58
p38 γ	-9.9	ARAF	-10.72	PKC β	-18.33	ROCK2	-12.53	Eg2	-11.4
CHK1	-9.8	PKC η	-10.71	FGFR1	-17.94	GSK3 β	-12.37	CDK2	-11.38
PKA	-9.8	MYRK14	-10.63	CK2 α 1	-17.69	MSK1	-12.24	CDK3	-11.3
MKK1b	-9.7	CDK5	-10.53	cdc2	-17.6	LCK	-12.18	PKC δ	-11.28
CK2 α 1	-9.6	cdc28	-10.47	SRC	-17.5	S6K1	-11.97	PKC ζ	-11.28
PKC α	-9.6	CK2 α 1	-10.43	LYN	-17.5	PKC α	-11.74	CK2 α 1	-11.25
GSK3 β	-9.5	PKC γ	-10.4	EGFR	-17.42	AMPK α 1	-11.65	PKC γ	-11.2
S6K1	-9.5	CK1 α	-10.21	EGFR	-17.36	p38 δ	-11.62	PKC η	-11.16
MSK1	-9.5	CDK2	-10.18	ABL	-17.2	PHK γ 1	-11.62	cdc2	-11.08
PRAK	-9.4	PKC β	-10.09	PDGFR β	-17.14	PRAK	-11.57	PKG1	-11.05
AMPK α 1	-9.4	PKC α	-10.08	PKC ζ	-16.74	p38 γ	-11.56	CDK6	-10.97
p38 β	-9.3	PKC δ	-9.93	PKC γ	-16.01	PKB α	-11.53	ARAF	-10.94
LCK	-9.0	PKA	-9.84	PHK γ 1	-15.57	PDK1	-11.23	SRC	-10.79
MKK1	-8.9	INSR	-9.76	CK1 α	-15.47	MKK2	-11.17	ASK γ	-10.65
SGK	-8.8	ABL	-9.72			p38 β	-11.16	Erk2	-10.64
JNK1	-8.8	JNK1	-8.19			p38 α	-11.1	PKC β	-10.59
PDK1	-8.8	Erk1	-8.15			JNK1	-11.04	CDK4	-10.53
MKK2	-8.6	PKG1	-6.88			Erk2	-10.97	PKC ϵ	-10.11
PHK γ 1	-7.3					MKK1b	-10.73	CK1 α	-10.09
								INSR	-10.06
								Erk1	-9.96
								PKA	-9.94
								JNK1	-7.96

FIG. 26

An ordered list of the computed binding energies obtained with SCR

Indirubin ^{2.6}		Quercetin ^{2.6}	
Kinase	ΔE	Kinase	ΔE
GSK3 β	-10.35	PKB α	-9.35
CDC2	-9.24	PKA	-9.21
PHK γ 1	-8.98	PRAK	-9.17
CDK5	-8.88	SGK	-8.98
PKC α	-8.88	MKK1	-8.81
MKK1	-8.53	GSK3 β	-8.74
AMPK α 1	-8.40	AMPK α 1	-8.19
S6K1	-8.36	MSK1	-8.09
CHK1	-7.97	MKK2	-8.00
MSK1	-7.82	PKC α	-7.81
LCK	-7.78	CHK1	-7.69
MKK1b	-7.58	LCK	-7.50
PKA	-7.11	MKK1b	-7.35
SGK	-7.10	ROCK2	-7.05
ROCK2	-7.08	S6K1	-6.73
Erk2	-6.65	Erk2	-5.76
p38 γ	-6.53	p38 γ	-5.44
p38 δ	-6.36	p38 δ	-5.38
PKB α	-6.27	JNK1	-4.82
MKK2	-5.93	p38 α	-4.67
p38 α	-5.41	p38 β	-4.28
CK2 α 1	-5.33	PDK1	-1.86
PRAK	-5.16	CK2 α 1	13.14
p38 β	-5.00	PHK γ 1	54.13
PDK1	4.26		
JNK1	2.17		

FIG. 27

SEQUENCE LISTING

<110> University of Iowa Research Foundation
Elcock, Adrian H.
Rockey, William M.

<120> RAPID COMPUTATIONAL IDENTIFICATION OF
TARGETS

<130> 21087.0034P1

<150> 60/676,500

<151> 2005-04-29

<150> 60/618,211

<151> 2004-10-12

<160> 64

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 1130

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 1

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

Tyr Glu Gly Val Trp Lys Lys Tyr Ser Leu Thr Val Ala Val Lys Thr
 260 265 270
 Leu Lys Glu Asp Thr Met Glu Val Glu Glu Phe Leu Lys Glu Ala Ala
 275 280 285
 Val Met Lys Glu Ile Lys His Pro Asn Leu Val Gln Leu Leu Gly Val
 290 295 300
 Cys Thr Arg Glu Pro Pro Phe Tyr Ile Ile Thr Glu Phe Met Thr Tyr
 305 310 315 320
 Gly Asn Leu Leu Asp Tyr Leu Arg Glu Cys Asn Arg Gln Glu Val Asn
 325 330 335
 Ala Val Val Leu Leu Tyr Met Ala Thr Gln Ile Ser Ser Ala Met Glu
 340 345 350
 Tyr Leu Glu Lys Lys Asn Phe Ile His Arg Asp Leu Ala Ala Arg Asn
 355 360 365
 Cys Leu Val Gly Glu Asn His Leu Val Lys Val Ala Asp Phe Gly Leu
 370 375 380
 Ser Arg Leu Met Thr Gly Asp Thr Tyr Thr Ala His Ala Gly Ala Lys
 385 390 395 400
 Phe Pro Ile Lys Trp Thr Ala Pro Glu Ser Leu Ala Tyr Asn Lys Phe
 405 410 415
 Ser Ile Lys Ser Asp Val Trp Ala Phe Gly Val Leu Leu Trp Glu Ile
 420 425 430
 Ala Thr Tyr Gly Met Ser Pro Tyr Pro Gly Ile Asp Leu Ser Gln Val
 435 440 445
 Tyr Glu Leu Leu Glu Lys Asp Tyr Arg Met Glu Arg Pro Glu Gly Cys
 450 455 460
 Pro Glu Lys Val Tyr Glu Leu Met Arg Ala Cys Trp Gln Trp Asn Pro
 465 470 475 480
 Ser Asp Arg Pro Ser Phe Ala Glu Ile His Gln Ala Phe Glu Thr Met
 485 490 495
 Phe Gln Glu Ser Ser Ile Ser Asp Glu Val Glu Lys Glu Leu Gly Lys
 500 505 510
 Gln Gly Val Arg Gly Ala Val Ser Thr Leu Leu Gln Ala Pro Glu Leu
 515 520 525
 Pro Thr Lys Thr Arg Thr Ser Arg Arg Ala Ala Glu His Arg Asp Thr
 530 535 540
 Thr Asp Val Pro Glu Met Pro His Ser Lys Gly Gln Gly Glu Ser Asp
 545 550 555 560
 Pro Leu Asp His Glu Pro Ala Val Ser Pro Leu Leu Pro Arg Lys Glu
 565 570 575
 Arg Gly Pro Pro Glu Gly Gly Leu Asn Glu Asp Glu Arg Leu Leu Pro
 580 585 590
 Lys Asp Lys Lys Thr Asn Leu Phe Ser Ala Leu Ile Lys Lys Lys Lys
 595 600 605
 Lys Thr Ala Pro Thr Pro Pro Lys Arg Ser Ser Ser Phe Arg Glu Met
 610 615 620
 Asp Gly Gln Pro Glu Arg Arg Gly Ala Gly Glu Glu Gly Arg Asp
 625 630 635 640
 Ile Ser Asn Gly Ala Leu Ala Phe Thr Pro Leu Asp Thr Ala Asp Pro
 645 650 655
 Ala Lys Ser Pro Lys Pro Ser Asn Gly Ala Gly Val Pro Asn Gly Ala
 660 665 670
 Leu Arg Glu Ser Gly Gly Ser Gly Phe Arg Ser Pro His Leu Trp Lys
 675 680 685
 Lys Ser Ser Thr Leu Thr Ser Ser Arg Leu Ala Thr Gly Glu Glu Glu
 690 695 700
 Gly Gly Gly Ser Ser Ser Lys Arg Phe Leu Arg Ser Cys Ser Ala Ser
 705 710 715 720
 Cys Val Pro His Gly Ala Lys Asp Thr Glu Trp Arg Ser Val Thr Leu
 725 730 735

```

Pro Arg Asp Leu Gln Ser Thr Gly Arg Gln Phe Asp Ser Ser Thr Phe
      740      745      750
Gly Gly His Lys Ser Glu Lys Pro Ala Leu Pro Arg Lys Arg Ala Gly
      755      760      765
Glu Asn Arg Ser Asp Gln Val Thr Arg Gly Thr Val Thr Pro Pro Pro
      770      775      780
Arg Leu Val Lys Lys Asn Glu Glu Ala Ala Asp Glu Val Phe Lys Asp
      785      790      795      800
Ile Met Glu Ser Ser Pro Gly Ser Ser Pro Pro Asn Leu Thr Pro Lys
      805      810      815
Pro Leu Arg Arg Gln Val Thr Val Ala Pro Ala Ser Gly Leu Pro His
      820      825      830
Lys Glu Glu Ala Glu Lys Gly Ser Ala Leu Gly Thr Pro Ala Ala Ala
      835      840      845
Glu Pro Val Thr Pro Thr Ser Lys Ala Gly Ser Gly Ala Pro Gly Gly
      850      855      860
Thr Ser Lys Gly Pro Ala Glu Glu Ser Arg Val Arg Arg His Lys His
      865      870      875      880
Ser Ser Glu Ser Pro Gly Arg Asp Lys Gly Lys Leu Ser Arg Leu Lys
      885      890      895
Pro Ala Pro Pro Pro Pro Pro Ala Ala Ser Ala Gly Lys Ala Gly Gly
      900      905      910
Lys Pro Ser Gln Ser Pro Ser Gln Glu Ala Ala Gly Glu Ala Val Leu
      915      920      925
Gly Ala Lys Thr Lys Ala Thr Ser Leu Val Asp Ala Val Asn Ser Asp
      930      935      940
Ala Ala Lys Pro Ser Gln Pro Gly Glu Gly Leu Lys Lys Pro Val Leu
      945      950      955      960
Pro Ala Thr Pro Lys Pro Gln Ser Ala Lys Pro Ser Gly Thr Pro Ile
      965      970      975
Ser Pro Ala Pro Val Pro Ser Thr Leu Pro Ser Ala Ser Ser Ala Leu
      980      985      990
Ala Gly Asp Gln Pro Ser Ser Thr Ala Phe Ile Pro Leu Ile Ser Thr
      995      1000      1005
Arg Val Ser Leu Arg Lys Thr Arg Gln Pro Pro Glu Arg Ile Ala Ser
      1010      1015      1020
Gly Ala Ile Thr Lys Gly Val Val Leu Asp Ser Thr Glu Ala Leu Cys
      1025      1030      1035      1040
Leu Ala Ile Ser Arg Asn Ser Glu Gln Met Ala Ser His Ser Ala Val
      1045      1050      1055
Leu Glu Ala Gly Lys Asn Leu Tyr Thr Phe Cys Val Ser Tyr Val Asp
      1060      1065      1070
Ser Ile Gln Gln Met Arg Asn Lys Phe Ala Phe Arg Glu Ala Ile Asn
      1075      1080      1085
Lys Leu Glu Asn Asn Leu Arg Glu Leu Gln Ile Cys Pro Ala Thr Ala
      1090      1095      1100
Gly Ser Gly Pro Ala Ala Thr Gln Asp Phe Ser Lys Leu Leu Ser Ser
      1105      1110      1115      1120
Val Lys Glu Ile Ser Asp Ile Val Gln Arg
      1125      1130

```

<210> 2

<211> 1036

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 2

Met	Gln	Pro	Glu	Glu	Gly	Thr	Gly	Trp	Leu	Leu	Glu	Leu	Leu	Ser	Glu	1	5	10	15
Val	Gln	Leu	Gln	Gln	Tyr	Phe	Leu	Arg	Leu	Arg	Asp	Asp	Leu	Asn	Val	20	25	30	
Thr	Arg	Leu	Ser	His	Phe	Glu	Tyr	Val	Lys	Asn	Glu	Asp	Leu	Glu	Lys	35	40	45	
Ile	Gly	Met	Gly	Arg	Pro	Gly	Gln	Arg	Arg	Leu	Trp	Glu	Ala	Val	Lys	50	55	60	
Arg	Arg	Lys	Ala	Leu	Cys	Lys	Arg	Lys	Ser	Trp	Met	Ser	Lys	Val	Phe	65	70	75	80
Ser	Gly	Lys	Arg	Leu	Glu	Ala	Glu	Phe	Pro	Pro	His	His	Ser	Gln	Ser	85	90	95	
Thr	Phe	Arg	Lys	Thr	Ser	Pro	Ala	Pro	Gly	Gly	Pro	Ala	Gly	Glu	Gly	100	105	110	
Pro	Leu	Gln	Ser	Leu	Thr	Cys	Leu	Ile	Gly	Glu	Lys	Asp	Leu	Arg	Leu	115	120	125	
Leu	Glu	Lys	Leu	Gly	Asp	Gly	Ser	Phe	Gly	Val	Val	Arg	Arg	Gly	Glu	130	135	140	
Trp	Asp	Ala	Pro	Ser	Gly	Lys	Thr	Val	Ser	Val	Ala	Val	Lys	Cys	Leu	145	150	155	160
Lys	Pro	Asp	Val	Leu	Ser	Gln	Pro	Glu	Ala	Met	Asp	Asp	Phe	Ile	Arg	165	170	175	
Glu	Val	Asn	Ala	Met	His	Ser	Leu	Asp	His	Arg	Asn	Leu	Ile	Arg	Leu	180	185	190	
Tyr	Gly	Val	Val	Leu	Thr	Pro	Pro	Met	Lys	Met	Val	Thr	Glu	Leu	Ala	195	200	205	
Pro	Leu	Gly	Ser	Leu	Leu	Asp	Arg	Leu	Arg	Lys	His	Gln	Gly	His	Phe	210	215	220	
Leu	Leu	Gly	Thr	Leu	Ser	Arg	Tyr	Ala	Val	Gln	Val	Ala	Glu	Gly	Met	225	230	235	240
Gly	Tyr	Leu	Glu	Ser	Lys	Arg	Phe	Ile	His	Arg	Asp	Leu	Ala	Ala	Arg	245	250	255	
Asn	Leu	Leu	Leu	Ala	Thr	Arg	Asp	Leu	Val	Lys	Ile	Gly	Asp	Phe	Gly	260	265	270	
Leu	Met	Arg	Ala	Leu	Pro	Gln	Asn	Asp	Asp	His	Tyr	Val	Met	Gln	Glu	275	280	285	
His	Arg	Lys	Val	Pro	Phe	Ala	Trp	Cys	Ala	Pro	Glu	Ser	Leu	Lys	Thr	290	295	300	
Arg	Thr	Phe	Ser	His	Ala	Ser	Asp	Thr	Trp	Met	Phe	Gly	Val	Thr	Leu	305	310	315	320
Trp	Glu	Met	Phe	Thr	Tyr	Gly	Gln	Glu	Pro	Trp	Ile	Gly	Leu	Asn	Gly	325	330	335	
Ser	Gln	Ile	Leu	His	Lys	Ile	Asp	Lys	Glu	Gly	Glu	Arg	Leu	Pro	Arg	340	345	350	
Pro	Glu	Asp	Cys	Pro	Gln	Asp	Ile	Tyr	Asn	Val	Met	Val	Gln	Cys	Trp	355	360	365	
Ala	His	Lys	Pro	Glu	Asp	Arg	Pro	Thr	Phe	Val	Ala	Leu	Arg	Asp	Phe	370	375	380	
Leu	Leu	Glu	Ala	Gln	Pro	Thr	Asp	Met	Arg	Ala	Leu	Gln	Asp	Phe	Glu	385	390	395	400
Glu	Pro	Asp	Lys	Leu	His	Ile	Gln	Met	Asn	Asp	Val	Ile	Thr	Val	Ile	405	410	415	
Glu	Gly	Arg	Ala	Glu	Asn	Tyr	Trp	Trp	Arg	Gly	Gln	Asn	Thr	Arg	Thr	420	425	430	
Leu	Cys	Val	Gly	Pro	Phe	Pro	Arg	Asn	Val	Val	Thr	Ser	Val	Ala	Gly	435	440	445	
Leu	Ser	Ala	Gln	Asp	Ile	Ser	Gln	Pro	Leu	Gln	Asn	Ser	Phe	Ile	His	450	455	460	
Thr	Gly	His	Gly	Asp	Ser	Asp	Pro	Arg	His	Cys	Trp	Gly	Phe	Pro	Asp	465	470	475	480

Arg Ile Asp Glu Leu Tyr Leu Gly Asn Pro Met Asp Pro Pro Asp Leu
 485 490 495
 Leu Ser Val Glu Leu Ser Thr Ser Arg Pro Pro Gln His Leu Gly Gly
 500 505 510
 Val Lys Lys Pro Thr Tyr Asp Pro Val Ser Glu Asp Gln Asp Pro Leu
 515 520 525
 Ser Ser Asp Phe Lys Arg Leu Gly Leu Arg Lys Pro Gly Leu Pro Arg
 530 535 540
 Gly Leu Trp Leu Ala Lys Pro Ser Ala Arg Val Pro Gly Thr Lys Ala
 545 550 555 560
 Ser Arg Gly Ser Gly Ala Glu Val Thr Leu Ile Asp Phe Gly Glu Glu
 565 570 575
 Pro Val Val Pro Ala Leu Arg Pro Cys Pro Pro Ser Leu Ala Gln Leu
 580 585 590
 Ala Met Asp Ala Cys Ser Leu Leu Asp Glu Thr Pro Pro Gln Ser Pro
 595 600 605
 Thr Arg Ala Leu Pro Arg Pro Leu His Pro Thr Pro Val Val Asp Trp
 610 615 620
 Asp Ala Arg Pro Leu Pro Pro Pro Pro Ala Tyr Asp Asp Val Ala Gln
 625 630 635 640
 Asp Glu Asp Asp Phe Glu Ile Cys Ser Ile Asn Ser Thr Leu Val Gly
 645 650 655
 Ala Gly Val Pro Ala Gly Pro Ser Gln Gly Gln Thr Asn Tyr Ala Phe
 660 665 670
 Val Pro Glu Gln Ala Arg Pro Pro Pro Leu Glu Asp Asn Leu Phe
 675 680 685
 Leu Pro Pro Gln Gly Gly Gly Lys Pro Pro Ser Ser Ala Gln Thr Ala
 690 695 700
 Glu Ile Phe Gln Ala Leu Gln Gln Glu Cys Met Arg Gln Leu Gln Ala
 705 710 715 720
 Pro Gly Ser Pro Ala Pro Ser Pro Ser Pro Gly Gly Asp Asp Lys Pro
 725 730 735
 Gln Val Pro Pro Arg Val Pro Ile Pro Pro Arg Pro Thr Arg Pro His
 740 745 750
 Val Gln Leu Ser Pro Ala Pro Pro Gly Glu Glu Glu Thr Ser Gln Trp
 755 760 765
 Pro Gly Pro Ala Ser Pro Pro Arg Val Pro Pro Arg Glu Pro Leu Ser
 770 775 780
 Pro Gln Gly Ser Arg Thr Pro Ser Pro Leu Val Pro Pro Gly Ser Ser
 785 790 795 800
 Pro Leu Pro Pro Arg Leu Ser Ser Ser Pro Gly Lys Thr Met Pro Thr
 805 810 815
 Thr Gln Ser Phe Ala Ser Asp Pro Lys Tyr Ala Thr Pro Gln Val Ile
 820 825 830
 Gln Ala Pro Gly Ala Gly Gly Pro Cys Ile Leu Pro Ile Val Arg Asp
 835 840 845
 Gly Lys Lys Val Ser Ser Thr His Tyr Tyr Leu Leu Pro Glu Arg Pro
 850 855 860
 Ser Tyr Leu Glu Arg Tyr Gln Arg Phe Leu Arg Glu Ala Gln Ser Pro
 865 870 875 880
 Glu Glu Pro Thr Pro Leu Pro Val Pro Leu Leu Leu Pro Pro Pro Ser
 885 890 895
 Thr Pro Ala Pro Ala Ala Pro Thr Ala Thr Val Arg Pro Met Pro Gln
 900 905 910
 Ala Ala Leu Asp Pro Lys Ala Asn Phe Ser Thr Asn Asn Ser Asn Pro
 915 920 925
 Gly Ala Arg Pro Pro Pro Pro Arg Ala Thr Ala Arg Leu Pro Gln Arg
 930 935 940
 Gly Cys Pro Gly Asp Gly Pro Glu Ala Gly Arg Pro Ala Asp Lys Ile
 945 950 955 960

Gln Met Ala Met Val His Gly Val Thr Thr Glu Glu Cys Gln Ala Ala
 965 970 975
 Leu Gln Cys His Gly Trp Ser Val Gln Arg Ala Ala Gln Tyr Leu Lys
 980 985 990
 Val Glu Gln Leu Phe Gly Leu Gly Leu Arg Pro Arg Gly Glu Cys His
 995 1000 1005
 Lys Val Leu Glu Met Phe Asp Trp Asn Leu Glu Gln Ala Gly Cys His
 1010 1015 1020
 Leu Leu Gly Ser Trp Gly Pro Ala His His Lys Arg
 1025 1030 1035

<210> 3

<211> 2245

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 3

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

Thr	Gly	Ser	Val	Lys	Ile	Gly	Asp	Leu	Gly	Leu	Ala	Thr	Leu	Lys	Arg	325	330	335
Ala	Ser	Phe	Ala	Lys	Ser	Val	Ile	Gly	Thr	Pro	Glu	Phe	Met	Ala	Pro	340	345	350
Glu	Met	Tyr	Glu	Glu	His	Tyr	Asp	Glu	Ser	Val	Asp	Val	Tyr	Ala	Phe	355	360	365
Gly	Met	Cys	Met	Leu	Glu	Met	Ala	Thr	Ser	Glu	Tyr	Pro	Tyr	Ser	Glu	370	375	380
Cys	Gln	Asn	Ala	Ala	Gln	Ile	Tyr	Arg	Lys	Val	Thr	Cys	Gly	Ile	Lys	385	390	395
Pro	Ala	Ser	Phe	Glu	Lys	Val	His	Asp	Pro	Glu	Ile	Lys	Glu	Ile	Ile	405	410	415
Gly	Glu	Cys	Ile	Cys	Lys	Asn	Lys	Glu	Arg	Tyr	Glu	Ile	Lys	Asp		420	425	430
Leu	Leu	Ser	His	Ala	Phe	Phe	Ala	Glu	Asp	Thr	Gly	Val	Arg	Val	Glu	435	440	445
Leu	Ala	Glu	Glu	Asp	His	Gly	Arg	Lys	Ser	Thr	Ile	Ala	Leu	Arg	Leu	450	455	460
Trp	Val	Glu	Asp	Pro	Lys	Lys	Leu	Lys	Gly	Lys	Pro	Lys	Asp	Asn	Gly	465	470	475
Ala	Ile	Glu	Phe	Thr	Phe	Asp	Leu	Glu	Lys	Glu	Thr	Pro	Asp	Glu	Val	485	490	495
Ala	Gln	Glu	Met	Ile	Glu	Ser	Gly	Phe	Phe	His	Glu	Ser	Asp	Val	Lys	500	505	510
Ile	Val	Ala	Lys	Ser	Ile	Arg	Asp	Arg	Val	Ala	Leu	Ile	Gln	Trp	Arg	515	520	525
Arg	Glu	Arg	Ile	Trp	Pro	Ala	Leu	Gln	Pro	Lys	Glu	Gln	Gln	Asp	Val	530	535	540
Gly	Ser	Pro	Asp	Lys	Ala	Arg	Gly	Pro	Pro	Val	Pro	Leu	Gln	Val	Gln	545	550	555
Val	Thr	Tyr	His	Ala	Gln	Ala	Gly	Gln	Pro	Gly	Pro	Pro	Glu	Pro	Glu	565	570	575
Glu	Pro	Glu	Ala	Asp	Gln	His	Leu	Leu	Pro	Pro	Thr	Leu	Pro	Thr	Ser	580	585	590
Ala	Thr	Ser	Leu	Ala	Ser	Asp	Ser	Thr	Phe	Asp	Ser	Gly	Gln	Gly	Ser	595	600	605
Thr	Val	Tyr	Ser	Asp	Ser	Gln	Ser	Ser	Gln	Gln	Ser	Val	Met	Leu	Gly	610	615	620
Ser	Leu	Ala	Asp	Ala	Ala	Pro	Ser	Pro	Ala	Gln	Cys	Val	Cys	Ser	Pro	625	630	635
Pro	Val	Ser	Glu	Gly	Pro	Val	Leu	Pro	Gln	Ser	Leu	Pro	Ser	Leu	Gly	645	650	655
Ala	Tyr	Gln	Gln	Pro	Thr	Ala	Ala	Pro	Gly	Leu	Pro	Val	Gly	Ser	Val	660	665	670
Pro	Ala	Pro	Ala	Cys	Pro	Pro	Ser	Leu	Gln	Gln	His	Phe	Pro	Asp	Pro	675	680	685
Ala	Met	Ser	Phe	Ala	Pro	Val	Leu	Pro	Pro	Pro	Ser	Thr	Pro	Met	Pro	690	695	700
Thr	Gly	Pro	Gly	Gln	Pro	Ala	Pro	Pro	Gly	Gln	Pro	Pro	Pro	Leu		705	710	715
Ala	Gln	Pro	Thr	Pro	Leu	Pro	Gln	Val	Leu	Ala	Pro	Gln	Pro	Val	Val	725	730	735
Pro	Leu	Gln	Pro	Val	Pro	Pro	His	Leu	Pro	Pro	Tyr	Leu	Ala	Pro	Ala	740	745	750
Ser	Gln	Val	Gly	Ala	Pro	Ala	Gln	Leu	Lys	Pro	Leu	Gln	Met	Pro	Gln	755	760	765
Ala	Pro	Leu	Gln	Pro	Leu	Ala	Gln	Val	Pro	Pro	Gln	Met	Pro	Pro	Ile	770	775	780
Pro	Val	Val	Pro	Pro	Ile	Thr	Pro	Leu	Ala	Gly	Ile	Asp	Gly	Leu	Pro	785	790	795
																		800

Pro Ala Leu Pro Asp Leu Pro Thr Ala Thr Val Pro Pro Val Pro Pro
 805 810 815
 Pro Gln Tyr Phe Ser Pro Ala Val Ile Leu Pro Ser Leu Ala Ala Pro
 820 825 830
 Leu Pro Pro Ala Ser Pro Ala Leu Pro Leu Gln Ala Val Lys Leu Pro
 835 840 845
 His Pro Pro Gly Ala Pro Leu Ala Met Pro Cys Arg Thr Ile Val Pro
 850 855 860
 Asn Ala Pro Ala Thr Ile Pro Leu Leu Ala Val Ala Pro Pro Gly Val
 865 870 875 880
 Ala Ala Leu Ser Ile His Ser Ala Val Ala Gln Leu Pro Gly Gln Pro
 885 890 895
 Val Tyr Pro Ala Ala Phe Pro Gln Met Ala Pro Thr Asp Val Pro Pro
 900 905 910
 Ser Pro His His Thr Val Gln Asn Met Arg Ala Thr Pro Pro Gln Pro
 915 920 925
 Ala Leu Pro Pro Gln Pro Thr Leu Pro Pro Gln Pro Val Leu Pro Pro
 930 935 940
 Gln Pro Thr Leu Pro Pro Gln Pro Val Leu Pro Pro Gln Pro Thr Arg
 945 950 955 960
 Pro Pro Gln Pro Val Leu Pro Pro Gln Pro Met Leu Pro Pro Gln Pro
 965 970 975
 Val Leu Pro Pro Gln Pro Ala Leu Pro Val Arg Pro Glu Pro Leu Gln
 980 985 990
 Pro His Leu Pro Glu Gln Ala Ala Pro Ala Ala Thr Pro Gly Ser Gln
 995 1000 1005
 Ile Leu Leu Gly His Pro Ala Pro Tyr Ala Val Asp Val Ala Ala Gln
 1010 1015 1020
 Val Pro Thr Val Pro Val Pro Pro Ala Ala Val Leu Ser Pro Pro Leu
 1025 1030 1035 1040
 Pro Glu Val Leu Leu Pro Ala Ala Pro Glu Leu Leu Pro Gln Phe Pro
 1045 1050 1055
 Ser Ser Leu Ala Thr Val Ser Ala Ser Val Gln Ser Val Pro Thr Gln
 1060 1065 1070
 Thr Ala Thr Leu Leu Pro Pro Ala Asn Pro Pro Leu Pro Gly Gly Pro
 1075 1080 1085
 Gly Ile Ala Ser Pro Cys Pro Thr Val Gln Leu Thr Val Glu Pro Val
 1090 1095 1100
 Gln Glu Glu Gln Ala Ser Gln Asp Lys Pro Pro Gly Leu Pro Gln Ser
 1105 1110 1115 1120
 Cys Glu Ser Tyr Gly Gly Ser Asp Val Thr Ser Gly Lys Glu Leu Ser
 1125 1130 1135
 Asp Ser Cys Glu Gly Ala Phe Gly Gly Gly Arg Leu Glu Gly Arg Ala
 1140 1145 1150
 Ala Arg Lys His His Arg Arg Ser Thr Arg Ala Arg Ser Arg Gln Glu
 1155 1160 1165
 Arg Ala Ser Arg Pro Arg Leu Thr Ile Leu Asn Val Cys Asn Thr Gly
 1170 1175 1180
 Asp Lys Met Val Glu Cys Gln Leu Glu Thr His Asn His Lys Met Val
 1185 1190 1195 1200
 Thr Phe Lys Phe Asp Leu Asp Gly Asp Ala Pro Asp Glu Ile Ala Thr
 1205 1210 1215
 Tyr Met Val Glu His Asp Phe Ile Leu Gln Ala Glu Arg Glu Thr Phe
 1220 1225 1230
 Ile Glu Gln Met Lys Asp Val Met Asp Lys Ala Glu Asp Met Leu Ser
 1235 1240 1245
 Glu Asp Thr Asp Ala Asp Arg Gly Ser Asp Pro Gly Thr Ser Pro Pro
 1250 1255 1260
 His Leu Ser Thr Cys Gly Leu Gly Thr Gly Glu Glu Ser Arg Gln Ser
 1265 1270 1275 1280

Gln Ala Asn Ala Pro Val Tyr Gln Gln Asn Val Leu His Thr Gly Lys
 1285 1290 1295
 Arg Trp Phe Ile Ile Cys Pro Val Ala Glu His Pro Ala Pro Glu Ala
 1300 1305 1310
 Pro Glu Ser Ser Pro Pro Leu Pro Leu Ser Ser Leu Pro Pro Glu Ala
 1315 1320 1325
 Ser Gln Asp Ser Ala Pro Tyr Lys Asp Gln Leu Ser Ser Lys Glu Gln
 1330 1335 1340
 Pro Ser Phe Leu Ala Ser Gln Gln Leu Leu Ser Gln Ala Gly Pro Ser
 1345 1350 1355 1360
 Asn Pro Pro Gly Ala Pro Pro Ala Pro Leu Ala Pro Ser Ser Pro Pro
 1365 1370 1375
 Val Thr Ala Leu Pro Gln Asp Gly Ala Ala Pro Ala Thr Ser Thr Met
 1380 1385 1390
 Pro Glu Pro Ala Ser Gly Thr Ala Ser Gln Ala Gly Gly Pro Gly Thr
 1395 1400 1405
 Pro Gln Gly Leu Thr Ser Glu Leu Glu Thr Ser Gln Pro Leu Ala Glu
 1410 1415 1420
 Thr His Glu Ala Pro Leu Ala Val Gln Pro Leu Val Val Gly Leu Ala
 1425 1430 1435 1440
 Pro Cys Thr Pro Ala Pro Glu Ala Ala Ser Thr Arg Asp Ala Ser Ala
 1445 1450 1455
 Pro Arg Glu Pro Leu Pro Pro Pro Ala Pro Glu Pro Ser Pro His Ser
 1460 1465 1470
 Gly Thr Pro Gln Pro Ala Leu Gly Gln Pro Ala Pro Leu Leu Pro Ala
 1475 1480 1485
 Ala Val Gly Ala Val Ser Leu Ala Thr Ser Gln Leu Pro Ser Pro Pro
 1490 1495 1500
 Leu Gly Pro Thr Val Pro Pro Gln Pro Pro Ser Ala Leu Glu Ser Asp
 1505 1510 1515 1520
 Gly Glu Gly Pro Pro Pro Arg Val Gly Phe Val Asp Ser Thr Ile Lys
 1525 1530 1535
 Ser Leu Asp Glu Lys Leu Arg Thr Leu Leu Tyr Gln Glu His Val Pro
 1540 1545 1550
 Thr Ser Ser Ala Ser Ala Gly Thr Pro Val Glu Val Gly Asp Arg Asp
 1555 1560 1565
 Phe Thr Leu Glu Pro Leu Arg Gly Asp Gln Pro Arg Ser Glu Val Cys
 1570 1575 1580
 Gly Gly Asp Leu Ala Leu Pro Pro Val Pro Lys Glu Ala Val Ser Gly
 1585 1590 1595 1600
 Arg Val Gln Leu Pro Gln Pro Leu Val Glu Lys Ser Glu Leu Ala Pro
 1605 1610 1615
 Thr Arg Gly Ala Val Met Glu Gln Gly Thr Ser Ser Ser Met Thr Glu
 1620 1625 1630
 Ser Ser Pro Arg Ser Met Leu Gly Tyr Asp Arg Asp Gly Arg Gln Val
 1635 1640 1645
 Ala Ser Asp Ser His Val Val Pro Ser Val Pro Gln Asp Val Pro Ala
 1650 1655 1660
 Phe Val Arg Pro Ala Arg Val Glu Pro Thr Asp Arg Asp Gly Gly Glu
 1665 1670 1675 1680
 Ala Gly Glu Ser Ser Ala Glu Pro Pro Pro Ser Asp Met Gly Thr Val
 1685 1690 1695
 Gly Gly Gln Ala Ser His Pro Gln Thr Leu Gly Ala Arg Ala Leu Gly
 1700 1705 1710
 Ser Pro Arg Lys Arg Pro Glu Gln Gln Asp Val Ser Ser Pro Ala Lys
 1715 1720 1725
 Thr Val Gly Arg Phe Ser Val Val Ser Thr Gln Asp Glu Trp Thr Leu
 1730 1735 1740
 Ala Ser Pro His Ser Leu Arg Tyr Ser Ala Pro Pro Asp Val Tyr Leu
 1745 1750 1755 1760

Asp Glu Ala Pro Ser Ser Pro Asp Val Lys Leu Ala Val Arg Arg Ala
 1765 1770 1775
 Gln Thr Ala Ser Ser Ile Glu Val Gly Val Gly Glu Pro Val Ser Ser
 1780 1785 1790
 Asp Ser Gly Asp Glu Gly Pro Arg Ala Arg Pro Pro Val Gln Lys Gln
 1795 1800 1805
 Ala Ser Leu Pro Val Ser Gly Ser Val Ala Gly Asp Phe Val Lys Lys
 1810 1815 1820
 Ala Thr Ala Phe Leu Gln Arg Pro Ser Arg Ala Gly Ser Leu Gly Pro
 1825 1830 1835 1840
 Glu Thr Pro Ser Arg Val Gly Met Lys Val Pro Thr Ile Ser Val Thr
 1845 1850 1855
 Ser Phe His Ser Gln Ser Ser Tyr Ile Ser Ser Asp Asn Asp Ser Glu
 1860 1865 1870
 Leu Glu Asp Ala Asp Ile Lys Lys Glu Leu Gln Ser Leu Arg Glu Lys
 1875 1880 1885
 His Leu Lys Glu Ile Ser Glu Leu Gln Ser Gln Gln Lys Gln Glu Ile
 1890 1895 1900
 Glu Ala Leu Tyr Arg Arg Leu Gly Lys Pro Leu Pro Pro Asn Val Gly
 1905 1910 1915 1920
 Phe Phe His Thr Ala Pro Pro Thr Gly Arg Arg Arg Lys Thr Ser Lys
 1925 1930 1935
 Ser Lys Leu Lys Ala Gly Lys Leu Leu Asn Pro Leu Val Arg Gln Leu
 1940 1945 1950
 Lys Val Val Ala Ser Ser Thr Gly His Leu Ala Asp Ser Ser Arg Gly
 1955 1960 1965
 Pro Pro Ala Lys Asp Pro Ala Gln Ala Ser Val Gly Leu Thr Ala Asp
 1970 1975 1980
 Ser Thr Gly Leu Ser Gly Lys Ala Val Gln Thr Gln Gln Pro Cys Ser
 1985 1990 1995 2000
 Val Arg Ala Ser Leu Ser Ser Asp Ile Cys Ser Gly Leu Ala Ser Asp
 2005 2010 2015
 Gly Gly Gly Ala Arg Gly Gln Gly Trp Thr Val Tyr His Pro Thr Ser
 2020 2025 2030
 Glu Arg Val Thr Tyr Lys Ser Ser Lys Pro Arg Ala Arg Phe Leu
 2035 2040 2045
 Ser Gly Pro Val Ser Val Ser Ile Trp Ser Ala Leu Lys Arg Leu Cys
 2050 2055 2060
 Leu Gly Lys Glu His Ser Ser Arg Ser Ser Thr Ser Ser Leu Ala Pro
 2065 2070 2075 2080
 Gly Pro Glu Pro Gly Pro Gln Pro Ala Leu His Val Gln Ala Gln Val
 2085 2090 2095
 Asn Asn Ser Asn Asn Lys Lys Gly Thr Phe Thr Asp Asp Leu His Lys
 2100 2105 2110
 Leu Val Asp Glu Trp Thr Ser Lys Thr Val Gly Ala Ala Gln Leu Lys
 2115 2120 2125
 Pro Thr Leu Asn Gln Leu Lys Gln Thr Gln Lys Leu Gln Asp Met Glu
 2130 2135 2140
 Ala Gln Ala Gly Trp Ala Ala Pro Gly Glu Ala Arg Ala Met Thr Ala
 2145 2150 2155 2160
 Pro Arg Ala Gly Val Gly Met Pro Arg Leu Pro Pro Ala Pro Gly Pro
 2165 2170 2175
 Leu Ser Thr Thr Val Ile Pro Gly Ala Ala Pro Thr Leu Ser Val Pro
 2180 2185 2190
 Thr Pro Asp Gly Ala Leu Gly Thr Ala Arg Arg Asn Gln Val Trp Phe
 2195 2200 2205
 Gly Leu Arg Val Pro Pro Thr Ala Cys Cys Gly His Ser Thr Gln Pro
 2210 2215 2220
 Arg Gly Gly Gln Arg Val Gly Ser Lys Thr Ala Ser Phe Ala Ala Ser
 2225 2230 2235 2240

Asp Pro Val Arg Ser
2245

<210> 4
<211> 505
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 4
Met Gly Leu Val Ser Ser Lys Lys Pro Asp Lys Glu Lys Pro Ile Lys
1 5 10 15
Glu Lys Asp Lys Gly Gln Trp Ser Pro Leu Lys Val Ser Ala Gln Asp
20 25 30
Lys Asp Ala Pro Pro Leu Pro Pro Leu Val Val Phe Asn His Leu Thr
35 40 45
Pro Pro Pro Pro Asp Glu His Leu Asp Glu Asp Lys His Phe Val Val
50 55 60
Ala Leu Tyr Asp Tyr Thr Ala Met Asn Asp Arg Asp Leu Gln Met Leu
65 70 75 80
Lys Gly Glu Lys Leu Gln Val Leu Lys Gly Thr Gly Asp Trp Trp Leu
85 90 95
Ala Arg Ser Leu Val Thr Gly Arg Glu Gly Tyr Val Pro Ser Asn Phe
100 105 110
Val Ala Arg Val Glu Ser Leu Glu Met Glu Arg Trp Phe Phe Arg Ser
115 120 125
Gln Gly Arg Lys Glu Ala Glu Arg Gln Leu Leu Ala Pro Ile Asn Lys
130 135 140
Ala Gly Ser Phe Leu Ile Arg Glu Ser Glu Thr Asn Lys Gly Ala Phe
145 150 155 160
Ser Leu Ser Val Lys Asp Val Thr Thr Gln Gly Glu Leu Ile Lys His
165 170 175
Tyr Lys Ile Arg Cys Leu Asp Glu Gly Gly Tyr Tyr Ile Ser Pro Arg
180 185 190
Ile Thr Phe Pro Ser Leu Gln Ala Leu Val Gln His Tyr Ser Lys Lys
195 200 205
Gly Asp Gly Leu Cys Gln Arg Leu Thr Leu Pro Cys Val Arg Pro Ala
210 215 220
Pro Gln Asn Pro Trp Ala Gln Asp Glu Trp Glu Ile Pro Arg Gln Ser
225 230 235 240
Leu Arg Leu Val Arg Lys Leu Gly Ser Gly Gln Phe Gly Glu Val Trp
245 250 255
Met Gly Tyr Tyr Lys Asn Asn Met Lys Val Ala Ile Lys Thr Leu Lys
260 265 270
Glu Gly Thr Met Ser Pro Glu Ala Phe Leu Gly Glu Ala Asn Val Met
275 280 285
Lys Ala Leu Gln His Glu Arg Leu Val Arg Leu Tyr Ala Val Val Thr
290 295 300
Lys Glu Pro Ile Tyr Ile Val Thr Glu Tyr Met Ala Arg Gly Cys Leu
305 310 315 320
Leu Asp Phe Leu Lys Thr Asp Glu Gly Ser Arg Leu Ser Leu Pro Arg
325 330 335
Leu Ile Asp Met Ser Ala Gln Ile Ala Glu Gly Met Ala Tyr Ile Glu
340 345 350
Arg Met Asn Ser Ile His Arg Asp Leu Arg Ala Ala Asn Ile Leu Val
355 360 365
Ser Glu Ala Leu Cys Cys Lys Ile Ala Asp Phe Gly Leu Ala Arg Ile
370 375 380

```

Ile Asp Ser Glu Tyr Thr Ala Gln Glu Gly Ala Lys Phe Pro Ile Lys
385                      390                      395                      400
Trp Thr Ala Pro Glu Ala Tyr His Phe Gly Val Phe Thr Ile Lys Ala
                      405                      410                      415
Asp Val Trp Ser Phe Gly Val Leu Leu Met Glu Val Val Thr Tyr Gly
                      420                      425                      430
Arg Val Pro Tyr Pro Gly Met Ser Asn Pro Glu Val Ile Arg Asn Leu
                      435                      440                      445
Glu Arg Gly Tyr Arg Met Pro Arg Pro Asp Thr Cys Pro Pro Glu Leu
                      450                      455                      460
Tyr Arg Gly Val Ile Ala Glu Cys Trp Arg Ser Arg Pro Glu Glu Arg
465                      470                      475                      480
Pro Thr Phe Glu Phe Leu Gln Ser Val Leu Glu Asp Phe Tyr Thr Ala
                      485                      490                      495
Thr Glu Arg Gln Tyr Glu Leu Gln Pro
                    500                      505

```

<210> 5
 <211> 972
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

```

<400> 5
Met Gly Pro Gly Val Leu Leu Leu Leu Leu Val Ala Thr Ala Trp His
1                      5                      10                      15
Gly Gln Gly Ile Pro Val Ile Glu Pro Ser Val Pro Glu Leu Val Val
                20                      25                      30
Lys Pro Gly Ala Thr Val Thr Leu Arg Cys Val Gly Asn Gly Ser Val
                35                      40                      45
Glu Trp Asp Gly Pro Pro Ser Pro His Trp Thr Leu Tyr Ser Asp Gly
50                      55                      60
Ser Ser Ser Ile Leu Ser Thr Asn Asn Ala Thr Phe Gln Asn Thr Gly
65                      70                      75                      80
Thr Tyr Arg Cys Thr Glu Pro Gly Asp Pro Leu Gly Gly Ser Ala Ala
                85                      90                      95
Ile His Leu Tyr Val Lys Asp Pro Ala Arg Pro Trp Asn Val Leu Ala
                100                      105                      110
Gln Glu Val Val Val Phe Glu Asp Gln Asp Ala Leu Leu Pro Cys Leu
                115                      120                      125
Leu Thr Asp Pro Val Leu Glu Ala Gly Val Ser Leu Val Arg Val Arg
130                      135                      140
Gly Arg Pro Leu Met Arg His Thr Asn Tyr Ser Phe Ser Pro Trp His
145                      150                      155                      160
Gly Phe Thr Ile His Arg Ala Lys Phe Ile Gln Ser Gln Asp Tyr Gln
                165                      170                      175
Cys Ser Ala Leu Met Gly Gly Arg Lys Val Met Ser Ile Ser Ile Arg
                180                      185                      190
Leu Lys Val Gln Lys Val Ile Pro Gly Pro Pro Ala Leu Thr Leu Val
195                      200                      205
Pro Ala Glu Leu Val Arg Ile Arg Gly Glu Ala Ala Gln Ile Val Cys
210                      215                      220
Ser Ala Ser Ser Val Asp Val Asn Phe Asp Val Phe Leu Gln His Asn
225                      230                      235                      240
Asn Thr Lys Leu Ala Ile Pro Gln Gln Ser Asp Phe His Asn Asn Arg
                245                      250                      255
Tyr Gln Lys Val Leu Thr Leu Asn Leu Asp Gln Val Asp Phe Gln His
260                      265                      270

```

Ala	Gly	Asn	Tyr	Ser	Cys	Val	Ala	Ser	Asn	Val	Gln	Gly	Lys	His	Ser	275	280	285
Thr	Ser	Met	Phe	Phe	Arg	Val	Val	Glu	Ser	Ala	Tyr	Leu	Asn	Leu	Ser	290	295	300
Ser	Glu	Gln	Asn	Leu	Ile	Gln	Glu	Val	Thr	Val	Gly	Glu	Gly	Leu	Asn	305	310	315
Leu	Lys	Val	Met	Val	Glu	Ala	Tyr	Pro	Gly	Leu	Gln	Gly	Phe	Asn	Trp	320	325	330
Thr	Tyr	Leu	Gly	Pro	Phe	Ser	Asp	His	Gln	Pro	Glu	Pro	Lys	Leu	Ala	335	340	345
Asn	Ala	Thr	Thr	Lys	Asp	Thr	Tyr	Arg	His	Thr	Phe	Thr	Leu	Ser	Leu	350	355	360
Pro	Arg	Leu	Lys	Pro	Ser	Glu	Ala	Gly	Arg	Tyr	Ser	Phe	Leu	Ala	Arg	365	370	375
Asn	Pro	Gly	Gly	Trp	Arg	Ala	Leu	Thr	Phe	Glu	Leu	Thr	Leu	Arg	Tyr	380	385	390
Pro	Pro	Glu	Val	Ser	Val	Ile	Trp	Thr	Phe	Ile	Asn	Gly	Ser	Gly	Thr	395	400	405
Leu	Leu	Cys	Ala	Ala	Ser	Gly	Tyr	Pro	Gln	Pro	Asn	Val	Thr	Trp	Leu	410	415	420
Gln	Cys	Ser	Gly	His	Thr	Asp	Arg	Cys	Asp	Glu	Ala	Gln	Val	Leu	Gln	425	430	435
Val	Trp	Asp	Asp	Pro	Tyr	Pro	Glu	Val	Leu	Ser	Gln	Glu	Pro	Phe	His	440	445	450
Lys	Val	Thr	Val	Gln	Ser	Leu	Leu	Thr	Val	Glu	Thr	Leu	Glu	His	Asn	455	460	465
Gln	Thr	Tyr	Glu	Cys	Arg	Ala	His	Asn	Ser	Val	Gly	Ser	Gly	Ser	Trp	470	475	480
Ala	Phe	Ile	Pro	Ile	Ser	Ala	Gly	Ala	His	Thr	His	Pro	Pro	Asp	Glu	485	490	495
Phe	Leu	Phe	Thr	Pro	Val	Val	Val	Ala	Cys	Met	Ser	Ile	Met	Ala	Leu	500	505	510
Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Leu	Tyr	Lys	Tyr	Lys	Gln	Lys	Pro	515	520	525
Lys	Tyr	Gln	Val	Arg	Trp	Lys	Ile	Ile	Glu	Ser	Tyr	Glu	Gly	Asn	Ser	530	535	540
Tyr	Thr	Phe	Ile	Asp	Pro	Thr	Gln	Leu	Pro	Tyr	Asn	Glu	Lys	Trp	Glu	545	550	555
Phe	Pro	Arg	Asn	Asn	Leu	Gln	Phe	Gly	Lys	Thr	Leu	Gly	Ala	Gly	Ala	560	565	570
Phe	Gly	Lys	Val	Val	Glu	Ala	Thr	Ala	Phe	Gly	Leu	Gly	Lys	Glu	Asp	575	580	585
Ala	Val	Leu	Lys	Val	Ala	Val	Lys	Met	Leu	Lys	Ser	Thr	Ala	His	Ala	590	595	600
Asp	Glu	Lys	Glu	Ala	Leu	Met	Ser	Glu	Leu	Lys	Ile	Met	Ser	His	Leu	605	610	615
Gly	Gln	His	Glu	Asn	Ile	Val	Asn	Leu	Leu	Gly	Ala	Cys	Thr	His	Gly	620	625	630
Gly	Pro	Val	Leu	Val	Ile	Thr	Glu	Tyr	Cys	Cys	Tyr	Gly	Asp	Leu	Leu	635	640	645
Asn	Phe	Leu	Arg	Arg	Lys	Ala	Glu	Ala	Met	Leu	Gly	Pro	Ser	Leu	Ser	650	655	660
Pro	Gly	Gln	Asp	Pro	Glu	Gly	Val	Asp	Tyr	Lys	Asn	Ile	His	Leu		665	670	675
Glu	Lys	Lys	Tyr	Val	Arg	Arg	Asp	Ser	Gly	Phe	Ser	Ser	Gln	Gly	Val	680	685	690
Asp	Thr	Tyr	Val	Glu	Met	Arg	Pro	Val	Ser	Thr	Ser	Ser	Asn	Asp	Ser	695	700	705
Phe	Ser	Glu	Gln	Asp	Leu	Asp	Lys	Glu	Asp	Gly	Arg	Pro	Leu	Glu	Leu	710	715	720
																725	730	735
																740	745	750

```

Arg Asp Leu Leu His Phe Ser Ser Gln Val Ala Gln Gly Met Ala Phe
      755                      760                      765
Leu Ala Ser Lys Asn Cys Ile His Arg Asp Val Ala Ala Arg Asn Val
      770                      775                      780
Leu Leu Thr Asn Gly His Val Ala Lys Ile Gly Asp Phe Gly Leu Ala
785                      790                      795                      800
Arg Asp Ile Met Asn Asp Ser Asn Tyr Ile Val Lys Gly Asn Ala Arg
      805                      810                      815
Leu Pro Val Lys Trp Met Ala Pro Glu Ser Ile Phe Asp Cys Val Tyr
      820                      825                      830
Thr Val Gln Ser Asp Val Trp Ser Tyr Gly Ile Leu Leu Trp Glu Ile
      835                      840                      845
Phe Ser Leu Gly Leu Asn Pro Tyr Pro Gly Ile Leu Val Asn Ser Lys
      850                      855                      860
Phe Tyr Lys Leu Val Lys Asp Gly Tyr Gln Met Ala Gln Pro Ala Phe
865                      870                      875                      880
Ala Pro Lys Asn Ile Tyr Ser Ile Met Gln Ala Cys Trp Ala Leu Glu
      885                      890                      895
Pro Thr His Arg Pro Thr Phe Gln Gln Ile Cys Ser Phe Leu Gln Glu
      900                      905                      910
Gln Ala Gln Glu Asp Arg Arg Glu Arg Asp Tyr Thr Asn Leu Pro Ser
      915                      920                      925
Ser Ser Arg Ser Gly Gly Ser Gly Ser Ser Ser Ser Glu Leu Glu Glu
      930                      935                      940
Glu Ser Ser Ser Glu His Leu Thr Cys Cys Glu Gln Gly Asp Ile Ala
945                      950                      955                      960
Gln Pro Leu Leu Gln Pro Asn Asn Tyr Gln Phe Cys
      965                      970

```

<210> 6
 <211> 450
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

```

<400> 6
Met Ser Ala Ile Gln Ala Ala Trp Pro Ser Gly Thr Glu Cys Ile Ala
 1                      5                      10                      15
Lys Tyr Asn Phe His Gly Thr Ala Glu Gln Asp Leu Pro Phe Cys Lys
      20                      25                      30
Gly Asp Val Leu Thr Ile Val Ala Val Thr Lys Asp Pro Asn Trp Tyr
      35                      40                      45
Lys Ala Lys Asn Lys Val Gly Arg Glu Gly Ile Ile Pro Ala Asn Tyr
      50                      55                      60
Val Gln Lys Arg Glu Gly Val Lys Ala Gly Thr Lys Leu Ser Leu Met
      65                      70                      75                      80
Pro Trp Phe His Gly Lys Ile Thr Arg Glu Gln Ala Glu Arg Leu Leu
      85                      90                      95
Tyr Pro Pro Glu Thr Gly Leu Phe Leu Val Arg Glu Ser Thr Asn Tyr
      100                      105                      110
Pro Gly Asp Tyr Thr Leu Cys Val Ser Cys Asp Gly Lys Val Glu His
      115                      120                      125
Tyr Arg Ile Met Tyr His Ala Ser Lys Leu Ser Ile Asp Glu Glu Val
      130                      135                      140
Tyr Phe Glu Asn Leu Met Gln Leu Val Glu His Tyr Thr Ser Asp Ala
      145                      150                      155                      160
Asp Gly Leu Cys Thr Arg Leu Ile Lys Pro Lys Val Met Glu Gly Thr
      165                      170                      175

```

Val Ala Ala Gln Asp Glu Phe Tyr Arg Ser Gly Trp Ala Leu Asn Met
 180 185 190
 Lys Glu Leu Lys Leu Leu Gln Thr Ile Gly Lys Gly Glu Phe Gly Asp
 195 200 205
 Val Met Leu Gly Asp Tyr Arg Gly Asn Lys Val Ala Val Lys Cys Ile
 210 215 220
 Lys Asn Asp Ala Thr Ala Gln Ala Phe Leu Ala Glu Ala Ser Val Met
 225 230 235 240
 Thr Gln Leu Arg His Ser Asn Leu Val Gln Leu Leu Gly Val Ile Val
 245 250 255
 Glu Glu Lys Gly Gly Leu Tyr Ile Val Thr Glu Tyr Met Ala Lys Gly
 260 265 270
 Ser Leu Val Asp Tyr Leu Arg Ser Arg Gly Arg Ser Val Leu Gly Gly
 275 280 285
 Asp Cys Leu Leu Lys Phe Ser Leu Asp Val Cys Glu Ala Met Glu Tyr
 290 295 300
 Leu Glu Gly Asn Asn Phe Val His Arg Asp Leu Ala Ala Arg Asn Val
 305 310 315 320
 Leu Val Ser Glu Asp Asn Val Ala Lys Val Ser Asp Phe Gly Leu Thr
 325 330 335
 Lys Glu Ala Ser Ser Thr Gln Asp Thr Gly Lys Leu Pro Val Lys Trp
 340 345 350
 Thr Ala Pro Glu Ala Leu Arg Glu Lys Lys Phe Ser Thr Lys Ser Asp
 355 360 365
 Val Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile Tyr Ser Phe Gly Arg
 370 375 380
 Val Pro Tyr Pro Arg Ile Pro Leu Lys Asp Val Val Pro Arg Val Glu
 385 390 395 400
 Lys Gly Tyr Lys Met Asp Ala Pro Asp Gly Cys Pro Pro Ala Val Tyr
 405 410 415
 Glu Val Met Lys Asn Cys Trp His Leu Asp Ala Ala Met Arg Pro Ser
 420 425 430
 Phe Leu Gln Leu Arg Glu Gln Leu Glu His Ile Lys Thr His Glu Leu
 435 440 445
 His Leu
 450

<210> 7
 <211> 1052
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 7
 Met Ala Ala Ala Tyr Leu Asp Pro Asn Leu Asn His Thr Pro Asn Ser
 1 5 10 15
 Ser Thr Lys Thr His Leu Gly Thr Gly Met Glu Arg Ser Pro Gly Ala
 20 25 30
 Met Glu Arg Val Leu Lys Val Phe His Tyr Phe Glu Ser Asn Ser Glu
 35 40 45
 Pro Thr Thr Trp Ala Ser Ile Ile Arg His Gly Asp Ala Thr Asp Val
 50 55 60
 Arg Gly Ile Ile Gln Lys Ile Val Asp Ser His Lys Val Lys His Val
 65 70 75 80
 Ala Cys Tyr Gly Phe Arg Leu Ser His Leu Arg Ser Glu Glu Val His
 85 90 95
 Trp Leu His Val Asp Met Gly Val Ser Ser Val Arg Glu Lys Tyr Glu
 100 105 110

Leu	Ala	His	Pro	Pro	Glu	Glu	Trp	Lys	Tyr	Glu	Leu	Arg	Ile	Arg	Tyr	115	120	125
Leu	Pro	Lys	Gly	Phe	Leu	Asn	Gln	Phe	Thr	Glu	Asp	Lys	Pro	Thr	Leu	130	135	140
Asn	Phe	Phe	Tyr	Gln	Gln	Val	Lys	Ser	Asp	Tyr	Met	Leu	Glu	Ile	Ala	145	150	155
Asp	Gln	Val	Asp	Gln	Glu	Ile	Ala	Leu	Lys	Leu	Gly	Cys	Leu	Glu	Ile	165	170	175
Arg	Arg	Ser	Tyr	Trp	Glu	Met	Arg	Gly	Asn	Ala	Leu	Glu	Lys	Lys	Ser	180	185	190
Asn	Tyr	Glu	Val	Leu	Glu	Lys	Asp	Val	Gly	Leu	Lys	Arg	Phe	Phe	Pro	195	200	205
Lys	Ser	Leu	Leu	Asp	Ser	Val	Lys	Ala	Lys	Thr	Leu	Arg	Lys	Leu	Ile	210	215	220
Gln	Gln	Thr	Phe	Arg	Gln	Phe	Ala	Asn	Leu	Asn	Arg	Glu	Glu	Ser	Ile	225	230	235
Leu	Lys	Phe	Phe	Glu	Ile	Leu	Ser	Pro	Val	Tyr	Arg	Phe	Asp	Lys	Glu	245	250	255
Cys	Phe	Lys	Cys	Ala	Leu	Gly	Ser	Ser	Trp	Ile	Ile	Ser	Val	Glu	Leu	260	265	270
Ala	Ile	Gly	Pro	Glu	Glu	Gly	Ile	Ser	Tyr	Leu	Thr	Asp	Lys	Gly	Cys	275	280	285
Asn	Pro	Thr	His	Leu	Ala	Asp	Phe	Thr	Gln	Val	Gln	Thr	Ile	Gln	Tyr	290	295	300
Ser	Asn	Ser	Glu	Asp	Lys	Asp	Arg	Lys	Gly	Met	Leu	Gln	Leu	Lys	Ile	305	310	315
Ala	Gly	Ala	Pro	Glu	Pro	Leu	Thr	Val	Thr	Ala	Pro	Ser	Leu	Thr	Ile	325	330	335
Ala	Glu	Asn	Met	Ala	Asp	Leu	Ile	Asp	Gly	Tyr	Cys	Arg	Leu	Val	Asn	340	345	350
Gly	Thr	Ser	Gln	Ser	Phe	Ile	Ile	Arg	Pro	Gln	Lys	Glu	Gly	Glu	Arg	355	360	365
Ala	Leu	Pro	Ser	Ile	Pro	Lys	Leu	Ala	Asn	Ser	Glu	Lys	Gln	Gly	Met	370	375	380
Arg	Thr	His	Ala	Val	Ser	Val	Ser	Glu	Thr	Asp	Asp	Tyr	Ala	Glu	Ile	385	390	395
Ile	Asp	Glu	Glu	Asp	Thr	Tyr	Thr	Met	Pro	Ser	Thr	Arg	Asp	Tyr	Glu	405	410	415
Ile	Gln	Arg	Glu	Arg	Ile	Glu	Leu	Gly	Arg	Cys	Ile	Gly	Glu	Gly	Gln	420	425	430
Phe	Gly	Asp	Val	His	Gln	Gly	Ile	Tyr	Met	Ser	Pro	Glu	Asn	Pro	Ala	435	440	445
Leu	Ala	Val	Ala	Ile	Lys	Thr	Cys	Lys	Asn	Cys	Thr	Ser	Asp	Ser	Val	450	455	460
Arg	Glu	Lys	Phe	Leu	Gln	Glu	Ala	Leu	Thr	Met	Arg	Gln	Phe	Asp	His	465	470	475
Pro	His	Ile	Val	Lys	Leu	Ile	Gly	Val	Ile	Thr	Glu	Asn	Pro	Val	Trp	485	490	495
Ile	Ile	Met	Glu	Leu	Cys	Thr	Leu	Gly	Glu	Leu	Arg	Ser	Phe	Leu	Gln	500	505	510
Val	Arg	Lys	Tyr	Ser	Leu	Asp	Leu	Ala	Ser	Leu	Ile	Leu	Tyr	Ala	Tyr	515	520	525
Gln	Leu	Ser	Thr	Ala	Leu	Ala	Tyr	Leu	Glu	Ser	Lys	Arg	Phe	Val	His	530	535	540
Arg	Asp	Ile	Ala	Ala	Arg	Asn	Val	Leu	Val	Ser	Ser	Asn	Asp	Cys	Val	545	550	555
Lys	Leu	Gly	Asp	Phe	Gly	Leu	Ser	Arg	Tyr	Met	Glu	Asp	Ser	Thr	Tyr	565	570	575
Tyr	Lys	Ala	Ser	Lys	Gly	Lys	Leu	Pro	Ile	Lys	Trp	Met	Ala	Pro	Glu	580	585	590

```

Ser Ile Asn Phe Arg Arg Phe Thr Ser Ala Ser Asp Val Trp Met Phe
      595                      600                      605
Gly Val Cys Met Trp Glu Ile Leu Met His Gly Val Lys Pro Phe Gln
      610                      615                      620
Gly Val Lys Asn Asn Asp Val Ile Gly Arg Ile Glu Asn Gly Glu Arg
625                      630                      635                      640
Leu Pro Met Pro Pro Asn Cys Pro Pro Thr Leu Tyr Ser Leu Met Thr
      645                      650                      655
Lys Cys Trp Ala Tyr Asp Pro Ser Arg Arg Pro Arg Phe Thr Glu Leu
      660                      665                      670
Lys Ala Gln Leu Ser Thr Ile Leu Glu Glu Glu Lys Ala Gln Gln Glu
      675                      680                      685
Glu Arg Met Arg Met Glu Ser Arg Arg Gln Ala Thr Val Ser Trp Asp
690                      695                      700
Ser Gly Gly Ser Asp Glu Ala Pro Pro Lys Pro Ser Arg Pro Gly Tyr
705                      710                      715                      720
Pro Ser Pro Arg Ser Ser Glu Gly Phe Tyr Pro Ser Pro Gln His Met
      725                      730                      735
Val Gln Thr Asn His Tyr Gln Val Ser Gly Tyr Pro Gly Ser His Gly
      740                      745                      750
Ile Thr Ala Met Ala Gly Ser Ile Tyr Pro Gly Gln Ala Ser Leu Leu
      755                      760                      765
Asp Gln Thr Asp Ser Trp Asn His Arg Pro Gln Glu Ile Ala Met Trp
770                      775                      780
Gln Pro Asn Val Glu Asp Ser Thr Val Leu Asp Leu Arg Gly Ile Gly
785                      790                      795                      800
Gln Val Leu Pro Thr His Leu Met Glu Glu Arg Leu Ile Arg Gln Gln
      805                      810                      815
Gln Glu Met Glu Glu Asp Gln Arg Trp Leu Glu Lys Glu Glu Arg Phe
      820                      825                      830
Leu Lys Pro Asp Val Arg Leu Ser Arg Gly Ser Ile Asp Arg Glu Asp
835                      840                      845
Gly Ser Leu Gln Gly Pro Ile Gly Asn Gln His Ile Tyr Gln Pro Val
850                      855                      860
Gly Lys Pro Asp Pro Ala Ala Pro Pro Lys Lys Pro Pro Arg Pro Gly
865                      870                      875                      880
Ala Pro Gly His Leu Gly Ser Leu Ala Ser Leu Ser Ser Pro Ala Asp
      885                      890                      895
Ser Tyr Asn Glu Gly Val Lys Leu Gln Pro Gln Glu Ile Ser Pro Pro
900                      905                      910
Pro Thr Ala Asn Leu Asp Arg Ser Asn Asp Lys Val Tyr Glu Asn Val
915                      920                      925
Thr Gly Leu Val Lys Ala Val Ile Glu Met Ser Ser Lys Ile Gln Pro
930                      935                      940
Ala Pro Pro Glu Glu Tyr Val Pro Met Val Lys Glu Val Gly Leu Ala
945                      950                      955                      960
Leu Arg Thr Leu Leu Ala Thr Val Asp Glu Thr Ile Pro Leu Leu Pro
      965                      970                      975
Ala Ser Thr His Arg Glu Ile Glu Met Ala Gln Lys Leu Leu Asn Ser
980                      985                      990
Asp Leu Gly Glu Leu Ile Asn Lys Met Lys Leu Ala Gln Gln Tyr Val
995                      1000                      1005
Met Thr Ser Leu Gln Gln Glu Tyr Lys Lys Gln Met Leu Thr Ala Ala
1010                      1015                      1020
His Ala Leu Ala Val Asp Ala Lys Asn Leu Leu Asp Val Ile Asp Gln
1025                      1030                      1035                      1040
Ala Arg Leu Lys Met Leu Gly Gln Thr Arg Pro His
      1045                      1050

```

<210> 8

<211> 821

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 8

```

Met Val Ser Trp Gly Arg Phe Ile Cys Leu Val Val Val Thr Met Ala
 1          5          10          15
Thr Leu Ser Leu Ala Arg Pro Ser Phe Ser Leu Val Glu Asp Thr Thr
          20          25          30
Leu Glu Pro Glu Glu Pro Pro Thr Lys Tyr Gln Ile Ser Gln Pro Glu
          35          40          45
Val Tyr Val Ala Ala Pro Gly Glu Ser Leu Glu Val Arg Cys Leu Leu
          50          55          60
Lys Asp Ala Ala Val Ile Ser Trp Thr Lys Asp Gly Val His Leu Gly
65          70          75          80
Pro Asn Asn Arg Thr Val Leu Ile Gly Glu Tyr Leu Gln Ile Lys Gly
          85          90          95
Ala Thr Pro Arg Asp Ser Gly Leu Tyr Ala Cys Thr Ala Ser Arg Thr
          100          105          110
Val Asp Ser Glu Thr Trp Tyr Phe Met Val Asn Val Thr Asp Ala Ile
          115          120          125
Ser Ser Gly Asp Asp Glu Asp Asp Thr Asp Gly Ala Glu Asp Phe Val
          130          135          140
Ser Glu Asn Ser Asn Asn Lys Arg Ala Pro Tyr Trp Thr Asn Thr Glu
145          150          155          160
Lys Met Glu Lys Arg Leu His Ala Val Pro Ala Ala Asn Thr Val Lys
          165          170          175
Phe Arg Cys Pro Ala Gly Gly Asn Pro Met Pro Thr Met Arg Trp Leu
          180          185          190
Lys Asn Gly Lys Glu Phe Lys Gln Glu His Arg Ile Gly Gly Tyr Lys
          195          200          205
Val Arg Asn Gln His Trp Ser Leu Ile Met Glu Ser Val Val Pro Ser
          210          215          220
Asp Lys Gly Asn Tyr Thr Cys Val Val Glu Asn Glu Tyr Gly Ser Ile
225          230          235          240
Asn His Thr Tyr His Leu Asp Val Val Glu Arg Ser Pro His Arg Pro
          245          250          255
Ile Leu Gln Ala Gly Leu Pro Ala Asn Ala Ser Thr Val Val Gly Gly
          260          265          270
Asp Val Glu Phe Val Cys Lys Val Tyr Ser Asp Ala Gln Pro His Ile
          275          280          285
Gln Trp Ile Lys His Val Glu Lys Asn Gly Ser Lys Tyr Gly Pro Asp
          290          295          300
Gly Leu Pro Tyr Leu Lys Val Leu Lys Ala Ala Gly Val Asn Thr Thr
305          310          315          320
Asp Lys Glu Ile Glu Val Leu Tyr Ile Arg Asn Val Thr Phe Glu Asp
          325          330          335
Ala Gly Glu Tyr Thr Cys Leu Ala Gly Asn Ser Ile Gly Ile Ser Phe
          340          345          350
His Ser Ala Trp Leu Thr Val Leu Pro Ala Pro Gly Arg Glu Lys Glu
          355          360          365
Ile Thr Ala Ser Pro Asp Tyr Leu Glu Ile Ala Ile Tyr Cys Ile Gly
          370          375          380
Val Phe Leu Ile Ala Cys Met Val Val Thr Val Ile Leu Cys Arg Met
385          390          395          400
Lys Asn Thr Thr Lys Lys Pro Asp Phe Ser Ser Gln Pro Ala Val His
          405          410          415

```

```

Lys Leu Thr Lys Arg Ile Pro Leu Arg Arg Gln Val Thr Val Ser Ala
      420      425      430
Glu Ser Ser Ser Ser Met Asn Ser Asn Thr Pro Leu Val Arg Ile Thr
      435      440      445
Thr Arg Leu Ser Ser Thr Ala Asp Thr Pro Met Leu Ala Gly Val Ser
      450      455      460
Glu Tyr Glu Leu Pro Glu Asp Pro Lys Trp Glu Phe Pro Arg Asp Lys
465      470      475      480
Leu Thr Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val Val
      485      490      495
Met Ala Glu Ala Val Gly Ile Asp Lys Asp Lys Pro Lys Glu Ala Val
      500      505      510
Thr Val Ala Val Lys Met Leu Lys Asp Asp Ala Thr Glu Lys Asp Leu
      515      520      525
Ser Asp Leu Val Ser Glu Met Glu Met Met Lys Met Ile Gly Lys His
      530      535      540
Lys Asn Ile Ile Asn Leu Leu Gly Ala Cys Thr Gln Asp Gly Pro Leu
545      550      555      560
Tyr Val Ile Val Glu Tyr Ala Ser Lys Gly Asn Leu Arg Glu Tyr Leu
      565      570      575
Arg Ala Arg Arg Pro Pro Gly Met Glu Tyr Ser Tyr Asp Ile Asn Arg
      580      585      590
Val Pro Glu Glu Gln Met Thr Phe Lys Asp Leu Val Ser Cys Thr Tyr
      595      600      605
Gln Leu Ala Arg Gly Met Glu Tyr Leu Ala Ser Gln Lys Cys Ile His
      610      615      620
Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asn Asn Val Met
625      630      635      640
Lys Ile Ala Asp Phe Gly Leu Ala Arg Asp Ile Asn Asn Ile Asp Tyr
      645      650      655
Tyr Lys Lys Thr Thr Asn Gly Arg Leu Pro Val Lys Trp Met Ala Pro
      660      665      670
Glu Ala Leu Phe Asp Arg Val Tyr Thr His Gln Ser Asp Val Trp Ser
      675      680      685
Phe Gly Val Leu Met Trp Glu Ile Phe Thr Leu Gly Gly Ser Pro Tyr
      690      695      700
Pro Gly Ile Pro Val Glu Glu Leu Phe Lys Leu Lys Glu Gly His
705      710      715      720
Arg Met Asp Lys Pro Ala Asn Cys Thr Asn Glu Leu Tyr Met Met Met
      725      730      735
Arg Asp Cys Trp His Ala Val Pro Ser Gln Arg Pro Thr Phe Lys Gln
      740      745      750
Leu Val Glu Asp Leu Asp Arg Ile Leu Thr Leu Thr Thr Asn Glu Glu
      755      760      765
Tyr Leu Asp Leu Ser Gln Pro Leu Glu Gln Tyr Ser Pro Ser Tyr Pro
      770      775      780
Asp Thr Arg Ser Ser Cys Ser Ser Gly Asp Asp Ser Val Phe Ser Pro
785      790      795      800
Asp Pro Met Pro Tyr Glu Pro Cys Leu Pro Gln Tyr Pro His Ile Asn
      805      810      815
Gly Ser Val Lys Thr
      820

```

<210> 9
 <211> 802
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 9

```

Met Arg Leu Leu Leu Ala Leu Leu Gly Val Leu Leu Ser Val Pro Gly
 1          5          10          15
Pro Pro Val Leu Ser Leu Glu Ala Ser Glu Glu Val Glu Leu Glu Pro
      20          25          30
Cys Leu Ala Pro Ser Leu Glu Gln Gln Glu Gln Glu Leu Thr Val Ala
      35          40          45
Leu Gly Gln Pro Val Arg Leu Cys Cys Gly Arg Ala Glu Arg Gly Gly
 50          55          60
His Trp Tyr Lys Glu Gly Ser Arg Leu Ala Pro Ala Gly Arg Val Arg
65          70          75          80
Gly Trp Arg Gly Arg Leu Glu Ile Ala Ser Phe Leu Pro Glu Asp Ala
      85          90          95
Gly Arg Tyr Leu Cys Leu Ala Arg Gly Ser Met Ile Val Leu Gln Asn
      100          105          110
Leu Thr Leu Ile Thr Gly Asp Ser Leu Thr Ser Ser Asn Asp Asp Glu
      115          120          125
Asp Pro Lys Ser His Arg Asp Pro Ser Asn Arg His Ser Tyr Pro Gln
      130          135          140
Gln Ala Pro Tyr Trp Thr His Pro Gln Arg Met Glu Lys Lys Leu His
      145          150          155          160
Ala Val Pro Ala Gly Asn Thr Val Lys Phe Arg Cys Pro Ala Ala Gly
      165          170          175
Asn Pro Thr Pro Thr Ile Arg Trp Leu Lys Asp Gly Gln Ala Phe His
      180          185          190
Gly Glu Asn Arg Ile Gly Gly Ile Arg Leu Arg His Gln His Trp Ser
      195          200          205
Leu Val Met Glu Ser Val Val Pro Ser Asp Arg Gly Thr Tyr Thr Cys
      210          215          220
Leu Val Glu Asn Ala Val Gly Ser Ile Arg Tyr Asn Tyr Leu Leu Asp
      225          230          235          240
Val Leu Glu Arg Ser Pro His Arg Pro Ile Leu Gln Ala Gly Leu Pro
      245          250          255
Ala Asn Thr Thr Ala Val Val Gly Ser Asp Val Glu Leu Leu Cys Lys
      260          265          270
Val Tyr Ser Asp Ala Gln Pro His Ile Gln Trp Leu Lys His Ile Val
      275          280          285
Ile Asn Gly Ser Ser Phe Gly Ala Val Gly Phe Pro Tyr Val Gln Val
      290          295          300
Leu Lys Thr Ala Asp Ile Asn Ser Ser Glu Val Glu Val Leu Tyr Leu
      305          310          315          320
Arg Asn Val Ser Ala Glu Asp Ala Gly Glu Tyr Thr Cys Leu Ala Gly
      325          330          335
Asn Ser Ile Gly Leu Ser Tyr Gln Ser Ala Trp Leu Thr Val Leu Pro
      340          345          350
Glu Glu Asp Pro Thr Trp Thr Ala Ala Ala Pro Glu Ala Arg Tyr Thr
      355          360          365
Asp Ile Ile Leu Tyr Ala Ser Gly Ser Leu Ala Leu Ala Val Leu Leu
      370          375          380
Leu Leu Ala Gly Leu Tyr Arg Gly Gln Ala Leu His Gly Arg His Pro
      385          390          395          400
Arg Pro Pro Ala Thr Val Gln Lys Leu Ser Arg Phe Pro Leu Ala Arg
      405          410          415
Gln Phe Ser Leu Glu Ser Gly Ser Ser Gly Lys Ser Ser Ser Ser Leu
      420          425          430
Val Arg Gly Val Arg Leu Ser Ser Ser Gly Pro Ala Leu Leu Ala Gly
      435          440          445
Leu Val Ser Leu Asp Leu Pro Leu Asp Pro Leu Trp Glu Phe Pro Arg
      450          455          460

```

```

Asp Arg Leu Val Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln
465                               470                               475                               480
Val Val Arg Ala Glu Ala Phe Gly Met Asp Pro Ala Arg Pro Asp Gln
                               485                               490                               495
Ala Ser Thr Val Ala Val Lys Met Leu Lys Asp Asn Ala Ser Asp Lys
                               500                               505                               510
Asp Leu Ala Asp Leu Val Ser Glu Met Glu Val Met Lys Leu Ile Gly
                               515                               520                               525
Arg His Lys Asn Ile Ile Asn Leu Leu Gly Val Cys Thr Gln Glu Gly
                               530                               535                               540
Pro Leu Tyr Val Ile Val Glu Cys Ala Ala Lys Gly Asn Leu Arg Glu
545                               550                               555                               560
Phe Leu Arg Ala Arg Arg Pro Pro Gly Pro Asp Leu Ser Pro Asp Gly
                               565                               570                               575
Pro Arg Ser Ser Glu Gly Pro Leu Ser Phe Pro Val Leu Val Ser Cys
                               580                               585                               590
Ala Tyr Gln Val Ala Arg Gly Met Gln Tyr Leu Glu Ser Arg Lys Cys
                               595                               600                               605
Ile His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Thr Glu Asp Asn
610                               615                               620
Val Met Lys Ile Ala Asp Phe Gly Leu Ala Arg Gly Val His His Ile
625                               630                               635                               640
Asp Tyr Tyr Lys Lys Thr Ser Asn Gly Arg Leu Pro Val Lys Trp Met
                               645                               650                               655
Ala Pro Glu Ala Leu Phe Asp Arg Val Tyr Thr His Gln Ser Asp Val
                               660                               665                               670
Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Ser
675                               680                               685
Pro Tyr Pro Gly Ile Pro Val Glu Glu Leu Phe Ser Leu Leu Arg Glu
690                               695                               700
Gly His Arg Met Asp Arg Pro Pro His Cys Pro Pro Glu Leu Tyr Gly
705                               710                               715                               720
Leu Met Arg Glu Cys Trp His Ala Ala Pro Ser Gln Arg Pro Thr Phe
                               725                               730                               735
Lys Gln Leu Val Glu Ala Leu Asp Lys Val Leu Leu Ala Val Ser Glu
740                               745                               750
Glu Tyr Leu Asp Leu Arg Leu Thr Phe Gly Pro Tyr Ser Pro Ser Gly
755                               760                               765
Gly Asp Ala Ser Ser Thr Cys Ser Ser Ser Asp Ser Val Phe Ser His
770                               775                               780
Asp Pro Leu Pro Leu Gly Ser Ser Ser Phe Pro Phe Gly Ser Gly Val
785                               790                               795                               800
Gln Thr

```

<210> 10

<211> 529

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 10

```

Met Gly Cys Val Phe Cys Lys Lys Leu Glu Pro Val Ala Thr Ala Lys
 1           5           10           15
Glu Asp Ala Gly Leu Glu Gly Asp Phe Arg Ser Tyr Gly Ala Ala Asp
          20           25           30
His Tyr Gly Pro Asp Pro Thr Lys Ala Arg Pro Ala Ser Phe Ala
        35           40           45

```

His Ile Pro Asn Tyr Ser Asn Phe Ser Ser Gln Ala Ile Asn Pro Gly
 50 55 60
 Phe Leu Asp Ser Gly Thr Ile Arg Gly Val Ser Gly Ile Gly Val Thr
 65 70 75 80
 Leu Phe Ile Ala Leu Tyr Asp Tyr Glu Ala Arg Thr Glu Asp Asp Leu
 85 90 95
 Thr Phe Thr Lys Gly Glu Lys Phe His Ile Leu Asn Asn Thr Glu Gly
 100 105 110
 Asp Trp Trp Glu Ala Arg Ser Leu Ser Ser Gly Lys Thr Gly Cys Ile
 115 120 125
 Pro Ser Asn Tyr Val Ala Pro Val Asp Ser Ile Gln Ala Glu Glu Trp
 130 135 140
 Tyr Phe Gly Lys Ile Gly Arg Lys Asp Ala Glu Arg Gln Leu Leu Ser
 145 150 155 160
 Pro Gly Asn Pro Gln Gly Ala Phe Leu Ile Arg Glu Ser Glu Thr Thr
 165 170 175
 Lys Gly Ala Tyr Ser Leu Ser Ile Arg Asp Trp Asp Gln Thr Arg Gly
 180 185 190
 Asp His Val Lys His Tyr Lys Ile Arg Lys Leu Asp Met Gly Gly Tyr
 195 200 205
 Tyr Ile Thr Thr Arg Val Gln Phe Asn Ser Val Gln Glu Leu Val Gln
 210 215 220
 His Tyr Met Glu Val Asn Asp Gly Leu Cys Asn Leu Leu Ile Ala Pro
 225 230 235 240
 Cys Thr Ile Met Lys Pro Gln Thr Leu Gly Leu Ala Lys Asp Ala Trp
 245 250 255
 Glu Ile Ser Arg Ser Ser Ile Thr Leu Glu Arg Arg Leu Gly Thr Gly
 260 265 270
 Cys Phe Gly Asp Val Trp Leu Gly Thr Trp Asn Gly Ser Thr Lys Val
 275 280 285
 Ala Val Lys Thr Leu Lys Pro Gly Thr Met Ser Pro Lys Ala Phe Leu
 290 295 300
 Glu Glu Ala Gln Val Met Lys Leu Leu Arg His Asp Lys Leu Val Gln
 305 310 315 320
 Leu Tyr Ala Val Val Ser Glu Glu Pro Ile Tyr Ile Val Thr Glu Phe
 325 330 335
 Met Cys His Gly Ser Leu Leu Asp Phe Leu Lys Asn Pro Glu Gly Gln
 340 345 350
 Asp Leu Arg Leu Pro Gln Leu Val Asp Met Ala Ala Gln Val Ala Glu
 355 360 365
 Gly Met Ala Tyr Met Glu Arg Met Asn Tyr Ile His Arg Asp Leu Arg
 370 375 380
 Ala Ala Asn Ile Leu Val Gly Glu Arg Leu Ala Cys Lys Ile Ala Asp
 385 390 395 400
 Phe Gly Leu Ala Arg Leu Ile Lys Asp Asp Glu Tyr Asn Pro Cys Gln
 405 410 415
 Gly Ser Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ala Leu Phe
 420 425 430
 Gly Arg Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu Leu
 435 440 445
 Thr Glu Leu Ile Thr Lys Gly Arg Ile Pro Tyr Pro Gly Met Asn Lys
 450 455 460
 Arg Glu Val Leu Glu Gln Val Glu Gln Gly Tyr His Met Pro Cys Pro
 465 470 475 480
 Pro Gly Cys Pro Ala Ser Leu Tyr Glu Ala Met Glu Gln Thr Trp Arg
 485 490 495
 Leu Asp Pro Glu Glu Arg Pro Thr Phe Glu Tyr Leu Gln Ser Phe Leu
 500 505 510
 Glu Asp Tyr Phe Thr Ser Ala Glu Pro Gln Tyr Gln Pro Gly Asp Gln
 515 520 525
 Thr

<210> 11
 <211> 1338
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 11
 Met Val Ser Tyr Trp Asp Thr Gly Val Leu Leu Cys Ala Leu Leu Ser
 1 5 10 15
 Cys Leu Leu Leu Thr Gly Ser Ser Ser Gly Ser Lys Leu Lys Asp Pro
 20 25 30
 Glu Leu Ser Leu Lys Gly Thr Gln His Ile Met Gln Ala Gly Gln Thr
 35 40 45
 Leu His Leu Gln Cys Arg Gly Glu Ala Ala His Lys Trp Ser Leu Pro
 50 55 60
 Glu Met Val Ser Lys Glu Ser Glu Arg Leu Ser Ile Thr Lys Ser Ala
 65 70 75 80
 Cys Gly Arg Asn Gly Lys Gln Phe Cys Ser Thr Leu Thr Leu Asn Thr
 85 90 95
 Ala Gln Ala Asn His Thr Gly Phe Tyr Ser Cys Lys Tyr Leu Ala Val
 100 105 110
 Pro Thr Ser Lys Lys Lys Glu Thr Glu Ser Ala Ile Tyr Ile Phe Ile
 115 120 125
 Ser Asp Thr Gly Arg Pro Phe Val Glu Met Tyr Ser Glu Ile Pro Glu
 130 135 140
 Ile Ile His Met Thr Glu Gly Arg Glu Leu Val Ile Pro Cys Arg Val
 145 150 155 160
 Thr Ser Pro Asn Ile Thr Val Thr Leu Lys Lys Phe Pro Leu Asp Thr
 165 170 175
 Leu Ile Pro Asp Gly Lys Arg Ile Ile Trp Asp Ser Arg Lys Gly Phe
 180 185 190
 Ile Ile Ser Asn Ala Thr Tyr Lys Glu Ile Gly Leu Leu Thr Cys Glu
 195 200 205
 Ala Thr Val Asn Gly His Leu Tyr Lys Thr Asn Tyr Leu Thr His Arg
 210 215 220
 Gln Thr Asn Thr Ile Ile Asp Val Gln Ile Ser Thr Pro Arg Pro Val
 225 230 235 240
 Lys Leu Leu Arg Gly His Thr Leu Val Leu Asn Cys Thr Ala Thr Thr
 245 250 255
 Pro Leu Asn Thr Arg Val Gln Met Thr Trp Ser Tyr Pro Asp Glu Lys
 260 265 270
 Asn Lys Arg Ala Ser Val Arg Arg Arg Ile Asp Gln Ser Asn Ser His
 275 280 285
 Ala Asn Ile Phe Tyr Ser Val Leu Thr Ile Asp Lys Met Gln Asn Lys
 290 295 300
 Asp Lys Gly Leu Tyr Thr Cys Arg Val Arg Ser Gly Pro Ser Phe Lys
 305 310 315 320
 Ser Val Asn Thr Ser Val His Ile Tyr Asp Lys Ala Phe Ile Thr Val
 325 330 335
 Lys His Arg Lys Gln Gln Val Leu Glu Thr Val Ala Gly Lys Arg Ser
 340 345 350
 Tyr Arg Leu Ser Met Lys Val Lys Ala Phe Pro Ser Pro Glu Val Val
 355 360 365
 Trp Leu Lys Asp Gly Leu Pro Ala Thr Glu Lys Ser Ala Arg Tyr Leu
 370 375 380

Thr Arg Gly Tyr Ser Leu Ile Ile Lys Asp Val Thr Glu Glu Asp Ala
 385 390 395 400
 Gly Asn Tyr Thr Ile Leu Leu Ser Ile Lys Gln Ser Asn Val Phe Lys
 405 410 415
 Asn Leu Thr Ala Thr Leu Ile Val Asn Val Lys Pro Gln Ile Tyr Glu
 420 425 430
 Lys Ala Val Ser Ser Phe Pro Asp Pro Ala Leu Tyr Pro Leu Gly Ser
 435 440 445
 Arg Gln Ile Leu Thr Cys Thr Ala Tyr Gly Ile Pro Gln Pro Thr Ile
 450 455 460
 Lys Trp Phe Trp His Pro Cys Asn His Asn His Ser Glu Ala Arg Cys
 465 470 475 480
 Asp Phe Cys Ser Asn Asn Glu Glu Ser Phe Ile Leu Asp Ala Asp Ser
 485 490 495
 Asn Met Gly Asn Arg Ile Glu Ser Ile Thr Gln Arg Met Ala Ile Ile
 500 505 510
 Glu Gly Lys Asn Lys Met Ala Ser Thr Leu Val Val Ala Asp Ser Arg
 515 520 525
 Ile Ser Gly Ile Tyr Ile Cys Ile Ala Ser Asn Lys Val Gly Thr Val
 530 535 540
 Gly Arg Asn Ile Ser Phe Tyr Ile Thr Asp Val Pro Asn Gly Phe His
 545 550 555 560
 Val Asn Leu Glu Lys Met Pro Thr Glu Gly Glu Asp Leu Lys Leu Ser
 565 570 575
 Cys Thr Val Asn Lys Phe Leu Tyr Arg Asp Val Thr Trp Ile Leu Leu
 580 585 590
 Arg Thr Val Asn Asn Arg Thr Met His Tyr Ser Ile Ser Lys Gln Lys
 595 600 605
 Met Ala Ile Thr Lys Glu His Ser Ile Thr Leu Asn Leu Thr Ile Met
 610 615 620
 Asn Val Ser Leu Gln Asp Ser Gly Thr Tyr Ala Cys Arg Ala Arg Asn
 625 630 635 640
 Val Tyr Thr Gly Glu Glu Ile Leu Gln Lys Lys Glu Ile Thr Ile Arg
 645 650 655
 Asp Gln Glu Ala Pro Tyr Leu Leu Arg Asn Leu Ser Asp His Thr Val
 660 665 670
 Ala Ile Ser Ser Ser Thr Thr Leu Asp Cys His Ala Asn Gly Val Pro
 675 680 685
 Glu Pro Gln Ile Thr Trp Phe Lys Asn Asn His Lys Ile Gln Gln Glu
 690 695 700
 Pro Gly Ile Ile Leu Gly Pro Gly Ser Ser Thr Leu Phe Ile Glu Arg
 705 710 715 720
 Val Thr Glu Glu Asp Glu Gly Val Tyr His Cys Lys Ala Thr Asn Gln
 725 730 735
 Lys Gly Ser Val Glu Ser Ser Ala Tyr Leu Thr Val Gln Gly Thr Ser
 740 745 750
 Asp Lys Ser Asn Leu Glu Leu Ile Thr Leu Thr Cys Thr Cys Val Ala
 755 760 765
 Ala Thr Leu Phe Trp Leu Leu Leu Thr Leu Leu Ile Arg Lys Met Lys
 770 775 780
 Arg Ser Ser Ser Glu Ile Lys Thr Asp Tyr Leu Ser Ile Ile Met Asp
 785 790 795 800
 Pro Asp Glu Val Pro Leu Asp Glu Gln Cys Glu Arg Leu Pro Tyr Asp
 805 810 815
 Ala Ser Lys Trp Glu Phe Ala Arg Glu Arg Leu Lys Leu Gly Lys Ser
 820 825 830
 Leu Gly Arg Gly Ala Phe Gly Lys Val Val Gln Ala Ser Ala Phe Gly
 835 840 845
 Ile Lys Lys Ser Pro Thr Cys Arg Thr Val Ala Val Lys Met Leu Lys
 850 855 860

Glu Gly Ala Thr Ala Ser Glu Tyr Lys Ala Leu Met Thr Glu Leu Lys
 865 870 875 880
 Ile Leu Thr His Ile Gly His His Leu Asn Val Val Asn Leu Leu Gly
 885 890 895
 Ala Cys Thr Lys Gln Gly Gly Pro Leu Met Val Ile Val Glu Tyr Cys
 900 905 910
 Lys Tyr Gly Asn Leu Ser Asn Tyr Leu Lys Ser Lys Arg Asp Leu Phe
 915 920 925
 Phe Leu Asn Lys Asp Ala Ala Leu His Met Glu Pro Lys Lys Glu Lys
 930 935 940
 Met Glu Pro Gly Leu Glu Gln Gly Lys Lys Pro Arg Leu Asp Ser Val
 945 950 955 960
 Thr Ser Ser Glu Ser Phe Ala Ser Ser Gly Phe Gln Glu Asp Lys Ser
 965 970 975
 Leu Ser Asp Val Glu Glu Glu Glu Asp Ser Asp Gly Phe Tyr Lys Glu
 980 985 990
 Pro Ile Thr Met Glu Asp Leu Ile Ser Tyr Ser Phe Gln Val Ala Arg
 995 1000 1005
 Gly Met Glu Phe Leu Ser Ser Arg Lys Cys Ile His Arg Asp Leu Ala
 1010 1015 1020
 Ala Arg Asn Ile Leu Leu Ser Glu Asn Asn Val Val Lys Ile Cys Asp
 1025 1030 1035 1040
 Phe Gly Leu Ala Arg Asp Ile Tyr Lys Asn Pro Asp Tyr Val Arg Lys
 1045 1050 1055
 Gly Asp Thr Arg Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe
 1060 1065 1070
 Asp Lys Ile Tyr Ser Thr Lys Ser Asp Val Trp Ser Tyr Gly Val Leu
 1075 1080 1085
 Leu Trp Glu Ile Phe Ser Leu Gly Gly Ser Pro Tyr Pro Gly Val Gln
 1090 1095 1100
 Met Asp Glu Asp Phe Cys Ser Arg Leu Arg Glu Gly Met Arg Met Arg
 1105 1110 1115 1120
 Ala Pro Glu Tyr Ser Thr Pro Glu Ile Tyr Gln Ile Met Leu Asp Cys
 1125 1130 1135
 Trp His Arg Asp Pro Lys Glu Arg Pro Arg Phe Ala Glu Leu Val Glu
 1140 1145 1150
 Lys Leu Gly Asp Leu Leu Gln Ala Asn Val Gln Gln Asp Gly Lys Asp
 1155 1160 1165
 Tyr Ile Pro Ile Asn Ala Ile Leu Thr Gly Asn Ser Gly Phe Thr Tyr
 1170 1175 1180
 Ser Thr Pro Ala Phe Ser Glu Asp Phe Phe Lys Glu Ser Ile Ser Ala
 1185 1190 1195 1200
 Pro Lys Phe Asn Ser Gly Ser Ser Asp Asp Val Arg Tyr Val Asn Ala
 1205 1210 1215
 Phe Lys Phe Met Ser Leu Glu Arg Ile Lys Thr Phe Glu Glu Leu Leu
 1220 1225 1230
 Pro Asn Ala Thr Ser Met Phe Asp Asp Tyr Gln Gly Asp Ser Ser Thr
 1235 1240 1245
 Leu Leu Ala Ser Pro Met Leu Lys Arg Phe Thr Trp Thr Asp Ser Lys
 1250 1255 1260
 Pro Lys Ala Ser Leu Lys Ile Asp Leu Arg Val Thr Ser Lys Ser Lys
 1265 1270 1275 1280
 Glu Ser Gly Leu Ser Asp Val Ser Arg Pro Ser Phe Cys His Ser Ser
 1285 1290 1295
 Cys Gly His Val Ser Glu Gly Lys Arg Arg Phe Thr Tyr Asp His Ala
 1300 1305 1310
 Glu Leu Glu Arg Lys Ile Ala Cys Cys Ser Pro Pro Pro Asp Tyr Asn
 1315 1320 1325
 Ser Val Val Leu Tyr Ser Thr Pro Pro Ile
 1330 1335

<210> 12
 <211> 1298
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 12
 Met Gln Arg Gly Ala Ala Leu Cys Leu Arg Leu Trp Leu Cys Leu Gly
 1 5 10 15
 Leu Leu Asp Gly Leu Val Ser Gly Tyr Ser Met Thr Pro Pro Thr Leu
 20 25 30
 Asn Ile Thr Glu Glu Ser His Val Ile Asp Thr Gly Asp Ser Leu Ser
 35 40 45
 Ile Ser Cys Arg Gly Gln His Pro Leu Glu Trp Ala Trp Pro Gly Ala
 50 55 60
 Gln Glu Ala Pro Ala Thr Gly Asp Lys Asp Ser Glu Asp Thr Gly Val
 65 70 75 80
 Val Arg Asp Cys Glu Gly Thr Asp Ala Arg Pro Tyr Cys Lys Val Leu
 85 90 95
 Leu Leu His Glu Val His Ala Asn Asp Thr Gly Ser Tyr Val Cys Tyr
 100 105 110
 Tyr Lys Tyr Ile Lys Ala Arg Ile Glu Gly Thr Thr Ala Ala Ser Ser
 115 120 125
 Tyr Val Phe Val Arg Asp Phe Glu Gln Pro Phe Ile Asn Lys Pro Asp
 130 135 140
 Thr Leu Leu Val Asn Arg Lys Asp Ala Met Trp Val Pro Cys Leu Val
 145 150 155 160
 Ser Ile Pro Gly Leu Asn Val Thr Leu Arg Ser Gln Ser Ser Val Leu
 165 170 175
 Trp Pro Asp Gly Gln Glu Val Val Trp Asp Asp Arg Arg Gly Met Leu
 180 185 190
 Val Ser Thr Pro Leu Leu His Asp Ala Leu Tyr Leu Gln Cys Glu Thr
 195 200 205
 Thr Trp Gly Asp Gln Asp Phe Leu Ser Asn Pro Phe Leu Val His Ile
 210 215 220
 Thr Gly Asn Glu Leu Tyr Asp Ile Gln Leu Leu Pro Arg Lys Ser Leu
 225 230 235 240
 Glu Leu Leu Val Gly Glu Lys Leu Val Leu Asn Cys Thr Val Trp Ala
 245 250 255
 Glu Phe Asn Ser Gly Val Thr Phe Asp Trp Asp Tyr Pro Gly Lys Gln
 260 265 270
 Ala Glu Arg Gly Lys Trp Val Pro Glu Arg Arg Ser Gln Gln Thr His
 275 280 285
 Thr Glu Leu Ser Ser Ile Leu Thr Ile His Asn Val Ser Gln His Asp
 290 295 300
 Leu Gly Ser Tyr Val Cys Lys Ala Asn Asn Gly Ile Gln Arg Phe Arg
 305 310 315 320
 Glu Ser Thr Glu Val Ile Val His Glu Asn Pro Phe Ile Ser Val Glu
 325 330 335
 Trp Leu Lys Gly Pro Ile Leu Glu Ala Thr Ala Gly Asp Glu Leu Val
 340 345 350
 Lys Leu Pro Val Lys Leu Ala Ala Tyr Pro Pro Pro Glu Phe Gln Trp
 355 360 365
 Tyr Lys Asp Gly Lys Ala Leu Ser Gly Arg His Ser Pro His Ala Leu
 370 375 380
 Val Leu Lys Glu Val Thr Glu Ala Ser Thr Gly Thr Tyr Thr Leu Ala
 385 390 395 400

```

Leu Trp Asn Ser Ala Ala Gly Leu Arg Arg Asn Ile Ser Leu Glu Leu
405 410 415
Val Val Asn Val Pro Pro Gln Ile His Glu Lys Glu Ala Ser Ser Pro
420 425 430
Ser Ile Tyr Ser Arg His Ser Arg Gln Ala Leu Thr Cys Thr Ala Tyr
435 440 445
Gly Val Pro Leu Pro Leu Ser Ile Gln Trp His Trp Arg Pro Trp Thr
450 455 460
Pro Cys Lys Met Phe Ala Gln Arg Ser Leu Arg Arg Arg Gln Gln Gln
465 470 475 480
Asp Leu Met Pro Gln Cys Arg Asp Trp Arg Ala Val Thr Thr Gln Asp
485 490 495
Ala Val Asn Pro Ile Glu Ser Leu Asp Thr Trp Thr Glu Phe Val Glu
500 505 510
Gly Lys Asn Lys Thr Val Ser Lys Leu Val Ile Gln Asn Ala Asn Val
515 520 525
Ser Ala Met Tyr Lys Cys Val Val Ser Asn Lys Val Gly Gln Asp Glu
530 535 540
Arg Leu Ile Tyr Phe Tyr Val Thr Thr Ile Pro Asp Gly Phe Thr Ile
545 550 555 560
Glu Ser Lys Pro Ser Glu Glu Leu Leu Glu Gly Gln Pro Val Leu Leu
565 570 575
Ser Cys Gln Ala Asp Ser Tyr Lys Tyr Glu His Leu Arg Trp Tyr Arg
580 585 590
Leu Asn Leu Ser Thr Leu His Asp Ala His Gly Asn Pro Leu Leu Leu
595 600 605
Asp Cys Lys Asn Val His Leu Phe Ala Thr Pro Leu Ala Ala Ser Leu
610 615 620
Glu Glu Val Ala Pro Gly Ala Arg His Ala Thr Leu Ser Leu Ser Ile
625 630 635 640
Pro Arg Val Ala Pro Glu His Glu Gly His Tyr Val Cys Glu Val Gln
645 650 655
Asp Arg Arg Ser His Asp Lys His Cys His Lys Lys Tyr Leu Ser Val
660 665 670
Gln Ala Leu Glu Ala Pro Arg Leu Thr Gln Asn Leu Thr Asp Leu Leu
675 680 685
Val Asn Val Ser Asp Ser Leu Glu Met Gln Cys Leu Val Ala Gly Ala
690 695 700
His Ala Pro Ser Ile Val Trp Tyr Lys Asp Glu Arg Leu Leu Glu Glu
705 710 715 720
Lys Ser Gly Val Asp Leu Ala Asp Ser Asn Gln Lys Leu Ser Ile Gln
725 730 735
Arg Val Arg Glu Glu Asp Ala Gly Arg Tyr Leu Cys Ser Val Cys Asn
740 745 750
Ala Lys Gly Cys Val Asn Ser Ser Ala Ser Val Ala Val Glu Gly Ser
755 760 765
Glu Asp Lys Gly Ser Met Glu Ile Val Ile Leu Val Gly Thr Gly Val
770 775 780
Ile Ala Val Phe Phe Trp Val Leu Leu Leu Leu Ile Phe Cys Asn Met
785 790 795 800
Arg Arg Pro Ala His Ala Asp Ile Lys Thr Gly Tyr Leu Ser Ile Ile
805 810 815
Met Asp Pro Gly Glu Val Pro Leu Glu Glu Gln Cys Glu Tyr Leu Ser
820 825 830
Tyr Asp Ala Ser Gln Trp Glu Phe Pro Arg Glu Arg Leu His Leu Gly
835 840 845
Arg Val Leu Gly Tyr Gly Ala Phe Gly Lys Val Val Glu Ala Ser Ala
850 855 860
Phe Gly Ile His Lys Gly Ser Ser Cys Asp Thr Val Ala Val Lys Met
865 870 875 880

```

Leu Lys Glu Gly Ala Thr Ala Ser Glu His Arg Ala Leu Met Ser Glu
 885 890 895
 Leu Lys Ile Leu Ile His Ile Gly Asn His Leu Asn Val Val Asn Leu
 900 905 910
 Leu Gly Ala Cys Thr Lys Pro Gln Gly Pro Leu Met Val Ile Val Glu
 915 920 925
 Phe Cys Lys Tyr Gly Asn Leu Ser Asn Phe Leu Arg Ala Lys Arg Asp
 930 935 940
 Ala Phe Ser Pro Cys Ala Glu Lys Ser Pro Glu Gln Arg Gly Arg Phe
 945 950 955 960
 Arg Ala Met Val Glu Leu Ala Arg Leu Asp Arg Arg Arg Pro Gly Ser
 965 970 975
 Ser Asp Arg Val Leu Phe Ala Arg Phe Ser Lys Thr Glu Gly Gly Ala
 980 985 990
 Arg Arg Ala Ser Pro Asp Gln Glu Ala Glu Asp Leu Trp Leu Ser Pro
 995 1000 1005
 Leu Thr Met Glu Asp Leu Val Cys Tyr Ser Phe Gln Val Ala Arg Gly
 1010 1015 1020
 Met Glu Phe Leu Ala Ser Arg Lys Cys Ile His Arg Asp Leu Ala Ala
 1025 1030 1035 1040
 Arg Asn Ile Leu Leu Ser Glu Ser Asp Val Val Lys Ile Cys Asp Phe
 1045 1050 1055
 Gly Leu Ala Arg Asp Ile Tyr Lys Asp Pro Asp Tyr Val Arg Lys Gly
 1060 1065 1070
 Ser Ala Arg Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe Asp
 1075 1080 1085
 Lys Val Tyr Thr Thr Gln Ser Asp Val Trp Ser Phe Gly Val Leu Leu
 1090 1095 1100
 Trp Glu Ile Phe Ser Leu Gly Ala Ser Pro Tyr Pro Gly Val Gln Ile
 1105 1110 1115 1120
 Asn Glu Glu Phe Cys Gln Arg Leu Arg Asp Gly Thr Arg Met Arg Ala
 1125 1130 1135
 Pro Glu Leu Ala Thr Pro Ala Ile Arg Arg Ile Met Leu Asn Cys Trp
 1140 1145 1150
 Ser Gly Asp Pro Lys Ala Arg Pro Ala Phe Ser Glu Leu Val Glu Ile
 1155 1160 1165
 Leu Gly Asp Leu Leu Gln Gly Arg Gly Leu Gln Glu Glu Glu Val
 1170 1175 1180
 Cys Met Ala Pro Arg Ser Ser Gln Ser Ser Glu Glu Gly Ser Phe Ser
 1185 1190 1195 1200
 Gln Val Ser Thr Met Ala Leu His Ile Ala Gln Ala Asp Ala Glu Asp
 1205 1210 1215
 Ser Pro Pro Ser Leu Gln Arg His Ser Leu Ala Ala Arg Tyr Tyr Asn
 1220 1225 1230
 Trp Val Ser Phe Pro Gly Cys Leu Ala Arg Gly Ala Glu Thr Arg Gly
 1235 1240 1245
 Ser Ser Arg Met Lys Thr Phe Glu Glu Phe Pro Met Thr Pro Thr Thr
 1250 1255 1260
 Tyr Lys Gly Ser Val Asp Asn Gln Thr Asp Ser Gly Met Val Leu Ala
 1265 1270 1275 1280
 Ser Glu Glu Phe Glu Gln Ile Glu Ser Arg His Arg Gln Glu Ser Gly
 1285 1290 1295
 Phe Arg

<210> 13

<211> 537

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 13

```

Met Gly Cys Val Gln Cys Lys Asp Lys Glu Ala Thr Lys Leu Thr Glu
 1           5           10           15
Glu Arg Asp Gly Ser Leu Asn Gln Ser Ser Gly Tyr Arg Tyr Gly Thr
          20           25           30
Asp Pro Thr Pro Gln His Tyr Pro Ser Phe Gly Val Thr Ser Ile Pro
          35           40           45
Asn Tyr Asn Asn Phe His Ala Gly Gly Gln Gly Leu Thr Val Phe
          50           55           60
Gly Gly Val Asn Ser Ser His Thr Gly Thr Leu Arg Thr Arg Gly
65          70          75          80
Gly Thr Gly Val Thr Leu Phe Val Ala Leu Tyr Asp Tyr Glu Ala Arg
          85          90          95
Thr Glu Asp Asp Leu Ser Phe His Lys Gly Glu Lys Phe Gln Ile Leu
          100         105         110
Asn Ser Ser Glu Gly Asp Trp Trp Glu Ala Arg Ser Leu Thr Thr Gly
          115         120         125
Glu Thr Gly Tyr Ile Pro Ser Asn Tyr Val Ala Pro Val Asp Ser Ile
          130         135         140
Gln Ala Glu Glu Trp Tyr Phe Gly Lys Leu Gly Arg Lys Asp Ala Glu
145         150         155         160
Arg Gln Leu Leu Ser Phe Gly Asn Pro Arg Gly Thr Phe Leu Ile Arg
          165         170         175
Glu Ser Glu Thr Thr Lys Gly Ala Tyr Ser Leu Ser Ile Arg Asp Trp
          180         185         190
Asp Asp Met Lys Gly Asp His Val Lys His Tyr Lys Ile Arg Lys Leu
          195         200         205
Asp Asn Gly Gly Tyr Tyr Ile Thr Thr Arg Ala Gln Phe Glu Thr Leu
          210         215         220
Gln Gln Leu Val Gln His Tyr Ser Glu Arg Ala Ala Gly Leu Cys Cys
225         230         235         240
Arg Leu Val Val Pro Cys His Lys Gly Met Pro Arg Leu Thr Asp Leu
          245         250         255
Ser Val Lys Thr Lys Asp Val Trp Glu Ile Pro Arg Glu Ser Leu Gln
          260         265         270
Leu Ile Lys Arg Leu Gly Asn Gly Gln Phe Gly Glu Val Trp Met Gly
          275         280         285
Thr Trp Asn Gly Asn Thr Lys Val Ala Ile Lys Thr Leu Lys Pro Gly
          290         295         300
Thr Met Ser Pro Glu Ser Phe Leu Glu Glu Ala Gln Ile Met Lys Lys
305         310         315         320
Leu Lys His Asp Lys Leu Val Gln Leu Tyr Ala Val Val Ser Glu Glu
          325         330         335
Pro Ile Tyr Ile Val Thr Glu Tyr Met Asn Lys Gly Ser Leu Leu Asp
          340         345         350
Phe Leu Lys Asp Gly Glu Gly Arg Ala Leu Lys Leu Pro Asn Leu Val
          355         360         365
Asp Met Ala Ala Gln Val Ala Ala Gly Met Ala Tyr Ile Glu Arg Met
          370         375         380
Asn Tyr Ile His Arg Asp Leu Arg Ser Ala Asn Ile Leu Val Gly Asn
385         390         395         400
Gly Leu Ile Cys Lys Ile Ala Asp Phe Gly Leu Ala Arg Leu Ile Glu
          405         410         415
Asp Asn Glu Tyr Thr Ala Arg Gln Gly Ala Lys Phe Pro Ile Lys Trp
          420         425         430
Thr Ala Pro Glu Ala Ala Leu Tyr Gly Arg Phe Thr Ile Lys Ser Asp
          435         440         445

```

```

Val Trp Ser Phe Gly Ile Leu Leu Thr Glu Leu Val Thr Lys Gly Arg
  450                      455                      460
Val Pro Tyr Pro Gly Met Asn Asn Arg Glu Val Leu Glu Gln Val Glu
465                      470                      475                      480
Arg Gly Tyr Arg Met Pro Cys Pro Gln Asp Cys Pro Ile Ser Leu His
                      485                      490                      495
Glu Leu Met Ile His Cys Trp Lys Lys Asp Pro Glu Glu Arg Pro Thr
                      500                      505                      510
Phe Glu Tyr Leu Gln Ser Phe Leu Glu Asp Tyr Phe Thr Ala Thr Glu
                      515                      520                      525
Pro Gln Tyr Gln Pro Gly Glu Asn Leu
  530                      535

```

<210> 14

<211> 526

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 14

```

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

```

```

Glu Ala Asn Val Met Lys Thr Leu Gln His Asp Lys Leu Val Lys Leu
305                      310                      315                      320
His Ala Val Val Thr Lys Glu Pro Ile Tyr Ile Ile Thr Glu Phe Met
                      325                      330                      335
Ala Lys Gly Ser Leu Leu Asp Phe Leu Lys Ser Asp Glu Gly Ser Lys
                      340                      345                      350
Gln Pro Leu Pro Lys Leu Ile Asp Phe Ser Ala Gln Ile Ala Glu Gly
                      355                      360                      365
Met Ala Phe Ile Glu Gln Arg Asn Tyr Ile His Arg Asp Leu Arg Ala
                      370                      375                      380
Ala Asn Ile Leu Val Ser Ala Ser Leu Val Cys Lys Ile Ala Asp Phe
385                      390                      395                      400
Gly Leu Ala Arg Val Ile Glu Asp Asn Glu Tyr Thr Ala Arg Glu Gly
                      405                      410                      415
Ala Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ile Asn Phe Gly
                      420                      425                      430
Ser Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Met
                      435                      440                      445
Glu Ile Val Thr Tyr Gly Arg Ile Pro Tyr Pro Gly Met Ser Asn Pro
                      450                      455                      460
Glu Val Ile Arg Ala Leu Glu Arg Gly Tyr Arg Met Pro Arg Pro Glu
465                      470                      475                      480
Asn Cys Pro Glu Glu Leu Tyr Asn Ile Met Met Arg Cys Trp Lys Asn
                      485                      490                      495
Arg Pro Glu Glu Arg Pro Thr Phe Glu Tyr Ile Gln Ser Val Leu Asp
                      500                      505                      510
Asp Phe Tyr Thr Ala Thr Glu Ser Gln Tyr Gln Gln Gln Pro
                      515                      520                      525

```

<210> 15

<211> 1308

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 15

```

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

```

Leu Val Ser Thr Asn Gly Ser Ser Gly Cys Gly Arg Cys His Lys Ser
 180 185 190
 Cys Thr Gly Arg Cys Trp Gly Pro Thr Glu Asn His Cys Gln Thr Leu
 195 200 205
 Thr Arg Thr Val Cys Ala Glu Gln Cys Asp Gly Arg Cys Tyr Gly Pro
 210 215 220
 Tyr Val Ser Asp Cys Cys His Arg Glu Cys Ala Gly Gly Cys Ser Gly
 225 230 235 240
 Pro Lys Asp Thr Asp Cys Phe Ala Cys Met Asn Phe Asn Asp Ser Gly
 245 250 255
 Ala Cys Val Thr Gln Cys Pro Gln Thr Phe Val Tyr Asn Pro Thr Thr
 260 265 270
 Phe Gln Leu Glu His Asn Phe Asn Ala Lys Tyr Thr Tyr Gly Ala Phe
 275 280 285
 Cys Val Lys Lys Cys Pro His Asn Phe Val Val Asp Ser Ser Ser Cys
 290 295 300
 Val Arg Ala Cys Pro Ser Ser Lys Met Glu Val Glu Glu Asn Gly Ile
 305 310 315 320
 Lys Met Cys Lys Pro Cys Thr Asp Ile Cys Pro Lys Ala Cys Asp Gly
 325 330 335
 Ile Gly Thr Gly Ser Leu Met Ser Ala Gln Thr Val Asp Ser Ser Asn
 340 345 350
 Ile Asp Lys Phe Ile Asn Cys Thr Lys Ile Asn Gly Asn Leu Ile Phe
 355 360 365
 Leu Val Thr Gly Ile His Gly Asp Pro Tyr Asn Ala Ile Glu Ala Ile
 370 375 380
 Asp Pro Glu Lys Leu Asn Val Phe Arg Thr Val Arg Glu Ile Thr Gly
 385 390 395 400
 Phe Leu Asn Ile Gln Ser Trp Pro Pro Asn Met Thr Asp Phe Ser Val
 405 410 415
 Phe Ser Asn Leu Val Thr Ile Gly Gly Arg Val Leu Tyr Ser Gly Leu
 420 425 430
 Ser Leu Leu Ile Leu Lys Gln Gln Gly Ile Thr Ser Leu Gln Phe Gln
 435 440 445
 Ser Leu Lys Glu Ile Ser Ala Gly Asn Ile Tyr Ile Thr Asp Asn Ser
 450 455 460
 Asn Leu Cys Tyr Tyr His Thr Ile Asn Trp Thr Thr Leu Phe Ser Thr
 465 470 475 480
 Ile Asn Gln Arg Ile Val Ile Arg Asp Asn Arg Lys Ala Glu Asn Cys
 485 490 495
 Thr Ala Glu Gly Met Val Cys Asn His Leu Cys Ser Ser Asp Gly Cys
 500 505 510
 Trp Gly Pro Gly Pro Asp Gln Cys Leu Ser Cys Arg Arg Phe Ser Arg
 515 520 525
 Gly Arg Ile Cys Ile Glu Ser Cys Asn Leu Tyr Asp Gly Glu Phe Arg
 530 535 540
 Glu Phe Glu Asn Gly Ser Ile Cys Val Glu Cys Asp Pro Gln Cys Glu
 545 550 555 560
 Lys Met Glu Asp Gly Leu Leu Thr Cys His Gly Pro Gly Pro Asp Asn
 565 570 575
 Cys Thr Lys Cys Ser His Phe Lys Asp Gly Pro Asn Cys Val Glu Lys
 580 585 590
 Cys Pro Asp Gly Leu Gln Gly Ala Asn Ser Phe Ile Phe Lys Tyr Ala
 595 600 605
 Asp Pro Asp Arg Glu Cys His Pro Cys His Pro Asn Cys Thr Gln Gly
 610 615 620
 Cys Asn Gly Pro Thr Ser His Asp Cys Ile Tyr Tyr Pro Trp Thr Gly
 625 630 635 640
 His Ser Thr Leu Pro Gln His Ala Arg Thr Pro Leu Ile Ala Ala Gly
 645 650 655

Val Ile Gly Gly Leu Phe Ile Leu Val Ile Val Gly Leu Thr Phe Ala
 660 665 670
 Val Tyr Val Arg Arg Lys Ser Ile Lys Lys Lys Arg Ala Leu Arg Arg
 675 680 685
 Phe Leu Glu Thr Glu Leu Val Glu Pro Leu Thr Pro Ser Gly Thr Ala
 690 695 700
 Pro Asn Gln Ala Gln Leu Arg Ile Leu Lys Glu Thr Glu Leu Lys Arg
 705 710 715 720
 Val Lys Val Leu Gly Ser Gly Ala Phe Gly Thr Val Tyr Lys Gly Ile
 725 730 735
 Trp Val Pro Glu Gly Glu Thr Val Lys Ile Pro Val Ala Ile Lys Ile
 740 745 750
 Leu Asn Glu Thr Thr Gly Pro Lys Ala Asn Val Glu Phe Met Asp Glu
 755 760 765
 Ala Leu Ile Met Ala Ser Met Asp His Pro His Leu Val Arg Leu Leu
 770 775 780
 Gly Val Cys Leu Ser Pro Thr Ile Gln Leu Val Thr Gln Leu Met Pro
 785 790 795 800
 His Gly Cys Leu Leu Glu Tyr Val His Glu His Lys Asp Asn Ile Gly
 805 810 815
 Ser Gln Leu Leu Leu Asn Trp Cys Val Gln Ile Ala Lys Gly Met Met
 820 825 830
 Tyr Leu Glu Glu Arg Arg Leu Val His Arg Asp Leu Ala Ala Arg Asn
 835 840 845
 Val Leu Val Lys Ser Pro Asn His Val Lys Ile Thr Asp Phe Gly Leu
 850 855 860
 Ala Arg Leu Leu Glu Gly Asp Glu Lys Glu Tyr Asn Ala Asp Gly Gly
 865 870 875 880
 Lys Met Pro Ile Lys Trp Met Ala Leu Glu Cys Ile His Tyr Arg Lys
 885 890 895
 Phe Thr His Gln Ser Asp Val Trp Ser Tyr Gly Val Thr Ile Trp Glu
 900 905 910
 Leu Met Thr Phe Gly Gly Lys Pro Tyr Asp Gly Ile Pro Thr Arg Glu
 915 920 925
 Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln Pro Pro Ile
 930 935 940
 Cys Thr Ile Asp Val Tyr Met Val Met Val Lys Cys Trp Met Ile Asp
 945 950 955 960
 Ala Asp Ser Arg Pro Lys Phe Lys Glu Leu Ala Ala Glu Phe Ser Arg
 965 970 975
 Met Ala Arg Asp Pro Gln Arg Tyr Leu Val Ile Gln Gly Asp Asp Arg
 980 985 990
 Met Lys Leu Pro Ser Pro Asn Asp Ser Lys Phe Phe Gln Asn Leu Leu
 995 1000 1005
 Asp Glu Glu Asp Leu Glu Asp Met Met Asp Ala Glu Glu Tyr Leu Val
 1010 1015 1020
 Pro Gln Ala Phe Asn Ile Pro Pro Pro Ile Tyr Thr Ser Arg Ala Arg
 1025 1030 1035 1040
 Ile Asp Ser Asn Arg Ser Glu Ile Gly His Ser Pro Pro Pro Ala Tyr
 1045 1050 1055
 Thr Pro Met Ser Gly Asn Gln Phe Val Tyr Arg Asp Gly Gly Phe Ala
 1060 1065 1070
 Ala Glu Gln Gly Val Ser Val Pro Tyr Arg Ala Pro Thr Ser Thr Ile
 1075 1080 1085
 Pro Glu Ala Pro Val Ala Gln Gly Ala Thr Ala Glu Ile Phe Asp Asp
 1090 1095 1100
 Ser Cys Cys Asn Gly Thr Leu Arg Lys Pro Val Ala Pro His Val Gln
 1105 1110 1115 1120
 Glu Asp Ser Ser Thr Gln Arg Tyr Ser Ala Asp Pro Thr Val Phe Ala
 1125 1130 1135

Pro Glu Arg Ser Pro Arg Gly Glu Leu Asp Glu Glu Gly Tyr Met Thr
 1140 1145 1150
 Pro Met Arg Asp Lys Pro Lys Gln Glu Tyr Leu Asn Pro Val Glu Glu
 1155 1160 1165
 Asn Pro Phe Val Ser Arg Arg Lys Asn Gly Asp Leu Gln Ala Leu Asp
 1170 1175 1180
 Asn Pro Glu Tyr His Asn Ala Ser Asn Gly Pro Pro Lys Ala Glu Asp
 1185 1190 1195 1200
 Glu Tyr Val Asn Glu Pro Leu Tyr Leu Asn Thr Phe Ala Asn Thr Leu
 1205 1210 1215
 Gly Lys Ala Glu Tyr Leu Lys Asn Asn Ile Leu Ser Met Pro Glu Lys
 1220 1225 1230
 Ala Lys Lys Ala Phe Asp Asn Pro Asp Tyr Trp Asn His Ser Leu Pro
 1235 1240 1245
 Pro Arg Ser Thr Leu Gln His Pro Asp Tyr Leu Gln Glu Tyr Ser Thr
 1250 1255 1260
 Lys Tyr Phe Tyr Lys Gln Asn Gly Arg Ile Arg Pro Ile Val Ala Glu
 1265 1270 1275 1280
 Asn Pro Glu Tyr Leu Ser Glu Phe Ser Leu Lys Pro Gly Thr Val Leu
 1285 1290 1295
 Pro Pro Pro Pro Tyr Arg His Arg Asn Thr Val Val
 1300 1305

<210> 16

<211> 745

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 16

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

Leu	Gln	Pro	Phe	Thr	Trp	His	Glu	Lys	Ile	Lys	Lys	Lys	Asp	Pro	Lys	225	230	235	240
Cys	Ile	Phe	Ala	Cys	Glu	Glu	Met	Ser	Gly	Glu	Val	Arg	Phe	Ser	Ser	245	250		255
His	Leu	Pro	Gln	Pro	Asn	Ser	Leu	Cys	Ser	Leu	Ile	Val	Glu	Pro	Met	260	265		270
Glu	Asn	Trp	Leu	Gln	Leu	Met	Leu	Asn	Trp	Asp	Pro	Gln	Gln	Arg	Gly	275	280		285
Gly	Pro	Val	Asp	Leu	Thr	Leu	Lys	Gln	Pro	Arg	Cys	Phe	Val	Leu	Met	290	295		300
Asp	His	Ile	Leu	Asn	Leu	Lys	Ile	Val	His	Ile	Leu	Asn	Met	Thr	Ser	305	310		320
Ala	Lys	Ile	Ile	Ser	Phe	Leu	Leu	Pro	Pro	Asp	Glu	Ser	Leu	His	Ser	325	330		335
Leu	Gln	Ser	Arg	Ile	Glu	Arg	Glu	Thr	Gly	Ile	Asn	Thr	Gly	Ser	Gln	340	345		350
Glu	Leu	Leu	Ser	Glu	Thr	Gly	Ile	Ser	Leu	Asp	Pro	Arg	Lys	Pro	Ala	355	360		365
Ser	Gln	Cys	Val	Leu	Asp	Gly	Val	Arg	Gly	Cys	Asp	Ser	Tyr	Met	Val	370	375		380
Tyr	Leu	Phe	Asp	Lys	Ser	Lys	Thr	Val	Tyr	Glu	Gly	Pro	Phe	Ala	Ser	385	390		400
Arg	Ser	Leu	Ser	Asp	Cys	Val	Asn	Tyr	Ile	Val	Gln	Asp	Ser	Lys	Ile	405	410		415
Gln	Leu	Pro	Ile	Ile	Gln	Leu	Arg	Lys	Val	Trp	Ala	Glu	Ala	Val	His	420	425		430
Tyr	Val	Ser	Gly	Leu	Lys	Glu	Asp	Tyr	Ser	Arg	Leu	Phe	Gln	Gly	Gln	435	440		445
Arg	Ala	Ala	Met	Leu	Ser	Leu	Leu	Arg	Tyr	Asn	Ala	Asn	Leu	Thr	Lys	450	455		460
Met	Lys	Asn	Thr	Leu	Ile	Ser	Ala	Ser	Gln	Gln	Leu	Lys	Ala	Lys	Leu	465	470		480
Glu	Phe	Phe	His	Lys	Ser	Ile	Gln	Leu	Asp	Leu	Glu	Arg	Tyr	Ser	Glu	485	490		495
Gln	Met	Thr	Tyr	Gly	Ile	Ser	Ser	Glu	Lys	Met	Leu	Lys	Ala	Trp	Lys	500	505		510
Glu	Met	Glu	Glu	Lys	Ala	Ile	His	Tyr	Ala	Glu	Val	Gly	Val	Ile	Gly	515	520		525
Tyr	Leu	Glu	Asp	Gln	Ile	Met	Ser	Leu	His	Ala	Glu	Ile	Met	Gly	Leu	530	535		540
Gln	Lys	Ser	Pro	Tyr	Gly	Arg	Arg	Gln	Gly	Asp	Leu	Met	Glu	Ser	Leu	545	550		560
Glu	Gln	Arg	Ala	Ile	Asp	Leu	Tyr	Lys	Gln	Leu	Lys	His	Arg	Pro	Ser	565	570		575
Asp	His	Ser	Tyr	Ser	Asp	Ser	Thr	Glu	Met	Val	Lys	Ile	Ile	Val	His	580	585		590
Thr	Val	Gln	Ser	Gln	Asp	Arg	Val	Leu	Lys	Glu	Leu	Phe	Gly	His	Leu	595	600		605
Ser	Lys	Leu	Leu	Gly	Cys	Lys	Gln	Lys	Ile	Ile	Asp	Leu	Leu	Pro	Lys	610	615		620
Val	Glu	Val	Ala	Leu	Ser	Asn	Ile	Lys	Glu	Ala	Asp	Asn	Thr	Val	Met	625	630		640
Phe	Met	Gln	Gly	Lys	Arg	Gln	Lys	Glu	Ile	Trp	His	Leu	Leu	Lys	Ile	645	650		655
Ala	Cys	Thr	Gln	Ser	Ser	Ala	Arg	Ser	Leu	Val	Gly	Ser	Ser	Leu	Glu	660	665		670
Gly	Ala	Val	Thr	Pro	Gln	Thr	Ser	Ala	Trp	Leu	Pro	Pro	Thr	Ser	Ala	675	680		685
Glu	His	Asp	His	Ser	Leu	Ser	Cys	Val	Val	Thr	Pro	Gln	Asp	Gly	Glu	690	695		700

```

Thr Ser Ala Gln Met Ile Glu Glu Asn Leu Asn Cys Leu Gly His Leu
705                      710                      715                      720
Ser Thr Ile Ile His Glu Ala Asn Glu Glu Gln Gly Asn Ser Met Met
                      725                      730                      735
Asn Leu Asp Trp Ser Trp Leu Thr Glu
                      740                      745

```

```

<210> 17
<211> 756
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence:/note =
        synthetic construct

```

```

<400> 17
Met Ser Trp Ser Pro Ser Leu Thr Thr Gln Thr Cys Gly Ala Trp Glu
 1      5      10      15
Met Lys Glu Arg Leu Gly Thr Gly Gly Phe Gly Asn Val Ile Arg Trp
      20      25      30
His Asn Gln Glu Thr Gly Glu Gln Ile Ala Ile Lys Gln Cys Arg Gln
      35      40      45
Glu Leu Ser Pro Arg Asn Arg Glu Arg Trp Cys Leu Glu Ile Gln Ile
      50      55      60
Met Arg Arg Leu Thr His Pro Asn Val Val Ala Ala Arg Asp Val Pro
      65      70      75      80
Glu Gly Met Gln Asn Leu Ala Pro Asn Asp Leu Pro Leu Leu Ala Met
      85      90      95
Glu Tyr Cys Gln Gly Gly Asp Leu Arg Lys Tyr Leu Asn Gln Phe Glu
      100     105     110
Asn Cys Cys Gly Leu Arg Glu Gly Ala Ile Leu Thr Leu Leu Ser Asp
      115     120     125
Ile Ala Ser Ala Leu Arg Tyr Leu His Glu Asn Arg Ile Ile His Arg
      130     135     140
Asp Leu Lys Pro Glu Asn Ile Val Leu Gln Gln Gly Glu Gln Arg Leu
      145     150     155     160
Ile His Lys Ile Ile Asp Leu Gly Tyr Ala Lys Glu Leu Asp Gln Gly
      165     170     175
Ser Leu Cys Thr Ser Phe Val Gly Thr Leu Gln Tyr Leu Ala Pro Glu
      180     185     190
Leu Leu Glu Gln Gln Lys Tyr Thr Val Thr Val Asp Tyr Trp Ser Phe
      195     200     205
Gly Thr Leu Ala Phe Glu Cys Ile Thr Gly Phe Arg Pro Phe Leu Pro
      210     215     220
Asn Trp Gln Pro Val Gln Trp His Ser Lys Val Arg Gln Lys Ser Glu
      225     230     235     240
Val Asp Ile Val Val Ser Glu Asp Leu Asn Gly Thr Val Lys Phe Ser
      245     250     255
Ser Ser Leu Pro Tyr Pro Asn Asn Leu Asn Ser Val Leu Ala Glu Arg
      260     265     270
Leu Glu Lys Trp Leu Gln Leu Met Leu Met Trp His Pro Arg Gln Arg
      275     280     285
Gly Thr Asp Pro Thr Tyr Gly Pro Asn Gly Cys Phe Lys Ala Leu Asp
      290     295     300
Asp Ile Leu Asn Leu Lys Leu Val His Ile Leu Asn Met Val Thr Gly
      305     310     315     320
Thr Ile His Thr Tyr Pro Val Thr Glu Asp Glu Ser Leu Gln Ser Leu
      325     330     335
Lys Ala Arg Ile Gln Gln Asp Thr Gly Ile Pro Glu Glu Asp Gln Glu
      340     345     350

```

```

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

```

<210> 18

<211> 976

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 18

```

Met Arg Gly Ala Arg Gly Ala Trp Asp Phe Leu Cys Val Leu Leu Leu
 1           5           10           15
Leu Leu Arg Val Gln Thr Gly Ser Ser Gln Pro Ser Val Ser Pro Gly
          20           25           30
Glu Pro Ser Pro Pro Ser Ile His Pro Gly Lys Ser Asp Leu Ile Val
          35           40           45
Arg Val Gly Asp Glu Ile Arg Leu Leu Cys Thr Asp Pro Gly Phe Val
          50           55           60
Lys Trp Thr Phe Glu Ile Leu Asp Glu Thr Asn Glu Asn Lys Gln Asn
65           70           75           80
Glu Trp Ile Thr Glu Lys Ala Glu Ala Thr Asn Thr Gly Lys Tyr Thr
          85           90           95
Cys Thr Asn Lys His Gly Leu Ser Asn Ser Ile Tyr Val Phe Val Arg
          100          105          110
Asp Pro Ala Lys Leu Phe Leu Val Asp Arg Ser Leu Tyr Gly Lys Glu
          115          120          125
Asp Asn Asp Thr Leu Val Arg Cys Pro Leu Thr Asp Pro Glu Val Thr
          130          135          140
Asn Tyr Ser Leu Lys Gly Cys Gln Gly Lys Pro Leu Pro Lys Asp Leu
          145          150          155          160
Arg Phe Ile Pro Asp Pro Lys Ala Gly Ile Met Ile Lys Ser Val Lys
          165          170          175
Arg Ala Tyr His Arg Leu Cys Leu His Cys Ser Val Asp Gln Glu Gly
          180          185          190
Lys Ser Val Leu Ser Glu Lys Phe Ile Leu Lys Val Arg Pro Ala Phe
          195          200          205
Lys Ala Val Pro Val Val Ser Val Ser Lys Ala Ser Tyr Leu Leu Arg
          210          215          220
Glu Gly Glu Glu Phe Thr Val Thr Cys Thr Ile Lys Asp Val Ser Ser
          225          230          235          240
Ser Val Tyr Ser Thr Trp Lys Arg Glu Asn Ser Gln Thr Lys Leu Gln
          245          250          255
Glu Lys Tyr Asn Ser Trp His His Gly Asp Phe Asn Tyr Glu Arg Gln
          260          265          270
Ala Thr Leu Thr Ile Ser Ser Ala Arg Val Asn Asp Ser Gly Val Phe
          275          280          285
Met Cys Tyr Ala Asn Asn Thr Phe Gly Ser Ala Asn Val Thr Thr Thr
          290          295          300
Leu Glu Val Val Asp Lys Gly Phe Ile Asn Ile Phe Pro Met Ile Asn
          305          310          315          320
Thr Thr Val Phe Val Asn Asp Gly Glu Asn Val Asp Leu Ile Val Glu
          325          330          335
Tyr Glu Ala Phe Pro Lys Pro Glu His Gln Gln Trp Ile Tyr Met Asn
          340          345          350
Arg Thr Phe Thr Asp Lys Trp Glu Asp Tyr Pro Lys Ser Glu Asn Glu
          355          360          365
Ser Asn Ile Arg Tyr Val Ser Glu Leu His Leu Thr Arg Leu Lys Gly
          370          375          380
Thr Glu Gly Gly Thr Tyr Thr Phe Leu Val Ser Asn Ser Asp Val Asn
          385          390          395          400
Ala Ala Ile Ala Phe Asn Val Tyr Val Asn Thr Lys Pro Glu Ile Leu
          405          410          415
Thr Tyr Asp Arg Leu Val Asn Gly Met Leu Gln Cys Val Ala Ala Gly
          420          425          430
Phe Pro Glu Pro Thr Ile Asp Trp Tyr Phe Cys Pro Gly Thr Glu Gln
          435          440          445
Arg Cys Ser Ala Ser Val Leu Pro Val Asp Val Gln Thr Leu Asn Ser
          450          455          460

```

Ser	Gly	Pro	Pro	Phe	Gly	Lys	Leu	Val	Val	Gln	Ser	Ser	Ile	Asp	Ser	465	470	475	480
Ser	Ala	Phe	Lys	His	Asn	Gly	Thr	Val	Glu	Cys	Lys	Ala	Tyr	Asn	Asp	485	490	495	
Val	Gly	Lys	Thr	Ser	Ala	Tyr	Phe	Asn	Phe	Ala	Phe	Lys	Gly	Asn	Asn	500	505	510	
Lys	Glu	Gln	Ile	His	Pro	His	Thr	Leu	Phe	Thr	Pro	Leu	Leu	Ile	Gly	515	520	525	
Phe	Val	Ile	Val	Ala	Gly	Met	Met	Cys	Ile	Ile	Val	Met	Ile	Leu	Thr	530	535	540	
Tyr	Lys	Tyr	Leu	Gln	Lys	Pro	Met	Tyr	Glu	Val	Gln	Trp	Lys	Val	Val	545	550	555	560
Glu	Glu	Ile	Asn	Gly	Asn	Asn	Tyr	Val	Tyr	Ile	Asp	Pro	Thr	Gln	Leu	565	570	575	
Pro	Tyr	Asp	His	Lys	Trp	Glu	Phe	Pro	Arg	Asn	Arg	Leu	Ser	Phe	Gly	580	585	590	
Lys	Thr	Leu	Gly	Ala	Gly	Ala	Phe	Gly	Lys	Val	Val	Glu	Ala	Thr	Ala	595	600	605	
Tyr	Gly	Leu	Ile	Lys	Ser	Asp	Ala	Ala	Met	Thr	Val	Ala	Val	Lys	Met	610	615	620	
Leu	Lys	Pro	Ser	Ala	His	Leu	Thr	Glu	Arg	Glu	Ala	Leu	Met	Ser	Glu	625	630	635	640
Leu	Lys	Val	Leu	Ser	Tyr	Leu	Gly	Asn	His	Met	Asn	Ile	Val	Asn	Leu	645	650	655	
Leu	Gly	Ala	Cys	Thr	Ile	Gly	Gly	Pro	Thr	Leu	Val	Ile	Thr	Glu	Tyr	660	665	670	
Cys	Cys	Tyr	Gly	Asp	Leu	Leu	Asn	Phe	Leu	Arg	Arg	Lys	Arg	Asp	Ser	675	680	685	
Phe	Ile	Cys	Ser	Lys	Gln	Glu	Asp	His	Ala	Glu	Ala	Ala	Leu	Tyr	Lys	690	695	700	
Asn	Leu	Leu	His	Ser	Lys	Glu	Ser	Ser	Cys	Ser	Asp	Ser	Thr	Asn	Glu	705	710	715	720
Tyr	Met	Asp	Met	Lys	Pro	Gly	Val	Ser	Tyr	Val	Val	Pro	Thr	Lys	Ala	725	730	735	
Asp	Lys	Arg	Arg	Ser	Val	Arg	Ile	Gly	Ser	Tyr	Ile	Glu	Arg	Asp	Val	740	745	750	
Thr	Pro	Ala	Ile	Met	Glu	Asp	Asp	Glu	Leu	Ala	Leu	Asp	Leu	Glu	Asp	755	760	765	
Leu	Leu	Ser	Phe	Ser	Tyr	Gln	Val	Ala	Lys	Gly	Met	Ala	Phe	Leu	Ala	770	775	780	
Ser	Lys	Asn	Cys	Ile	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Ile	Leu	Leu	785	790	795	800
Thr	His	Gly	Arg	Ile	Thr	Lys	Ile	Cys	Asp	Phe	Gly	Leu	Ala	Arg	Asp	805	810	815	
Ile	Lys	Asn	Asp	Ser	Asn	Tyr	Val	Val	Lys	Gly	Asn	Ala	Arg	Leu	Pro	820	825	830	
Val	Lys	Trp	Met	Ala	Pro	Glu	Ser	Ile	Phe	Asn	Cys	Val	Tyr	Thr	Phe	835	840	845	
Glu	Ser	Asp	Val	Trp	Ser	Tyr	Gly	Ile	Phe	Leu	Trp	Glu	Leu	Phe	Ser	850	855	860	
Leu	Gly	Ser	Ser	Pro	Tyr	Pro	Gly	Met	Pro	Val	Asp	Ser	Lys	Phe	Tyr	865	870	875	880
Lys	Met	Ile	Lys	Glu	Gly	Phe	Arg	Met	Leu	Ser	Pro	Glu	His	Ala	Pro	885	890	895	
Ala	Glu	Met	Tyr	Asp	Ile	Met	Lys	Thr	Cys	Trp	Asp	Ala	Asp	Pro	Leu	900	905	910	
Lys	Arg	Pro	Thr	Phe	Lys	Gln	Ile	Val	Gln	Leu	Ile	Glu	Lys	Gln	Ile	915	920	925	
Ser	Glu	Ser	Thr	Asn	His	Ile	Tyr	Ser	Asn	Leu	Ala	Asn	Cys	Ser	Pro	930	935	940	

Asn	Arg	Gln	Lys	Pro	Val	Val	Asp	His	Ser	Val	Arg	Ile	Asn	Ser	Val
945					950					955					960
Gly	Ser	Thr	Ala	Ser	Ser	Ser	Gln	Pro	Leu	Leu	Val	His	Asp	Asp	Val
				965					970					975	

<210> 19

<211> 509

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 19

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

```

Asn Ile Leu Val Ser Asp Thr Leu Ser Cys Lys Ile Ala Asp Phe Gly
  370                      375                      380
Leu Ala Arg Leu Ile Glu Asp Asn Glu Tyr Thr Ala Arg Glu Gly Ala
385                      390                      395                      400
Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ile Asn Tyr Gly Thr
                      405                      410                      415
Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Thr Glu
                      420                      425                      430
Ile Val Thr His Gly Arg Ile Pro Tyr Pro Gly Met Thr Asn Pro Glu
                      435                      440                      445
Val Ile Gln Asn Leu Glu Arg Gly Tyr Arg Met Val Arg Pro Asp Asn
450                      455                      460
Cys Pro Glu Glu Leu Tyr Gln Leu Met Arg Leu Cys Trp Lys Glu Arg
465                      470                      475                      480
Pro Glu Asp Arg Pro Thr Phe Asp Tyr Leu Arg Ser Val Leu Glu Asp
                      485                      490                      495
Phe Phe Thr Ala Thr Glu Gly Gln Tyr Gln Pro Gln Pro
          500                      505

```

<210> 20

<211> 512

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 20

```

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

```

Gln Phe Gly Glu Val Trp Met Gly Tyr Tyr Asn Asn Ser Thr Lys Val
 260 265 270
 Ala Val Lys Thr Leu Lys Pro Gly Thr Met Ser Val Gln Ala Phe Leu
 275 280 285
 Glu Glu Ala Asn Leu Met Lys Thr Leu Gln His Asp Lys Leu Val Arg
 290 295 300
 Leu Tyr Ala Val Val Thr Arg Glu Glu Pro Ile Tyr Ile Ile Thr Glu
 305 310 315 320
 Tyr Met Ala Lys Gly Ser Leu Leu Asp Phe Leu Lys Ser Asp Glu Gly
 325 330 335
 Gly Lys Val Leu Leu Pro Lys Leu Ile Asp Phe Ser Ala Gln Ile Ala
 340 345 350
 Glu Gly Met Ala Tyr Ile Glu Arg Lys Asn Tyr Ile His Arg Asp Leu
 355 360 365
 Arg Ala Ala Asn Val Leu Val Ser Glu Ser Leu Met Cys Lys Ile Ala
 370 375 380
 Asp Phe Gly Leu Ala Arg Val Ile Glu Asp Asn Glu Tyr Thr Ala Arg
 385 390 395 400
 Glu Gly Ala Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ile Asn
 405 410 415
 Phe Gly Cys Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu
 420 425 430
 Leu Tyr Glu Ile Val Thr Tyr Gly Lys Ile Pro Tyr Pro Gly Arg Thr
 435 440 445
 Asn Ala Asp Val Met Thr Ala Leu Ser Gln Gly Tyr Arg Met Pro Arg
 450 455 460
 Val Glu Asn Cys Pro Asp Glu Leu Tyr Asp Ile Met Lys Met Cys Trp
 465 470 475 480
 Lys Glu Lys Ala Glu Glu Arg Pro Thr Phe Asp Tyr Leu Gln Ser Val
 485 490 495
 Leu Asp Asp Phe Tyr Thr Ala Thr Glu Gly Gln Tyr Gln Gln Gln Pro
 500 505 510

<210> 21

<211> 499

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =

synthetic construct

<400> 21

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

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

<210> 22

<211> 1089

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 22

Met Gly Thr Ser His Pro Ala Phe Leu Val Leu Gly Cys Leu Leu Thr
 1 5 10 15

Gly Leu Ser Leu Ile Leu Cys Gln Leu Ser Leu Pro Ser Ile Leu Pro
 20 25 30
 Asn Glu Asn Glu Lys Val Val Gln Leu Asn Ser Ser Phe Ser Leu Arg
 35 40 45
 Cys Phe Gly Glu Ser Glu Val Ser Trp Gln Tyr Pro Met Ser Glu Glu
 50 55 60
 Glu Ser Ser Asp Val Glu Ile Arg Asn Glu Glu Asn Asn Ser Gly Leu
 65 70 75 80
 Phe Val Thr Val Leu Glu Val Ser Ser Ala Ser Ala Ala His Thr Gly
 85 90 95
 Leu Tyr Thr Cys Tyr Tyr Asn His Thr Gln Thr Glu Glu Asn Glu Leu
 100 105 110
 Glu Gly Arg His Ile Tyr Ile Tyr Val Pro Asp Pro Asp Val Ala Phe
 115 120 125
 Val Pro Leu Gly Met Thr Asp Tyr Leu Val Ile Val Glu Asp Asp Asp
 130 135 140
 Ser Ala Ile Ile Pro Cys Arg Thr Thr Asp Pro Glu Thr Pro Val Thr
 145 150 155 160
 Leu His Asn Ser Glu Gly Val Val Pro Ala Ser Tyr Asp Ser Arg Gln
 165 170 175
 Gly Phe Asn Gly Thr Phe Thr Val Gly Pro Tyr Ile Cys Glu Ala Thr
 180 185 190
 Val Lys Gly Lys Lys Phe Gln Thr Ile Pro Phe Asn Val Tyr Ala Leu
 195 200 205
 Lys Ala Thr Ser Glu Leu Asp Leu Glu Met Glu Ala Leu Lys Thr Val
 210 215 220
 Tyr Lys Ser Gly Glu Thr Ile Val Val Thr Cys Ala Val Phe Asn Asn
 225 230 235 240
 Glu Val Val Asp Leu Gln Trp Thr Tyr Pro Gly Glu Val Lys Gly Lys
 245 250 255
 Gly Ile Thr Met Leu Glu Glu Ile Lys Val Pro Ser Ile Lys Leu Val
 260 265 270
 Tyr Thr Leu Thr Val Pro Glu Ala Thr Val Lys Asp Ser Gly Asp Tyr
 275 280 285
 Glu Cys Ala Ala Arg Gln Ala Thr Arg Glu Val Lys Glu Met Lys Lys
 290 295 300
 Val Thr Ile Ser Val His Glu Lys Gly Phe Ile Glu Ile Lys Pro Thr
 305 310 315 320
 Phe Ser Gln Leu Glu Ala Val Asn Leu His Glu Val Lys His Phe Val
 325 330 335
 Val Glu Val Arg Ala Tyr Pro Pro Arg Ile Ser Trp Leu Lys Asn
 340 345 350
 Asn Leu Thr Leu Ile Glu Asn Leu Thr Glu Ile Thr Thr Asp Val Glu
 355 360 365
 Lys Ile Gln Glu Ile Arg Tyr Arg Ser Lys Leu Lys Leu Ile Arg Ala
 370 375 380
 Lys Glu Glu Asp Ser Gly His Tyr Thr Ile Val Ala Gln Asn Glu Asp
 385 390 395 400
 Ala Val Lys Ser Tyr Thr Phe Glu Leu Leu Thr Gln Val Pro Ser Ser
 405 410 415
 Ile Leu Asp Leu Val Asp Asp His His Gly Ser Thr Gly Gly Gln Thr
 420 425 430
 Val Arg Cys Thr Ala Glu Gly Thr Pro Leu Pro Asp Ile Glu Trp Met
 435 440 445
 Ile Cys Lys Asp Ile Lys Lys Cys Asn Asn Glu Thr Ser Trp Thr Ile
 450 455 460
 Leu Ala Asn Asn Val Ser Asn Ile Ile Thr Glu Ile His Ser Arg Asp
 465 470 475 480
 Arg Ser Thr Val Glu Gly Arg Val Thr Phe Ala Lys Val Glu Glu Thr
 485 490 495

Ile	Ala	Val	Arg	Cys	Leu	Ala	Lys	Asn	Leu	Leu	Gly	Ala	Glu	Asn	Arg		
			500					505					510				
Glu	Leu	Lys	Leu	Val	Ala	Pro	Thr	Leu	Arg	Ser	Glu	Leu	Thr	Val	Ala		
		515					520					525					
Ala	Ala	Val	Leu	Val	Leu	Leu	Val	Ile	Val	Ile	Ile	Ser	Leu	Ile	Val		
	530					535				540							
Leu	Val	Val	Ile	Trp	Lys	Gln	Lys	Pro	Arg	Tyr	Glu	Ile	Arg	Trp	Arg		
545					550					555					560		
Val	Ile	Glu	Ser	Ile	Ser	Pro	Asp	Gly	His	Glu	Tyr	Ile	Tyr	Val	Asp		
			565					570						575			
Pro	Met	Gln	Leu	Pro	Tyr	Asp	Ser	Arg	Trp	Glu	Phe	Pro	Arg	Asp	Gly		
		580						585					590				
Leu	Val	Leu	Gly	Arg	Val	Leu	Gly	Ser	Gly	Ala	Phe	Gly	Lys	Val	Val		
	595						600					605					
Glu	Gly	Thr	Ala	Tyr	Gly	Leu	Ser	Arg	Ser	Gln	Pro	Val	Met	Lys	Val		
	610					615				620							
Ala	Val	Lys	Met	Leu	Lys	Pro	Thr	Ala	Arg	Ser	Ser	Glu	Lys	Gln	Ala		
625					630					635					640		
Leu	Met	Ser	Glu	Leu	Lys	Ile	Met	Thr	His	Leu	Gly	Pro	His	Leu	Asn		
			645					650						655			
Ile	Val	Asn	Leu	Leu	Gly	Ala	Cys	Thr	Lys	Ser	Gly	Pro	Ile	Tyr	Ile		
		660						665					670				
Ile	Thr	Glu	Tyr	Cys	Phe	Tyr	Gly	Asp	Leu	Val	Asn	Tyr	Leu	His	Lys		
	675						680					685					
Asn	Arg	Asp	Ser	Phe	Leu	Ser	His	His	Pro	Glu	Lys	Pro	Lys	Lys	Glu		
	690					695					700						
Leu	Asp	Ile	Phe	Gly	Leu	Asn	Pro	Ala	Asp	Glu	Ser	Thr	Arg	Ser	Tyr		
705					710					715					720		
Val	Ile	Leu	Ser	Phe	Glu	Asn	Asn	Gly	Asp	Tyr	Met	Asp	Met	Lys	Gln		
			725					730						735			
Ala	Asp	Thr	Thr	Gln	Tyr	Val	Pro	Met	Leu	Glu	Arg	Lys	Glu	Val	Ser		
		740						745					750				
Lys	Tyr	Ser	Asp	Ile	Gln	Arg	Ser	Leu	Tyr	Asp	Arg	Pro	Ala	Ser	Tyr		
	755						760					765					
Lys	Lys	Lys	Ser	Met	Leu	Asp	Ser	Glu	Val	Lys	Asn	Leu	Leu	Ser	Asp		
	770					775				780							
Asp	Asn	Ser	Glu	Gly	Leu	Thr	Leu	Leu	Asp	Leu	Leu	Ser	Phe	Thr	Tyr		
785					790					795					800		
Gln	Val	Ala	Arg	Gly	Met	Glu	Phe	Leu	Ala	Ser	Lys	Asn	Cys	Val	His		
			805						810					815			
Arg	Asp	Leu	Ala	Arg	Asn	Val	Leu	Leu	Ala	Gln	Gly	Lys	Ile	Val			
		820					825					830					
Lys	Ile	Cys	Asp	Phe	Gly	Leu	Ala	Arg	Asp	Ile	Met	His	Asp	Ser	Asn		
	835						840					845					
Tyr	Val	Ser	Lys	Gly	Ser	Thr	Phe	Leu	Pro	Val	Lys	Trp	Met	Ala	Pro		
	850					855					860						
Glu	Ser	Ile	Phe	Asp	Asn	Leu	Tyr	Thr	Thr	Leu	Ser	Asp	Val	Trp	Ser		
865					870					875					880		
Tyr	Gly	Ile	Leu	Leu	Trp	Glu	Ile	Phe	Ser	Leu	Gly	Gly	Thr	Pro	Tyr		
			885						890					895			
Pro	Gly	Met	Met	Val	Asp	Ser	Thr	Phe	Tyr	Asn	Lys	Ile	Lys	Ser	Gly		
		900						905					910				
Tyr	Arg	Met	Ala	Lys	Pro	Asp	His	Ala	Thr	Ser	Glu	Val	Tyr	Glu	Ile		
	915						920					925					
Met	Val	Lys	Cys	Trp	Asn	Ser	Glu	Pro	Glu	Lys	Arg	Pro	Ser	Phe	Tyr		
	930					935					940						
His	Leu	Ser	Glu	Ile	Val	Glu	Asn	Leu	Leu	Pro	Gly	Gln	Tyr	Lys	Lys		
945					950					955					960		
Ser	Tyr	Glu	Lys	Ile	His	Leu	Asp	Phe	Leu	Lys	Ser	Asp	His	Pro	Ala		
			965						970					975			

Val Ala Arg Met Arg Val Asp Ser Asp Asn Ala Tyr Ile Gly Val Thr
 980 985 990
 Tyr Lys Asn Glu Glu Asp Lys Leu Lys Asp Trp Glu Gly Gly Leu Asp
 995 1000 1005
 Glu Gln Arg Leu Ser Ala Asp Ser Gly Tyr Ile Ile Pro Leu Pro Asp
 1010 1015 1020
 Ile Asp Pro Val Pro Glu Glu Asp Leu Gly Lys Arg Asn Arg His
 1025 1030 1035 1040
 Ser Ser Gln Thr Ser Glu Glu Ser Ala Ile Glu Thr Gly Ser Ser Ser
 1045 1050 1055
 Ser Thr Phe Ile Lys Arg Glu Asp Glu Thr Ile Glu Asp Ile Asp Met
 1060 1065 1070
 Met Asp Asp Ile Gly Ile Asp Ser Ser Asp Leu Val Glu Asp Ser Phe
 1075 1080 1085
 Leu

<210> 23

<211> 1106

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 23

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

Arg	Ser	Ile	Leu	His	Ile	Pro	Ser	Ala	Glu	Leu	Glu	Asp	Ser	Gly	Thr		
		275					280					285					
Tyr	Thr	Cys	Asn	Val	Thr	Glu	Ser	Val	Asn	Asp	His	Gln	Asp	Glu	Lys		
		290				295					300						
Ala	Ile	Asn	Ile	Thr	Val	Val	Glu	Ser	Gly	Tyr	Val	Arg	Leu	Leu	Gly		
305					310					315					320		
Glu	Val	Gly	Thr	Leu	Gln	Phe	Ala	Glu	Leu	His	Arg	Ser	Arg	Thr	Leu		
				325					330					335			
Gln	Val	Val	Phe	Glu	Ala	Tyr	Pro	Pro	Pro	Thr	Val	Leu	Trp	Phe	Lys		
			340				345						350				
Asp	Asn	Arg	Thr	Leu	Gly	Asp	Ser	Ser	Ala	Gly	Glu	Ile	Ala	Leu	Ser		
		355				360						365					
Thr	Arg	Asn	Val	Ser	Glu	Thr	Arg	Tyr	Val	Ser	Glu	Leu	Thr	Leu	Val		
		370				375					380						
Arg	Val	Lys	Val	Ala	Glu	Ala	Gly	His	Tyr	Thr	Met	Arg	Ala	Phe	His		
385					390					395					400		
Glu	Asp	Ala	Glu	Val	Gln	Leu	Ser	Phe	Gln	Leu	Gln	Ile	Asn	Val	Pro		
				405					410					415			
Val	Arg	Val	Leu	Glu	Leu	Ser	Glu	Ser	His	Pro	Asp	Ser	Gly	Glu	Gln		
			420					425					430				
Thr	Val	Arg	Cys	Arg	Gly	Arg	Gly	Met	Pro	Gln	Pro	Asn	Ile	Ile	Trp		
		435					440					445					
Ser	Ala	Cys	Arg	Asp	Leu	Lys	Arg	Cys	Pro	Arg	Glu	Leu	Pro	Pro	Thr		
		450				455					460						
Leu	Leu	Gly	Asn	Ser	Ser	Glu	Glu	Glu	Ser	Gln	Leu	Glu	Thr	Asn	Val		
465					470					475					480		
Thr	Tyr	Trp	Glu	Glu	Gln	Glu	Phe	Glu	Val	Val	Ser	Thr	Leu	Arg			
			485					490					495				
Leu	Gln	His	Val	Asp	Arg	Pro	Leu	Ser	Val	Arg	Cys	Thr	Leu	Arg	Asn		
			500					505					510				
Ala	Val	Gly	Gln	Asp	Thr	Gln	Glu	Val	Ile	Val	Val	Pro	His	Ser	Leu		
		515						520				525					
Pro	Phe	Lys	Val	Val	Val	Ile	Ser	Ala	Ile	Leu	Ala	Leu	Val	Val	Leu		
		530				535					540						
Thr	Ile	Ile	Ser	Leu	Ile	Ile	Leu	Ile	Met	Leu	Trp	Gln	Lys	Lys	Pro		
545					550					555					560		
Arg	Tyr	Glu	Ile	Arg	Trp	Lys	Val	Ile	Glu	Ser	Val	Ser	Ser	Asp	Gly		
				565					570					575			
His	Glu	Tyr	Ile	Tyr	Val	Asp	Pro	Met	Gln	Leu	Pro	Tyr	Asp	Ser	Thr		
			580					585					590				
Trp	Glu	Leu	Pro	Arg	Asp	Gln	Leu	Val	Leu	Gly	Arg	Thr	Leu	Gly	Ser		
		595					600					605					
Gly	Ala	Phe	Gly	Gln	Val	Val	Glu	Ala	Thr	Ala	His	Gly	Leu	Ser	His		
		610				615					620						
Ser	Gln	Ala	Thr	Met	Lys	Val	Ala	Val	Lys	Met	Leu	Lys	Ser	Thr	Ala		
625					630					635					640		
Arg	Ser	Ser	Glu	Lys	Gln	Ala	Leu	Met	Ser	Glu	Leu	Lys	Ile	Met	Ser		
				645					650					655			
His	Leu	Gly	Pro	His	Leu	Asn	Val	Val	Asn	Leu	Leu	Gly	Ala	Cys	Thr		
			660					665					670				
Lys	Gly	Gly	Pro	Ile	Tyr	Ile	Ile	Thr	Glu	Tyr	Cys	Arg	Tyr	Gly	Asp		
		675						680				685					
Leu	Val	Asp	Tyr	Leu	His	Arg	Asn	Lys	His	Thr	Phe	Leu	Gln	His	His		
		690				695					700						
Ser	Asp	Lys	Arg	Arg	Pro	Pro	Ser	Ala	Glu	Leu	Tyr	Ser	Asn	Ala	Leu		
705					710					715					720		
Pro	Val	Gly	Leu	Pro	Leu	Pro	Ser	His	Val	Ser	Leu	Thr	Gly	Glu	Ser		
				725					730					735			
Asp	Gly	Gly	Tyr	Met	Asp	Met	Ser	Lys	Asp	Glu	Ser	Val	Asp	Tyr	Val		
			740					745					750				

Pro Met Leu Asp Met Lys Gly Asp Val Lys Tyr Ala Asp Ile Glu Ser
 755 760 765
 Ser Asn Tyr Met Ala Pro Tyr Asp Asn Tyr Val Pro Ser Ala Pro Glu
 770 775 780
 Arg Thr Cys Arg Ala Thr Leu Ile Asn Glu Ser Pro Val Leu Ser Tyr
 785 790 795 800
 Met Asp Leu Val Gly Phe Ser Tyr Gln Val Ala Asn Gly Met Glu Phe
 805 810 815
 Leu Ala Ser Lys Asn Cys Val His Arg Asp Leu Ala Ala Arg Asn Val
 820 825 830
 Leu Ile Cys Glu Gly Lys Leu Val Lys Ile Cys Asp Phe Gly Leu Ala
 835 840 845
 Arg Asp Ile Met Arg Asp Ser Asn Tyr Ile Ser Lys Gly Ser Thr Phe
 850 855 860
 Leu Pro Leu Lys Trp Met Ala Pro Glu Ser Ile Phe Asn Ser Leu Tyr
 865 870 875 880
 Thr Thr Leu Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Trp Glu Ile
 885 890 895
 Phe Thr Leu Gly Gly Thr Pro Tyr Pro Glu Leu Pro Met Asn Glu Gln
 900 905 910
 Phe Tyr Asn Ala Ile Lys Arg Gly Tyr Arg Met Ala Gln Pro Ala His
 915 920 925
 Ala Ser Asp Glu Ile Tyr Glu Ile Met Gln Lys Cys Trp Glu Glu Lys
 930 935 940
 Phe Glu Ile Arg Pro Pro Phe Ser Gln Leu Val Leu Leu Glu Arg
 945 950 955 960
 Leu Leu Gly Glu Gly Tyr Lys Lys Lys Tyr Gln Gln Val Asp Glu Glu
 965 970 975
 Phe Leu Arg Ser Asp His Pro Ala Ile Leu Arg Ser Gln Ala Arg Leu
 980 985 990
 Pro Gly Phe His Gly Leu Arg Ser Pro Leu Asp Thr Ser Ser Val Leu
 995 1000 1005
 Tyr Thr Ala Val Gln Pro Asn Glu Gly Asp Asn Asp Tyr Ile Ile Pro
 1010 1015 1020
 Leu Pro Asp Pro Lys Pro Glu Val Ala Asp Glu Gly Pro Leu Glu Gly
 1025 1030 1035 1040
 Ser Pro Ser Leu Ala Ser Ser Thr Leu Asn Glu Val Asn Thr Ser Ser
 1045 1050 1055
 Thr Ile Ser Cys Asp Ser Pro Leu Glu Pro Gln Asp Glu Pro Glu Pro
 1060 1065 1070
 Glu Pro Gln Leu Glu Leu Gln Val Glu Pro Glu Pro Glu Leu Glu Gln
 1075 1080 1085
 Leu Pro Asp Ser Gly Cys Pro Ala Pro Arg Ala Glu Ala Glu Asp Ser
 1090 1095 1100
 Phe Leu
 1105

<210> 24

<211> 1114

<212> PRT

<213> Artificial Sequence

<220>

 <223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 24

Met Ala Lys Ala Thr Ser Gly Ala Ala Gly Leu Arg Leu Leu Leu
 1 5 10 15
 Leu Leu Leu Pro Leu Leu Gly Lys Val Ala Leu Gly Leu Tyr Phe Ser
 20 25 30

Arg	Asp	Ala	Tyr	Trp	Glu	Lys	Leu	Tyr	Val	Asp	Gln	Ala	Ala	Gly	Thr
		35					40					45			
Pro	Leu	Leu	Tyr	Val	His	Ala	Leu	Arg	Asp	Ala	Pro	Glu	Glu	Val	Pro
	50					55					60				
Ser	Phe	Arg	Leu	Gly	Gln	His	Leu	Tyr	Gly	Thr	Tyr	Arg	Thr	Arg	Leu
65					70					75					80
His	Glu	Asn	Asn	Trp	Ile	Cys	Ile	Gln	Glu	Asp	Thr	Gly	Leu	Leu	Tyr
				85						90				95	
Leu	Asn	Arg	Ser	Leu	Asp	His	Ser	Ser	Trp	Glu	Lys	Leu	Ser	Val	Arg
			100					105					110		
Asn	Arg	Gly	Phe	Pro	Leu	Leu	Thr	Val	Tyr	Leu	Lys	Val	Phe	Leu	Ser
		115					120					125			
Pro	Thr	Ser	Leu	Arg	Glu	Gly	Glu	Cys	Gln	Trp	Pro	Gly	Cys	Ala	Arg
	130					135					140				
Val	Tyr	Phe	Ser	Phe	Phe	Asn	Thr	Ser	Phe	Pro	Ala	Cys	Ser	Ser	Leu
145					150					155					160
Lys	Pro	Arg	Glu	Leu	Cys	Phe	Pro	Glu	Thr	Arg	Pro	Ser	Phe	Arg	Ile
				165					170					175	
Arg	Glu	Asn	Arg	Pro	Pro	Gly	Thr	Phe	His	Gln	Phe	Arg	Leu	Leu	Pro
			180				185					190			
Val	Gln	Phe	Leu	Cys	Pro	Asn	Ile	Ser	Val	Ala	Tyr	Arg	Leu	Leu	Glu
		195					200					205			
Gly	Glu	Gly	Leu	Pro	Phe	Arg	Cys	Ala	Pro	Asp	Ser	Leu	Glu	Val	Ser
	210					215					220				
Thr	Arg	Trp	Ala	Leu	Asp	Arg	Glu	Gln	Arg	Glu	Lys	Tyr	Glu	Leu	Val
225					230					235					240
Ala	Val	Cys	Thr	Val	His	Ala	Gly	Ala	Arg	Glu	Glu	Val	Val	Met	Val
				245					250					255	
Pro	Phe	Pro	Val	Thr	Val	Tyr	Asp	Glu	Asp	Asp	Ser	Ala	Pro	Thr	Phe
			260					265					270		
Pro	Ala	Gly	Val	Asp	Thr	Ala	Ser	Ala	Val	Val	Glu	Phe	Lys	Arg	Lys
		275					280					285			
Glu	Asp	Thr	Val	Val	Ala	Thr	Leu	Arg	Val	Phe	Asp	Ala	Asp	Val	Val
	290					295					300				
Pro	Ala	Ser	Gly	Glu	Leu	Val	Arg	Arg	Tyr	Thr	Ser	Thr	Leu	Leu	Pro
305					310					315					320
Gly	Asp	Thr	Trp	Ala	Gln	Gln	Thr	Phe	Arg	Val	Glu	His	Trp	Pro	Asn
				325					330					335	
Glu	Thr	Ser	Val	Gln	Ala	Asn	Gly	Ser	Phe	Val	Arg	Ala	Thr	Val	His
			340					345					350		
Asp	Tyr	Arg	Leu	Val	Leu	Asn	Arg	Asn	Leu	Ser	Ile	Ser	Glu	Asn	Arg
		355					360					365			
Thr	Met	Gln	Leu	Ala	Val	Leu	Val	Asn	Asp	Ser	Asp	Phe	Gln	Gly	Pro
	370					375					380				
Gly	Ala	Gly	Val	Leu	Leu	Leu	His	Phe	Asn	Val	Ser	Val	Leu	Pro	Val
385					390					395					400
Ser	Leu	His	Leu	Pro	Ser	Thr	Tyr	Ser	Leu	Ser	Val	Ser	Arg	Arg	Ala
				405					410					415	
Arg	Arg	Phe	Ala	Gln	Ile	Gly	Lys	Val	Cys	Val	Glu	Asn	Cys	Gln	Ala
			420					425					430		
Phe	Ser	Gly	Ile	Asn	Val	Gln	Tyr	Lys	Leu	His	Ser	Ser	Gly	Ala	Asn
		435					440					445			
Cys	Ser	Thr	Leu	Gly	Val	Val	Thr	Ser	Ala	Glu	Asp	Thr	Ser	Gly	Ile
	450					455					460				
Leu	Phe	Val	Asn	Asp	Thr	Lys	Ala	Leu	Arg	Arg	Pro	Lys	Cys	Ala	Glu
465					470					475					480
Leu	His	Tyr	Met	Val	Val	Ala	Thr	Asp	Gln	Gln	Thr	Ser	Arg	Gln	Ala
				485					490					495	
Gln	Ala	Gln	Leu	Leu	Val	Thr	Val	Glu	Gly	Ser	Tyr	Val	Ala	Glu	Glu
			500					505						510	

Ala	Gly	Cys	Pro	Leu	Ser	Cys	Ala	Val	Ser	Lys	Arg	Arg	Leu	Glu	Cys		
		515					520					525					
Glu	Glu	Cys	Gly	Gly	Leu	Gly	Ser	Pro	Thr	Gly	Arg	Cys	Glu	Trp	Arg		
		530				535					540						
Gln	Gly	Asp	Gly	Lys	Gly	Ile	Thr	Arg	Asn	Phe	Ser	Thr	Cys	Ser	Pro		
545					550					555					560		
Ser	Thr	Lys	Thr	Cys	Pro	Asp	Gly	His	Cys	Asp	Val	Val	Glu	Thr	Gln		
				565					570						575		
Asp	Ile	Asn	Ile	Cys	Pro	Gln	Asp	Cys	Leu	Arg	Gly	Ser	Ile	Val	Gly		
		580						585					590				
Gly	His	Glu	Pro	Gly	Glu	Pro	Arg	Gly	Ile	Lys	Ala	Gly	Tyr	Gly	Thr		
		595					600					605					
Cys	Asn	Cys	Phe	Pro	Glu	Glu	Lys	Cys	Phe	Cys	Glu	Pro	Glu	Asp			
		610				615				620							
Ile	Gln	Asp	Pro	Leu	Cys	Asp	Glu	Leu	Cys	Arg	Thr	Val	Ile	Ala	Ala		
625					630					635					640		
Ala	Val	Leu	Phe	Ser	Phe	Ile	Val	Ser	Val	Leu	Leu	Ser	Ala	Phe	Cys		
				645					650					655			
Ile	His	Cys	Tyr	His	Lys	Phe	Ala	His	Lys	Pro	Pro	Ile	Ser	Ser	Ala		
		660						665					670				
Glu	Met	Thr	Phe	Arg	Arg	Pro	Ala	Gln	Ala	Phe	Pro	Val	Ser	Tyr	Ser		
		675					680					685					
Ser	Ser	Gly	Ala	Arg	Arg	Pro	Ser	Leu	Asp	Ser	Met	Glu	Asn	Gln	Val		
		690				695					700						
Ser	Val	Asp	Ala	Phe	Lys	Ile	Leu	Glu	Asp	Pro	Lys	Trp	Glu	Phe	Pro		
705					710					715					720		
Arg	Lys	Asn	Leu	Val	Leu	Gly	Lys	Thr	Leu	Gly	Glu	Gly	Glu	Phe	Gly		
				725					730						735		
Lys	Val	Val	Lys	Ala	Thr	Ala	Phe	His	Leu	Lys	Gly	Arg	Ala	Gly	Tyr		
		740						745					750				
Thr	Thr	Val	Ala	Val	Lys	Met	Leu	Lys	Glu	Asn	Ala	Ser	Pro	Ser	Glu		
		755					760					765					
Leu	Arg	Asp	Leu	Leu	Ser	Glu	Phe	Asn	Val	Leu	Lys	Gln	Val	Asn	His		
		770				775					780						
Pro	His	Val	Ile	Lys	Leu	Tyr	Gly	Ala	Cys	Ser	Gln	Asp	Gly	Pro	Leu		
785					790					795					800		
Leu	Leu	Ile	Val	Glu	Tyr	Ala	Lys	Tyr	Gly	Ser	Leu	Arg	Gly	Phe	Leu		
				805					810					815			
Arg	Glu	Ser	Arg	Lys	Val	Gly	Pro	Gly	Tyr	Leu	Gly	Ser	Gly	Gly	Ser		
				820				825					830				
Arg	Asn	Ser	Ser	Ser	Leu	Asp	His	Pro	Asp	Glu	Arg	Ala	Leu	Thr	Met		
		835					840					845					
Gly	Asp	Leu	Ile	Ser	Phe	Ala	Trp	Gln	Ile	Ser	Gln	Gly	Met	Gln	Tyr		
		850				855				860							
Leu	Ala	Glu	Met	Lys	Leu	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Ile		
865					870					875					880		
Leu	Val	Ala	Glu	Gly	Arg	Lys	Met	Lys	Ile	Ser	Asp	Phe	Gly	Leu	Ser		
				885					890					895			
Arg	Asp	Val	Tyr	Glu	Glu	Asp	Ser	Tyr	Val	Lys	Arg	Ser	Gln	Gly	Arg		
		900						905					910				
Ile	Pro	Val	Lys	Trp	Met	Ala	Ile	Glu	Ser	Leu	Phe	Asp	His	Ile	Tyr		
		915					920						925				
Thr	Thr	Gln	Ser	Asp	Val	Trp	Ser	Phe	Gly	Val	Leu	Leu	Trp	Glu	Ile		
		930				935						940					
Val	Thr	Leu	Gly	Gly	Asn	Pro	Tyr	Pro	Gly	Ile	Pro	Pro	Glu	Arg	Leu		
945					950					955					960		
Phe	Asn	Leu	Leu	Lys	Thr	Gly	His	Arg	Met	Glu	Arg	Pro	Asp	Asn	Cys		
				965					970					975			
Ser	Glu	Glu	Met	Tyr	Arg	Leu	Met	Leu	Gln	Cys	Trp	Lys	Gln	Glu	Pro		
			980					985						990			

```

Asp Lys Arg Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys Met
    995                      1000                      1005
Met Val Lys Arg Arg Asp Tyr Leu Asp Leu Ala Ala Ser Thr Pro Ser
    1010                      1015                      1020
Asp Ser Leu Ile Tyr Asp Asp Gly Leu Ser Glu Glu Glu Thr Pro Leu
1025                      1030                      1035                      1040
Val Asp Cys Asn Asn Ala Pro Leu Pro Arg Ala Leu Pro Ser Thr Trp
    1045                      1050                      1055
Ile Glu Asn Lys Leu Tyr Gly Met Ser Asp Pro Asn Trp Pro Gly Glu
    1060                      1065                      1070
Ser Pro Val Pro Leu Thr Arg Ala Asp Gly Thr Asn Thr Gly Phe Pro
    1075                      1080                      1085
Arg Tyr Pro Asn Asp Ser Val Tyr Ala Asn Trp Met Leu Ser Pro Ser
    1090                      1095                      1100
Ala Ala Lys Leu Met Asp Thr Phe Asp Ser
1105                      1110

```

<210> 25

<211> 540

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 25

```

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

```

```

Trp Ala Gln Asn Pro Asp Glu Arg Pro Ser Phe Leu Lys Cys Leu Ile
      275                      280                      285
Glu Leu Glu Pro Val Leu Arg Thr Phe Glu Glu Ile Thr Phe Leu Glu
      290                      295                      300
Ala Val Ile Gln Leu Lys Lys Thr Lys Leu Gln Ser Val Ser Ser Ala
305                      310                      315                      320
Ile His Leu Cys Asp Lys Lys Lys Met Glu Leu Ser Leu Asn Ile Pro
      325                      330                      335
Val Asn His Gly Pro Gln Glu Glu Ser Cys Gly Ser Ser Gln Leu His
      340                      345                      350
Glu Asn Ser Gly Ser Pro Glu Thr Ser Arg Ser Leu Pro Ala Pro Gln
      355                      360                      365
Asp Asn Asp Phe Leu Ser Arg Lys Ala Gln Asp Cys Tyr Phe Met Lys
      370                      375                      380
Leu His His Cys Pro Gly Asn His Ser Trp Asp Ser Thr Ile Ser Gly
385                      390                      395                      400
Ser Gln Arg Ala Ala Phe Cys Asp His Lys Thr Thr Pro Cys Ser Ser
      405                      410                      415
Ala Ile Ile Asn Pro Leu Ser Thr Ala Gly Asn Ser Glu Arg Leu Gln
      420                      425                      430
Pro Gly Ile Ala Gln Gln Trp Ile Gln Ser Lys Arg Glu Asp Ile Val
      435                      440                      445
Asn Gln Met Thr Glu Ala Cys Leu Asn Gln Ser Leu Asp Ala Leu Leu
      450                      455                      460
Ser Arg Asp Leu Ile Met Lys Glu Asp Tyr Glu Leu Val Ser Thr Lys
465                      470                      475                      480
Pro Thr Arg Thr Ser Lys Val Arg Gln Leu Leu Asp Thr Thr Asp Ile
      485                      490                      495
Gln Gly Glu Glu Phe Ala Lys Val Ile Val Gln Lys Leu Lys Asp Asn
      500                      505                      510
Lys Gln Met Gly Leu Gln Pro Tyr Pro Glu Ile Leu Val Val Ser Arg
      515                      520                      525
Ser Pro Ser Leu Asn Leu Leu Gln Asn Lys Ser Met
      530                      535                      540

```

<210> 26

<211> 536

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 26

```

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

```

```

Thr Gly Gln Thr Gly Tyr Ile Pro Ser Asn Tyr Val Ala Pro Ser Asp
  130                      135                      140
Ser Ile Gln Ala Glu Glu Trp Tyr Phe Gly Lys Ile Thr Arg Arg Glu
145                      150                      155                      160
Ser Glu Arg Leu Leu Leu Asn Ala Glu Asn Pro Arg Gly Thr Phe Leu
                      165                      170                      175
Val Arg Glu Ser Glu Thr Thr Lys Gly Ala Tyr Cys Leu Ser Val Ser
                      180                      185                      190
Asp Phe Asp Asn Ala Lys Gly Leu Asn Val Lys His Tyr Lys Ile Arg
                      195                      200                      205
Lys Leu Asp Ser Gly Gly Phe Tyr Ile Thr Ser Arg Thr Gln Phe Asn
210                      215                      220
Ser Leu Gln Gln Leu Val Ala Tyr Tyr Ser Lys His Ala Asp Gly Leu
225                      230                      235                      240
Cys His Arg Leu Thr Thr Val Cys Pro Thr Ser Lys Pro Gln Thr Gln
                      245                      250                      255
Gly Leu Ala Lys Asp Ala Trp Glu Ile Pro Arg Glu Ser Leu Arg Leu
                      260                      265                      270
Glu Val Lys Leu Gly Gln Gly Cys Phe Gly Glu Val Trp Met Gly Thr
                      275                      280                      285
Trp Asn Gly Thr Thr Arg Val Ala Ile Lys Thr Leu Lys Pro Gly Thr
290                      295                      300
Met Ser Pro Glu Ala Phe Leu Gln Glu Ala Gln Val Met Lys Lys Leu
305                      310                      315                      320
Arg His Glu Lys Leu Val Gln Leu Tyr Ala Val Val Ser Glu Glu Pro
                      325                      330                      335
Ile Tyr Ile Val Thr Glu Tyr Met Ser Lys Gly Ser Leu Leu Asp Phe
                      340                      345                      350
Leu Lys Gly Glu Thr Gly Lys Tyr Leu Arg Leu Pro Gln Leu Val Asp
                      355                      360                      365
Met Ala Ala Gln Ile Ala Ser Gly Met Ala Tyr Val Glu Arg Met Asn
370                      375                      380
Tyr Val His Arg Asp Leu Arg Ala Ala Asn Ile Leu Val Gly Glu Asn
385                      390                      395                      400
Leu Val Cys Lys Val Ala Asp Phe Gly Leu Ala Arg Leu Ile Glu Asp
                      405                      410                      415
Asn Glu Tyr Thr Ala Arg Gln Gly Ala Lys Phe Pro Ile Lys Trp Thr
420                      425                      430
Ala Pro Glu Ala Ala Leu Tyr Gly Arg Phe Thr Ile Lys Ser Asp Val
435                      440                      445
Trp Ser Phe Gly Ile Leu Leu Thr Glu Leu Thr Thr Lys Gly Arg Val
450                      455                      460
Pro Tyr Pro Gly Met Val Asn Arg Glu Val Leu Asp Gln Val Glu Arg
465                      470                      475                      480
Gly Tyr Arg Met Pro Cys Pro Pro Glu Cys Pro Glu Ser Leu His Asp
                      485                      490                      495
Leu Met Cys Gln Cys Trp Arg Lys Glu Pro Glu Glu Arg Pro Thr Phe
500                      505                      510
Glu Tyr Leu Gln Ala Phe Leu Glu Asp Tyr Phe Thr Ser Thr Glu Pro
515                      520                      525
Gln Tyr Gln Pro Gly Glu Asn Leu
530                      535

```

<210> 27

<211> 635

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 27

```

Met Ala Ser Ser Gly Met Ala Asp Ser Ala Asn His Leu Pro Phe Phe
 1          5          10          15
Phe Gly Asn Ile Thr Arg Glu Glu Ala Glu Asp Tyr Leu Val Gln Gly
          20          25          30
Gly Met Ser Asp Gly Leu Tyr Leu Leu Arg Gln Ser Arg Asn Tyr Leu
          35          40          45
Gly Gly Phe Ala Leu Ser Val Ala His Gly Arg Lys Ala His His Tyr
 50          55          60
Thr Ile Glu Arg Glu Leu Asn Gly Thr Tyr Ala Ile Ala Gly Gly Arg
65          70          75          80
Thr His Ala Ser Pro Ala Asp Leu Cys His Tyr His Ser Gln Glu Ser
          85          90          95
Asp Gly Leu Val Cys Leu Leu Lys Lys Pro Phe Asn Arg Pro Gln Gly
          100          105          110
Val Gln Pro Lys Thr Gly Pro Phe Glu Asp Leu Lys Glu Asn Leu Ile
          115          120          125
Arg Glu Tyr Val Lys Gln Thr Trp Asn Leu Gln Gly Gln Ala Leu Glu
130          135          140
Gln Ala Ile Ile Ser Gln Lys Pro Gln Leu Glu Lys Leu Ile Ala Thr
145          150          155          160
Thr Ala His Glu Lys Met Pro Trp Phe His Gly Lys Ile Ser Arg Glu
          165          170          175
Glu Ser Glu Gln Ile Val Leu Ile Gly Ser Lys Thr Asn Gly Lys Phe
          180          185          190
Leu Ile Arg Ala Arg Asp Asn Asn Gly Ser Tyr Ala Leu Cys Leu Leu
          195          200          205
His Glu Gly Lys Val Leu His Tyr Arg Ile Asp Lys Asp Lys Thr Gly
210          215          220
Lys Leu Ser Ile Pro Glu Gly Lys Lys Phe Asp Thr Leu Trp Gln Leu
225          230          235          240
Val Glu His Tyr Ser Tyr Lys Ala Asp Gly Leu Leu Arg Val Leu Thr
          245          250          255
Val Pro Cys Gln Lys Ile Gly Thr Gln Gly Asn Val Asn Phe Gly Gly
          260          265          270
Arg Pro Gln Leu Pro Gly Ser His Pro Ala Thr Trp Ser Ala Gly Gly
          275          280          285
Ile Ile Ser Arg Ile Lys Ser Tyr Ser Phe Pro Lys Pro Gly His Arg
290          295          300
Lys Ser Ser Pro Ala Gln Gly Asn Arg Gln Glu Ser Thr Val Ser Phe
305          310          315          320
Asn Pro Tyr Glu Pro Glu Leu Ala Pro Trp Ala Ala Asp Lys Gly Pro
          325          330          335
Gln Arg Glu Ala Leu Pro Met Asp Thr Glu Val Tyr Glu Ser Pro Tyr
          340          345          350
Ala Asp Pro Glu Glu Ile Arg Pro Lys Glu Val Tyr Leu Asp Arg Lys
          355          360          365
Leu Leu Thr Leu Glu Asp Lys Glu Leu Gly Ser Gly Asn Phe Gly Thr
          370          375          380
Val Lys Lys Gly Tyr Tyr Gln Met Lys Lys Val Val Lys Thr Val Ala
385          390          395          400
Val Lys Ile Leu Lys Asn Glu Ala Asn Asp Pro Ala Leu Lys Asp Glu
          405          410          415
Leu Leu Ala Glu Ala Asn Val Met Gln Gln Leu Asp Asn Pro Tyr Ile
          420          425          430
Val Arg Met Ile Gly Ile Cys Glu Ala Glu Ser Trp Met Leu Val Met
          435          440          445
Glu Met Ala Glu Leu Gly Pro Leu Asn Lys Tyr Leu Gln Gln Asn Arg
450          455          460

```

```

His Val Lys Asp Lys Asn Ile Ile Glu Leu Val His Gln Val Ser Met
465                               470                               475                               480
Gly Met Lys Tyr Leu Glu Glu Ser Asn Phe Val His Arg Asp Leu Ala
                               485                               490                               495
Ala Arg Asn Val Leu Leu Val Thr Gln His Tyr Ala Lys Ile Ser Asp
                               500                               505                               510
Phe Gly Leu Ser Lys Ala Leu Arg Ala Asp Glu Asn Tyr Tyr Lys Ala
                               515                               520                               525
Gln Thr His Gly Lys Trp Pro Val Lys Trp Tyr Ala Pro Glu Cys Ile
                               530                               535                               540
Asn Tyr Tyr Lys Phe Ser Ser Lys Ser Asp Val Trp Ser Phe Gly Val
545                               550                               555                               560
Leu Met Trp Glu Ala Phe Ser Tyr Gly Gln Lys Pro Tyr Arg Gly Met
                               565                               570                               575
Lys Gly Ser Glu Val Thr Ala Met Leu Glu Lys Gly Glu Arg Met Gly
                               580                               585                               590
Cys Pro Ala Gly Cys Pro Arg Glu Met Tyr Asp Leu Met Asn Leu Cys
                               595                               600                               605
Trp Thr Tyr Asp Val Glu Asn Arg Pro Gly Phe Ala Ala Val Glu Leu
610                               615                               620
Arg Leu Arg Asn Tyr Tyr Tyr Asp Val Val Asn
625                               630                               635

```

<210> 28

<211> 631

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 28

```

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

```

```

Asn Glu Gly Tyr Ile Pro Ser Asn Tyr Val Thr Gly Lys Lys Ser Asn
225                230                235                240
Asn Leu Asp Gln Tyr Glu Trp Tyr Cys Arg Asn Met Asn Arg Ser Lys
                245                250                255
Ala Glu Gln Leu Leu Arg Ser Glu Asp Lys Glu Gly Gly Phe Met Val
                260                265                270
Arg Asp Ser Ser Gln Pro Gly Leu Tyr Thr Val Ser Leu Tyr Thr Lys
                275                280                285
Phe Gly Gly Glu Gly Ser Ser Gly Phe Arg His Tyr His Ile Lys Glu
290                295                300
Thr Thr Thr Ser Pro Lys Lys Tyr Tyr Leu Ala Glu Lys His Ala Phe
305                310                315                320
Gly Ser Ile Pro Glu Ile Ile Glu Tyr His Lys His Asn Ala Ala Gly
                325                330                335
Leu Val Thr Arg Leu Arg Tyr Pro Val Ser Val Lys Gly Lys Asn Ala
                340                345                350
Pro Thr Thr Ala Gly Phe Ser Tyr Glu Lys Trp Glu Ile Asn Pro Ser
                355                360                365
Glu Leu Thr Phe Met Arg Glu Leu Gly Ser Gly Leu Phe Gly Val Val
370                375                380
Arg Leu Gly Lys Trp Arg Ala Gln Tyr Lys Val Ala Ile Lys Ala Ile
385                390                395                400
Arg Glu Gly Ala Met Cys Glu Glu Asp Phe Ile Glu Glu Ala Lys Val
                405                410                415
Met Met Lys Leu Thr His Pro Lys Leu Val Gln Leu Tyr Gly Val Cys
                420                425                430
Thr Gln Gln Lys Pro Ile Tyr Ile Val Thr Glu Phe Met Glu Arg Gly
                435                440                445
Cys Leu Leu Asn Phe Leu Arg Gln Arg Gln Gly His Phe Ser Arg Asp
450                455                460
Val Leu Leu Ser Met Cys Gln Asp Val Cys Glu Gly Met Glu Tyr Leu
465                470                475                480
Glu Arg Asn Ser Phe Ile His Arg Asp Leu Ala Ala Arg Asn Cys Leu
                485                490                495
Val Ser Glu Ala Gly Val Val Lys Val Ser Asp Phe Gly Met Ala Arg
500                505                510
Tyr Phe Leu Asp Asp Gln Tyr Thr Ser Ser Ser Gly Ala Lys Phe Pro
515                520                525
Val Lys Trp Cys Pro Pro Glu Val Phe Asn Tyr Ser Arg Phe Ser Ser
530                535                540
Lys Ser Asp Val Trp Ser Phe Gly Val Leu Met Trp Glu Val Phe Thr
545                550                555                560
Glu Gly Arg Met Pro Phe Glu Lys Tyr Thr Asn Tyr Glu Val Val Thr
                565                570                575
Met Val Thr Arg Gly His Arg Leu Tyr Gln Pro Lys Leu Ala Ser Asn
580                585                590
Tyr Val Tyr Glu Val Met Leu Arg Cys Trp Gln Glu Lys Pro Glu Gly
595                600                605
Arg Pro Ser Phe Glu Asp Leu Leu Arg Thr Ile Asp Glu Leu Val Glu
610                615                620
Cys Glu Glu Thr Phe Gly Arg
625                630

```

<210> 29

<211> 543

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 29

```

Met Gly Cys Ile Lys Ser Lys Glu Asn Lys Ser Pro Ala Ile Lys Tyr
 1           5           10           15
Arg Pro Glu Asn Thr Pro Glu Pro Val Ser Thr Ser Val Ser His Tyr
          20           25           30
Gly Ala Glu Pro Thr Thr Val Ser Pro Cys Pro Ser Ser Ser Ala Lys
          35           40           45
Gly Thr Ala Val Asn Phe Ser Ser Leu Ser Met Thr Pro Phe Gly Gly
 50           55           60
Ser Ser Gly Val Thr Pro Phe Gly Gly Ala Ser Ser Ser Phe Ser Val
65           70           75           80
Val Pro Ser Ser Tyr Pro Ala Gly Leu Thr Gly Gly Val Thr Ile Phe
          85           90           95
Val Ala Leu Tyr Asp Tyr Glu Ala Arg Thr Thr Glu Asp Leu Ser Phe
          100          105          110
Lys Lys Gly Glu Arg Phe Gln Ile Ile Asn Asn Thr Glu Gly Asp Trp
          115          120          125
Trp Glu Ala Arg Ser Ile Ala Thr Gly Lys Asn Gly Tyr Ile Pro Ser
          130          135          140
Asn Tyr Val Ala Pro Ala Asp Ser Ile Gln Ala Glu Glu Trp Tyr Phe
145          150          155          160
Gly Lys Met Gly Arg Lys Asp Ala Glu Arg Leu Leu Leu Asn Pro Gly
          165          170          175
Asn Gln Arg Gly Ile Phe Leu Val Arg Glu Ser Glu Thr Thr Lys Gly
          180          185          190
Ala Tyr Ser Leu Ser Ile Arg Asp Trp Asp Glu Ile Arg Gly Asp Asn
          195          200          205
Val Lys His Tyr Lys Ile Arg Lys Leu Asp Asn Gly Gly Tyr Tyr Ile
          210          215          220
Thr Thr Arg Ala Gln Phe Asp Thr Leu Gln Lys Leu Val Lys His Tyr
225          230          235          240
Thr Glu His Ala Asp Gly Leu Cys His Lys Leu Thr Thr Val Cys Pro
          245          250          255
Thr Val Lys Pro Gln Thr Gln Gly Leu Ala Lys Asp Ala Trp Glu Ile
          260          265          270
Pro Arg Glu Ser Leu Arg Leu Glu Val Lys Leu Gly Gln Gly Cys Phe
          275          280          285
Gly Glu Val Trp Met Gly Thr Trp Asn Gly Thr Thr Lys Val Ala Ile
          290          295          300
Lys Thr Leu Lys Pro Gly Thr Met Met Pro Glu Ala Phe Leu Gln Glu
305          310          315          320
Ala Gln Ile Met Lys Lys Leu Arg His Asp Lys Leu Val Pro Leu Tyr
          325          330          335
Ala Val Val Ser Glu Glu Pro Ile Tyr Ile Val Thr Glu Phe Met Ser
          340          345          350
Lys Gly Ser Leu Leu Asp Phe Leu Lys Glu Gly Asp Gly Lys Tyr Leu
          355          360          365
Lys Leu Pro Gln Leu Val Asp Met Ala Ala Gln Ile Ala Asp Gly Met
          370          375          380
Ala Tyr Ile Glu Arg Met Asn Tyr Ile His Arg Asp Leu Arg Ala Ala
385          390          395          400
Asn Ile Leu Val Gly Glu Asn Leu Val Cys Lys Ile Ala Asp Phe Gly
          405          410          415
Leu Ala Arg Leu Ile Glu Asp Asn Glu Tyr Thr Ala Arg Gln Gly Ala
          420          425          430
Lys Phe Pro Ile Lys Trp Thr Ala Pro Glu Ala Ala Leu Tyr Gly Arg
          435          440          445
Phe Thr Ile Lys Ser Asp Val Trp Ser Phe Gly Ile Leu Gln Thr Glu
          450          455          460

```

```

Leu Val Thr Lys Gly Arg Val Pro Tyr Pro Gly Met Val Asn Arg Glu
465                      470                      475                      480
Val Leu Glu Gln Val Glu Arg Gly Tyr Arg Met Pro Cys Pro Gln Gly
                      485                      490                      495
Cys Pro Glu Ser Leu His Glu Leu Met Asn Leu Cys Trp Lys Lys Asp
                      500                      505                      510
Pro Asp Glu Arg Pro Thr Phe Glu Tyr Ile Gln Ser Phe Leu Glu Asp
                      515                      520                      525
Tyr Phe Thr Ala Thr Glu Pro Gln Tyr Gln Pro Gly Glu Asn Leu
                      530                      535                      540

```

<210> 30

<211> 680

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 30

```

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

```

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

<210> 31

<211> 505

<212> PRT

<213> Artificial Sequence

$\langle 220 \rangle$

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 31

Met Ser Asn Ile Cys Gln Arg Leu Trp Glu Tyr Leu Glu Pro Tyr Leu
1 5 10 15
Pro Cys Leu Ser Thr Glu Ala Asp Lys Ser Thr Val Ile Glu Asn Pro
20 25 30

Gly	Ala	Leu	Cys	Ser	Pro	Gln	Ser	Gln	Arg	His	Gly	His	Tyr	Phe	Val
	35					40					45				
Ala	Leu	Phe	Asp	Tyr	Gln	Ala	Arg	Thr	Ala	Glu	Asp	Leu	Ser	Phe	Arg
50					55					60					
Ala	Gly	Asp	Lys	Leu	Gln	Val	Leu	Asp	Thr	Leu	His	Glu	Gly	Trp	Trp
65					70					75				80	
Phe	Ala	Arg	His	Leu	Glu	Lys	Arg	Arg	Asp	Gly	Ser	Ser	Gln	Gln	Leu
			85					90					95		
Gln	Gly	Tyr	Ile	Pro	Ser	Asn	Tyr	Val	Ala	Glu	Asp	Arg	Ser	Leu	Gln
		100					105					110			
Ala	Glu	Pro	Trp	Phe	Phe	Gly	Ala	Ile	Gly	Arg	Ser	Asp	Ala	Glu	Lys
	115					120					125				
Gln	Leu	Leu	Tyr	Ser	Glu	Asn	Lys	Thr	Gly	Ser	Phe	Leu	Ile	Arg	Glu
	130					135				140					
Ser	Glu	Ser	Gln	Lys	Gly	Glu	Phe	Ser	Leu	Ser	Val	Leu	Asp	Gly	Ala
145				150					155					160	
Val	Val	Lys	His	Tyr	Arg	Ile	Lys	Arg	Leu	Asp	Glu	Gly	Gly	Phe	Phe
			165					170						175	
Leu	Thr	Arg	Arg	Arg	Ile	Phe	Ser	Thr	Leu	Asn	Glu	Phe	Val	Ser	His
		180						185					190		
Tyr	Thr	Lys	Thr	Ser	Asp	Gly	Leu	Cys	Val	Lys	Leu	Gly	Lys	Pro	Cys
	195					200					205				
Leu	Lys	Ile	Gln	Val	Pro	Ala	Pro	Phe	Asp	Leu	Ser	Tyr	Lys	Thr	Val
	210					215				220					
Asp	Gln	Trp	Glu	Ile	Asp	Arg	Asn	Ser	Ile	Gln	Leu	Leu	Lys	Arg	Leu
225					230					235				240	
Gly	Ser	Gly	Gln	Phe	Gly	Glu	Val	Trp	Glu	Gly	Leu	Trp	Asn	Asn	Thr
			245						250				255		
Thr	Pro	Val	Ala	Val	Lys	Thr	Leu	Lys	Pro	Gly	Ser	Met	Asp	Pro	Asn
		260						265					270		
Asp	Phe	Leu	Arg	Glu	Ala	Gln	Ile	Met	Lys	Asn	Leu	Arg	His	Pro	Lys
	275						280					285			
Leu	Ile	Gln	Leu	Tyr	Ala	Val	Cys	Thr	Leu	Glu	Asp	Pro	Ile	Tyr	Ile
	290					295				300					
Ile	Thr	Glu	Leu	Met	Arg	His	Gly	Ser	Leu	Gln	Glu	Tyr	Leu	Gln	Asn
305				310						315				320	
Asp	Thr	Gly	Ser	Lys	Ile	His	Leu	Thr	Gln	Gln	Val	Asp	Met	Ala	Ala
			325						330					335	
Gln	Val	Ala	Ser	Gly	Met	Ala	Tyr	Leu	Glu	Ser	Arg	Asn	Tyr	Ile	His
		340						345					350		
Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu	Val	Gly	Glu	His	Asn	Ile	Tyr
	355					360						365			
Lys	Val	Ala	Asp	Phe	Gly	Leu	Ala	Arg	Val	Phe	Lys	Val	Asp	Asn	Glu
	370					375				380					
Asp	Ile	Tyr	Glu	Ser	Arg	His	Glu	Ile	Lys	Leu	Pro	Val	Lys	Trp	Thr
385				390						395				400	
Ala	Pro	Glu	Ala	Ile	Arg	Ser	Asn	Lys	Phe	Ser	Ile	Lys	Ser	Asp	Val
			405						410					415	
Trp	Ser	Phe	Gly	Ile	Leu	Leu	Tyr	Glu	Ile	Ile	Thr	Tyr	Gly	Lys	Met
	420							425					430		
Pro	Tyr	Ser	Gly	Met	Thr	Gly	Ala	Gln	Val	Ile	Gln	Met	Leu	Ala	Gln
	435						440					445			
Asn	Tyr	Arg	Leu	Pro	Gln	Pro	Ser	Asn	Cys	Pro	Gln	Gln	Phe	Tyr	Asn
	450					455					460				
Ile	Met	Leu	Glu	Cys	Trp	Asn	Ala	Glu	Pro	Lys	Glu	Arg	Pro	Thr	Phe
465				470						475				480	
Glu	Thr	Leu	Arg	Trp	Lys	Leu	Glu	Asp	Tyr	Phe	Glu	Thr	Asp	Ser	Ser
			485						490					495	
Tyr	Ser	Asp	Ala	Asn	Asn	Phe	Ile	Arg							
		500						505							

<210> 32
 <211> 1743
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 32

```

Met Ala Thr Asp Ser Gly Asp Pro Ala Ser Thr Glu Asp Ser Glu Lys
 1          5          10          15
Pro Asp Gly Ile Ser Phe Glu Asn Arg Val Pro Gln Val Ala Thr
 20          25          30
Leu Thr Val Glu Ala Arg Leu Lys Glu Lys Asn Ser Thr Phe Ser Ala
 35          40          45
Ser Gly Glu Thr Val Glu Arg Lys Arg Phe Phe Arg Lys Ser Val Glu
 50          55          60
Met Thr Glu Asp Asp Lys Val Ala Glu Ser Ser Pro Lys Asp Glu Arg
 65          70          75          80
Ile Lys Ala Ala Met Asn Ile Pro Arg Val Asp Lys Leu Pro Ser Asn
 85          90          95
Val Leu Arg Gly Gly Gln Glu Val Lys Tyr Glu Gln Cys Ser Lys Ser
100          105          110
Thr Ser Glu Ile Ser Lys Asp Cys Phe Lys Glu Lys Asn Glu Lys Glu
115          120          125
Met Glu Glu Glu Ala Glu Met Lys Ala Val Ala Thr Ser Pro Ser Gly
130          135          140
Arg Phe Leu Lys Phe Asp Ile Glu Leu Gly Arg Gly Ala Phe Lys Thr
145          150          155          160
Val Tyr Lys Gly Leu Asp Thr Glu Thr Trp Val Glu Val Ala Trp Cys
165          170          175
Glu Leu Gln Asp Arg Lys Leu Thr Lys Ala Glu Gln Gln Arg Phe Lys
180          185          190
Glu Glu Ala Glu Met Leu Lys Gly Leu Gln His Pro Asn Ile Val Arg
195          200          205
Phe Tyr Asp Ser Trp Glu Ser Ile Leu Lys Gly Lys Lys Cys Ile Val
210          215          220
Leu Val Thr Glu Leu Met Thr Ser Gly Thr Leu Lys Thr Tyr Leu Lys
225          230          235          240
Arg Phe Lys Val Met Lys Pro Lys Val Leu Arg Ser Trp Cys Arg Gln
245          250          255
Ile Leu Lys Gly Leu Gln Phe Leu His Thr Arg Thr Pro Pro Ile Ile
260          265          270
His Arg Asp Leu Lys Cys Asp Asn Ile Phe Ile Thr Gly Pro Thr Gly
275          280          285
Ser Val Lys Ile Gly Asp Leu Gly Leu Ala Thr Leu Met Arg Thr Ser
290          295          300
Phe Ala Lys Ser Val Ile Gly Thr Pro Glu Phe Met Ala Pro Glu Met
305          310          315          320
Tyr Glu Glu His Tyr Asp Glu Ser Val Asp Val Tyr Ala Phe Gly Met
325          330          335
Cys Met Leu Glu Met Ala Thr Ser Glu Tyr Pro Tyr Ser Glu Cys Gln
340          345          350
Asn Ala Ala Gln Ile Tyr Arg Lys Val Thr Ser Gly Ile Lys Pro Ala
355          360          365
Ser Phe Asn Lys Val Thr Asp Pro Glu Val Lys Glu Ile Ile Glu Gly
370          375          380
Cys Ile Arg Gln Asn Lys Ser Glu Arg Leu Ser Ile Arg Asp Leu Leu
385          390          395          400

```

Asn His Ala Phe Phe Ala Glu Asp Thr Gly Leu Arg Val Glu Leu Ala
 405 410 415
 Glu Glu Asp Asp Cys Ser Asn Ser Ser Leu Ala Leu Arg Leu Trp Val
 420 425 430
 Glu Asp Pro Lys Lys Leu Lys Gly Lys His Lys Asp Asn Glu Ala Ile
 435 440 445
 Glu Phe Ser Phe Asn Leu Glu Thr Asp Thr Pro Glu Glu Val Ala Tyr
 450 455 460
 Glu Met Val Lys Ser Gly Phe Phe His Glu Ser Asp Ser Lys Ala Val
 465 470 475 480
 Ala Lys Ser Ile Arg Asp Arg Val Thr Pro Ile Lys Lys Thr Arg Glu
 485 490 495
 Lys Lys Pro Ala Gly Cys Leu Glu Glu Arg Arg Asp Ser Gln Cys Lys
 500 505 510
 Ser Met Gly Asn Val Phe Pro Gln Pro Gln Asn Thr Thr Leu Pro Leu
 515 520 525
 Ala Pro Ala Gln Gln Thr Gly Ala Glu Cys Glu Glu Thr Glu Val Asp
 530 535 540
 Gln His Val Arg Gln Gln Leu Leu Gln Arg Lys Pro Gln Gln His Cys
 545 550 555 560
 Ser Ser Val Thr Gly Asp Asn Leu Ser Glu Ala Gly Ala Ala Ser Val
 565 570 575
 Ile His Ser Asp Thr Ser Ser Gln Pro Ser Val Ala Tyr Ser Ser Asn
 580 585 590
 Gln Thr Met Gly Ser Gln Met Val Ser Asn Ile Pro Gln Ala Glu Val
 595 600 605
 Asn Val Pro Gly Gln Ile Tyr Ser Ser Gln Gln Leu Val Gly His Tyr
 610 615 620
 Gln Gln Val Ser Gly Leu Gln Lys His Ser Lys Leu Thr Gln Pro Gln
 625 630 635 640
 Ile Leu Pro Leu Val Gln Gly Gln Ser Thr Val Leu Pro Val His Val
 645 650 655
 Leu Gly Pro Thr Val Val Ser Gln Pro Gln Val Ser Pro Leu Thr Val
 660 665 670
 Gln Lys Val Pro Gln Ile Lys Pro Val Ser Gln Pro Val Gly Ala Glu
 675 680 685
 Gln Gln Ala Ala Leu Leu Lys Pro Asp Leu Val Arg Ser Leu Asn Gln
 690 695 700
 Asp Val Ala Thr Thr Lys Glu Asn Val Ser Ser Pro Asp Asn Pro Ser
 705 710 715 720
 Gly Asn Gly Lys Gln Asp Arg Ile Lys Gln Arg Arg Ala Ser Cys Pro
 725 730 735
 Arg Pro Glu Lys Gly Thr Lys Phe Gln Leu Thr Val Leu Gln Val Ser
 740 745 750
 Thr Ser Gly Asp Asn Met Val Glu Cys Gln Leu Glu Thr His Asn Asn
 755 760 765
 Lys Met Val Thr Phe Lys Phe Asp Val Asp Gly Asp Ala Pro Glu Asp
 770 775 780
 Ile Ala Asp Tyr Met Val Glu Asp Asn Phe Val Leu Glu Ser Glu Lys
 785 790 795 800
 Glu Lys Phe Val Glu Leu Arg Ala Ile Val Gly Gln Ala Gln Glu
 805 810 815
 Ile Leu His Val His Phe Ala Thr Glu Arg Ala Thr Gly Val Asp Ser
 820 825 830
 Ile Thr Val Asp Ser Asn Ser Ser Gln Thr Gly Ser Ser Glu Gln Val
 835 840 845
 Gln Ile Asn Ser Thr Ser Thr Gln Thr Ser Asn Glu Ser Ala Pro Gln
 850 855 860
 Ser Ser Pro Val Gly Arg Trp Arg Phe Cys Ile Asn Gln Thr Ile Arg
 865 870 875 880

```

Asn Arg Glu Thr Gln Ser Pro Pro Ser Leu Gln His Ser Met Ser Ala
      885                      890                      895
Val Pro Gly Arg His Pro Leu Pro Ser Pro Lys Asn Thr Ser Asn Lys
      900                      905                      910
Glu Ile Ser Arg Asp Thr Leu Leu Thr Ile Glu Asn Asn Pro Cys His
      915                      920                      925
Arg Ala Leu Phe Thr Ser Lys Ser Glu His Lys Asp Val Val Asp Gly
      930                      935                      940
Lys Ile Ser Glu Cys Ala Ser Val Glu Thr Lys Gln Pro Ala Ile Leu
945      950                      955                      960
Tyr Gln Val Glu Asp Asn Arg Gln Ile Met Ala Pro Val Thr Asn Ser
      965                      970                      975
Ser Ser Tyr Ser Thr Thr Ser Val Arg Ala Val Pro Ala Glu Cys Glu
      980                      985                      990
Gly Leu Thr Lys Gln Ala Ser Ile Phe Ile Pro Val Tyr Pro Cys His
      995                      1000                      1005
Gln Thr Ala Ser Gln Ala Asp Ala Leu Met Ser His Pro Gly Glu Ser
      1010                      1015                      1020
Thr Gln Thr Ser Gly Asn Ser Leu Thr Thr Leu Ala Phe Asp Gln Lys
1025      1030                      1035                      1040
Pro Gln Thr Leu Ser Val Gln Gln Pro Ala Met Asp Ala Glu Phe Ile
      1045                      1050                      1055
Ser Gln Glu Gly Glu Thr Thr Val Asn Thr Glu Ala Ser Ser Pro Lys
      1060                      1065                      1070
Thr Val Ile Pro Thr Gln Thr Pro Gly Leu Glu Pro Thr Thr Leu Gln
      1075                      1080                      1085
Pro Thr Thr Val Leu Glu Ser Asp Gly Glu Arg Pro Pro Lys Leu Glu
      1090                      1095                      1100
Phe Ala Asp Asn Arg Ile Lys Thr Leu Asp Glu Lys Leu Arg Asn Leu
1105      1110                      1115                      1120
Leu Tyr Gln Glu His Ser Ile Ser Ser Ile Tyr Pro Glu Ser Gln Lys
      1125                      1130                      1135
Asp Thr Gln Ser Ile Asp Ser Pro Phe Ser Ser Ser Ala Glu Asp Thr
      1140                      1145                      1150
Leu Ser Cys Pro Val Thr Glu Val Ile Ala Ile Ser His Cys Gly Ile
      1155                      1160                      1165
Lys Asp Ser Pro Val Gln Ser Pro Asn Phe Gln Gln Thr Gly Ser Lys
      1170                      1175                      1180
Leu Leu Ser Asn Val Ala Ala Ser Gln Pro Ala Asn Ile Ser Val Phe
1185      1190                      1195                      1200
Lys Arg Asp Leu Asn Val Ile Thr Ser Val Pro Ser Glu Leu Cys Leu
      1205                      1210                      1215
His Glu Met Ser Ser Asp Ala Ser Leu Pro Gly Asp Pro Glu Ala Tyr
      1220                      1225                      1230
Pro Ala Ala Val Ser Ser Gly Gly Ala Ile His Leu Gln Thr Gly Val
      1235                      1240                      1245
Glu Thr Glu Glu Met Arg Ser Ala Ile Ala Pro Asp Pro Ile Pro Leu
      1250                      1255                      1260
Thr Arg Glu Ser Thr Ala Asp Thr Arg Ala Leu Asn Arg Cys Lys Ala
1265      1270                      1275                      1280
Met Ser Gly Ser Phe Gln Arg Gly Arg Phe Gln Val Ile Thr Ile Pro
      1285                      1290                      1295
Gln Gln Gln Ser Ala Lys Met Thr Ser Phe Gly Ile Glu His Ile Ser
      1300                      1305                      1310
Val Phe Ser Glu Thr Asn His Ser Ser Glu Glu Ala Phe Ile Lys Thr
      1315                      1320                      1325
Ala Lys Ser Gln Leu Val Glu Ile Glu Pro Ala Thr Gln Asn Pro Lys
      1330                      1335                      1340
Thr Ser Phe Ser Tyr Glu Lys Leu Gln Ala Leu Gln Glu Thr Cys Lys
1345      1350                      1355                      1360

```

Glu	Asn	Lys	Gly	Val	Pro	Lys	Gln	Gly	Asp	Asn	Phe	Leu	Ser	Phe	Ser
				1365							1370			1375	
Ala	Ala	Cys	Glu	Thr	Asp	Val	Ser	Ser	Val	Thr	Pro	Glu	Lys	Glu	Phe
				1380							1385			1390	
Glu	Glu	Thr	Ser	Ala	Thr	Gly	Ser	Ser	Met	Gln	Ser	Gly	Ser	Glu	Leu
				1395							1400			1405	
Leu	Leu	Lys	Glu	Arg	Glu	Ile	Leu	Thr	Ala	Gly	Lys	Gln	Pro	Ser	Ser
				1410							1415			1420	
Asp	Ser	Glu	Phe	Ser	Ala	Ser	Leu	Ala	Gly	Ser	Gly	Lys	Ser	Val	Ala
				1425							1430			1435	
Lys	Thr	Gly	Pro	Glu	Ser	Asn	Gln	Cys	Leu	Pro	His	His	Glu	Glu	Gln
				1445							1450			1455	
Ala	Tyr	Ala	Gln	Thr	Gln	Ser	Ser	Leu	Phe	Tyr	Ser	Pro	Ser	Ser	Pro
				1460							1465			1470	
Met	Ser	Ser	Asp	Asp	Glu	Ser	Glu	Ile	Glu	Asp	Glu	Asp	Leu	Lys	Val
				1475							1480			1485	
Glu	Leu	Gln	Arg	Leu	Arg	Glu	Lys	His	Ile	Gln	Glu	Val	Val	Asn	Leu
				1490							1495			1500	
Gln	Thr	Gln	Gln	Asn	Lys	Glu	Leu	Gln	Glu	Leu	Tyr	Glu	Arg	Leu	Arg
				1505							1510			1515	
Ser	Ile	Lys	Asp	Ser	Lys	Thr	Gln	Ser	Thr	Glu	Ile	Pro	Leu	Pro	Pro
				1525							1530			1535	
Ala	Ser	Pro	Arg	Arg	Pro	Arg	Ser	Phe	Lys	Ser	Lys	Leu	Arg	Ser	Arg
				1540							1545			1550	
Pro	Gln	Ser	Leu	Thr	His	Val	Asp	Asn	Gly	Ile	Val	Ala	Thr	Asp	Pro
				1555							1560			1565	
Leu	Cys	Val	Glu	Ser	Asn	Ala	Ala	Ser	Cys	Gln	Gln	Ser	Pro	Ala	Ser
				1570							1575			1580	
Lys	Lys	Gly	Met	Phe	Thr	Asp	Asp	Leu	His	Lys	Leu	Val	Asp	Asp	Trp
				1585							1590			1595	
Thr	Lys	Glu	Ala	Val	Gly	Asn	Ser	Leu	Ile	Lys	Pro	Ser	Leu	Asn	Gln
				1605							1610			1615	
Leu	Lys	Gln	Ser	Gln	His	Lys	Leu	Glu	Thr	Glu	Asn	Trp	Asn	Lys	Val
				1620							1625			1630	
Ser	Glu	Asn	Thr	Pro	Ser	Thr	Met	Gly	Tyr	Thr	Ser	Thr	Trp	Ile	Ser
				1635							1640			1645	
Ser	Leu	Ser	Gln	Ile	Arg	Gly	Ala	Val	Pro	Thr	Ser	Leu	Pro	Gln	Gly
				1650							1655			1660	
Leu	Ser	Leu	Pro	Ser	Phe	Pro	Gly	Pro	Leu	Ser	Ser	Tyr	Gly	Met	Pro
				1665							1670			1675	
His	Val	Cys	Gln	Tyr	Asn	Ala	Val	Ala	Gly	Ala	Gly	Tyr	Pro	Val	Gln
				1685							1690			1695	
Trp	Val	Gly	Ile	Ser	Gly	Thr	Thr	Gln	Gln	Ser	Val	Val	Ile	Pro	Ala
				1700							1705			1710	
Gln	Ser	Gly	Pro	Phe	Gln	Pro	Gly	Met	Asn	Met	Gln	Ala	Phe	Pro	
				1715							1720			1725	
Thr	Ser	Ser	Val	Gln	Asn	Pro	Ala	Thr	Ile	Pro	Pro	Gly	Pro	Lys	
				1730							1735			1740	

<210> 33

<211> 252

<212> PRT

<213> Artificial Sequence

 $\langle 220 \rangle$

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 33

Ile Thr Met Lys His Lys Leu Gly Gly Gly Gln Tyr Gly Glu Val Tyr
1 5 10 15

```

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

```

<210> 34

<211> 260

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 34

```

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

```

```

Met Gln Glu His Arg Lys Val Pro Phe Ala Trp Cys Ala Pro Glu Ser
      165      170      175
Leu Lys Thr Arg Thr Phe Ser His Ala Ser Asp Thr Trp Met Phe Gly
      180      185      190
Val Thr Leu Trp Glu Met Phe Thr Tyr Gly Gln Glu Pro Trp Ile Gly
      195      200      205
Leu Asn Gly Ser Gln Ile Leu His Lys Ile Asp Lys Glu Gly Glu Arg
      210      215      220
Leu Pro Arg Pro Glu Asp Cys Pro Gln Asp Ile Tyr Asn Val Met Val
225      230      235      240
Gln Cys Trp Ala His Lys Pro Glu Asp Arg Pro Thr Phe Val Ala Leu
      245      250      255
Arg Asp Phe Leu
      260

```

```

<210> 35
<211> 259
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence:/note =
      synthetic construct

```

```

<400> 35
Leu Lys Phe Asp Ile Glu Leu Gly Arg Gly Ser Phe Lys Thr Val Tyr
1      5      10      15
Lys Gly Leu Asp Thr Glu Thr Trp Val Glu Val Ala Trp Cys Glu Leu
      20      25      30
Gln Asp Arg Lys Leu Thr Lys Leu Glu Arg Gln Arg Phe Lys Glu Glu
      35      40      45
Ala Glu Met Leu Lys Gly Leu Gln His Pro Asn Ile Val Arg Phe Tyr
      50      55      60
Asp Phe Trp Glu Ser Ser Ala Lys Gly Lys Arg Cys Ile Val Leu Val
65      70      75      80
Thr Glu Leu Met Thr Ser Gly Thr Leu Lys Thr Tyr Leu Lys Arg Phe
      85      90      95
Lys Val Met Lys Pro Lys Val Leu Arg Ser Trp Cys Arg Gln Ile Leu
      100      105      110
Lys Gly Leu Leu Phe Leu His Thr Arg Thr Pro Pro Ile Ile His Arg
      115      120      125
Asp Leu Lys Cys Asp Asn Ile Phe Ile Thr Gly Pro Thr Gly Ser Val
      130      135      140
Lys Ile Gly Asp Leu Gly Leu Ala Thr Leu Lys Arg Ala Ser Phe Ala
145      150      155      160
Lys Ser Val Ile Gly Thr Pro Glu Phe Met Ala Pro Glu Met Tyr Glu
      165      170      175
Glu His Tyr Asp Glu Ser Val Asp Val Tyr Ala Phe Gly Met Cys Met
      180      185      190
Leu Glu Met Ala Thr Ser Glu Tyr Pro Tyr Ser Glu Cys Gln Asn Ala
      195      200      205
Ala Gln Ile Tyr Arg Lys Val Thr Cys Gly Ile Lys Pro Ala Ser Phe
      210      215      220
Glu Lys Val His Asp Pro Glu Ile Lys Glu Ile Ile Gly Glu Cys Ile
225      230      235      240
Cys Lys Asn Lys Glu Glu Arg Tyr Glu Ile Lys Asp Leu Leu Ser His
      245      250      255
Ala Phe Phe

```

```

<210> 36

```

<211> 251

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 36

```

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

```

<210> 37

<211> 329

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 37

```

Leu Gln Phe Gly Lys Thr Leu Gly Ala Gly Ala Phe Gly Lys Val Val
 1           5           10           15
Glu Ala Thr Ala Phe Gly Leu Gly Lys Glu Asp Ala Val Leu Lys Val
      20           25           30
Ala Val Lys Met Leu Lys Ser Thr Ala His Ala Asp Glu Lys Glu Ala
      35           40           45
Leu Met Ser Glu Leu Lys Ile Met Ser His Leu Gly Gln His Glu Asn
50           55           60

```

```

Ile Val Asn Leu Leu Gly Ala Cys Thr His Gly Gly Pro Val Leu Val
65          70          75          80
Ile Thr Glu Tyr Cys Cys Tyr Gly Asp Leu Leu Asn Phe Leu Arg Arg
      85          90          95
Lys Ala Glu Ala Met Leu Gly Pro Ser Leu Ser Pro Gly Gln Asp Pro
      100        105        110
Glu Gly Gly Val Asp Tyr Lys Asn Ile His Leu Glu Lys Lys Tyr Val
      115        120        125
Arg Arg Asp Ser Gly Phe Ser Ser Gln Gly Val Asp Thr Tyr Val Glu
      130        135        140
Met Arg Pro Val Ser Thr Ser Ser Asn Asp Ser Phe Ser Glu Gln Asp
145          150        155        160
Leu Asp Lys Glu Asp Gly Arg Pro Leu Glu Leu Arg Asp Leu Leu His
      165        170        175
Phe Ser Ser Gln Val Ala Gln Gly Met Ala Phe Leu Ala Ser Lys Asn
      180        185        190
Cys Ile His Arg Asp Val Ala Ala Arg Asn Val Leu Leu Thr Asn Gly
      195        200        205
His Val Ala Lys Ile Gly Asp Phe Gly Leu Ala Arg Asp Ile Met Asn
      210        215        220
Asp Ser Asn Tyr Ile Val Lys Gly Asn Ala Arg Leu Pro Val Lys Trp
225          230        235        240
Met Ala Pro Glu Ser Ile Phe Asp Cys Val Tyr Thr Val Gln Ser Asp
      245        250        255
Val Trp Ser Tyr Gly Ile Leu Leu Trp Glu Ile Phe Ser Leu Gly Leu
      260        265        270
Asn Pro Tyr Pro Gly Ile Leu Val Asn Ser Lys Phe Tyr Lys Leu Val
      275        280        285
Lys Asp Gly Tyr Gln Met Ala Gln Pro Ala Phe Ala Pro Lys Asn Ile
      290        295        300
Tyr Ser Ile Met Gln Ala Cys Trp Ala Leu Glu Pro Thr His Arg Pro
305          310        315        320
Thr Phe Gln Gln Ile Cys Ser Phe Leu
      325

```

<210> 38

<211> 246

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 38

```

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

```

```

Gly Asn Asn Phe Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val
    115          120          125
Ser Glu Asp Asn Val Ala Lys Val Ser Asp Phe Gly Leu Thr Lys Glu
    130          135          140
Ala Ser Ser Thr Gln Asp Thr Gly Lys Leu Pro Val Lys Trp Thr Ala
145          150          155          160
Pro Glu Ala Leu Arg Glu Lys Lys Phe Ser Thr Lys Ser Asp Val Trp
    165          170          175
Ser Phe Gly Ile Leu Leu Trp Glu Ile Tyr Ser Phe Gly Arg Val Pro
    180          185          190
Tyr Pro Arg Ile Pro Leu Lys Asp Val Val Pro Arg Val Glu Lys Gly
    195          200          205
Tyr Lys Met Asp Ala Pro Asp Gly Cys Pro Pro Ala Val Tyr Glu Val
    210          215          220
Met Lys Asn Cys Trp His Leu Asp Ala Ala Met Arg Pro Ser Phe Leu
225          230          235          240
Gln Leu Arg Glu Gln Leu
    245

```

<210> 39

<211> 255

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 39

```

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

```

<210> 40
 <211> 277
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 40

```

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

```

<210> 41
 <211> 277
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 41

```

Leu Val Leu Gly Lys Pro Leu Gly Glu Gly Cys Phe Gly Gln Val Val
 1          5          10          15

```

```

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

```

<210> 42

<211> 250

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 42

```

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

```

```

Gly Glu Arg Leu Ala Cys Lys Ile Ala Asp Phe Gly Leu Ala Arg Leu
  130                135                140
Ile Lys Asp Asp Glu Tyr Asn Pro Cys Gln Gly Ser Lys Phe Pro Ile
145                150                155                160
Lys Trp Thr Ala Pro Glu Ala Ala Leu Phe Gly Arg Phe Thr Ile Lys
                165                170                175
Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Thr Glu Leu Ile Thr Lys
                180                185                190
Gly Arg Ile Pro Tyr Pro Gly Met Asn Lys Arg Glu Val Leu Glu Gln
                195                200                205
Val Glu Gln Gly Tyr His Met Pro Cys Pro Pro Gly Cys Pro Ala Ser
                210                215                220
Leu Tyr Glu Ala Met Glu Gln Thr Trp Arg Leu Asp Pro Glu Glu Arg
225                230                235                240
Pro Thr Phe Glu Tyr Leu Gln Ser Phe Leu
                245                250

```

<210> 43

<211> 328

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 43

```

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

```

```

Pro Tyr Pro Gly Val Gln Met Asp Glu Asp Phe Cys Ser Arg Leu Arg
      275                      280                      285
Glu Gly Met Arg Met Arg Ala Pro Glu Tyr Ser Thr Pro Glu Ile Tyr
      290                      295                      300
Gln Ile Met Leu Asp Cys Trp His Arg Asp Pro Lys Glu Arg Pro Arg
305                      310                      315                      320
Phe Ala Glu Leu Val Glu Lys Leu
      325

```

<210> 44

<211> 325

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 44

```

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

```

<210> 45
 <211> 250
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 45

```

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

```

<210> 46
 <211> 251
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 46

```

Leu Lys Leu Glu Lys Lys Leu Gly Ala Gly Gln Phe Gly Glu Val Trp
 1           5           10           15
Met Ala Thr Tyr Asn Lys His Thr Lys Val Ala Val Lys Thr Met Lys
          20           25           30
Pro Gly Ser Met Ser Val Glu Ala Phe Leu Ala Glu Ala Asn Val Met
          35           40           45

```

```

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

```

<210> 47

<211> 258

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 47

```

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

```

```

Ile Trp Glu Leu Met Thr Phe Gly Gly Lys Pro Tyr Asp Gly Ile Pro
    195                200                205
Thr Arg Glu Ile Pro Asp Leu Leu Glu Lys Gly Glu Arg Leu Pro Gln
    210                215                220
Pro Pro Ile Cys Thr Ile Asp Val Tyr Met Val Met Val Lys Cys Trp
225                230                235                240
Met Ile Asp Ala Asp Ser Arg Pro Lys Phe Lys Glu Leu Ala Ala Glu
    245                250                255
Phe Ser

```

<210> 48

<211> 274

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 48

```

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

```

<210> 49

<211> 275

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 49

```

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

```

<210> 50

<211> 336

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 50

```

Leu Ser Phe Gly Lys Thr Leu Gly Ala Gly Ala Phe Gly Lys Val Val
 1          5          10          15
Glu Ala Thr Ala Tyr Gly Leu Ile Lys Ser Asp Ala Ala Met Thr Val
          20          25          30
Ala Val Lys Met Leu Lys Pro Ser Ala His Leu Thr Glu Arg Glu Ala
          35          40          45

```

```

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

```

<210> 51

<211> 251

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 51

```

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

```

```

Glu Arg Asn Tyr Ile His Arg Asp Leu Arg Ala Ala Asn Ile Leu Val
    115                      120                      125
Ser Asp Thr Leu Ser Cys Lys Ile Ala Asp Phe Gly Leu Ala Arg Leu
    130                      135                      140
Ile Glu Asp Asn Glu Tyr Thr Ala Arg Glu Gly Ala Lys Phe Pro Ile
    145                      150                      155                      160
Lys Trp Thr Ala Pro Glu Ala Ile Asn Tyr Gly Thr Phe Thr Ile Lys
    165                      170                      175
Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Thr Glu Ile Val Thr His
    180                      185                      190
Gly Arg Ile Pro Tyr Pro Gly Met Thr Asn Pro Glu Val Ile Gln Asn
    195                      200                      205
Leu Glu Arg Gly Tyr Arg Met Val Arg Pro Asp Asn Cys Pro Glu Glu
    210                      215                      220
Leu Tyr Gln Leu Met Arg Leu Cys Trp Lys Glu Arg Pro Glu Asp Arg
    225                      230                      235                      240
Pro Thr Phe Asp Tyr Leu Arg Ser Val Leu Glu
    245                      250

```

<210> 52

<211> 252

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 52

```

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

```

<210> 53
 <211> 250
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 53

```

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

```

<210> 54
 <211> 358
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 54

```

Leu Val Leu Gly Arg Val Leu Gly Ser Gly Ala Phe Gly Lys Val Val
 1           5           10           15
Glu Gly Thr Ala Tyr Gly Leu Ser Arg Ser Gln Pro Val Met Lys Val
          20           25           30
Ala Val Lys Met Leu Lys Pro Thr Ala Arg Ser Ser Glu Lys Gln Ala
          35           40           45

```

```

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

```

```

<210> 55
<211> 359
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence:/note =
        synthetic construct

```

```

<400> 55
Leu Val Leu Gly Arg Thr Leu Gly Ser Gly Ala Phe Gly Gln Val Val
 1                      5                      10                      15
Glu Ala Thr Ala His Gly Leu Ser His Ser Gln Ala Thr Met Lys Val
                        20                      25                      30
Ala Val Lys Met Leu Lys Ser Thr Ala Arg Ser Ser Glu Lys Gln Ala
                        35                      40                      45
Leu Met Ser Glu Leu Lys Ile Met Ser His Leu Gly Pro His Leu Asn
50                      55                      60
Val Val Asn Leu Leu Gly Ala Cys Thr Lys Gly Gly Pro Ile Tyr Ile
65                      70                      75                      80

```

```

Ile Thr Glu Tyr Cys Arg Tyr Gly Asp Leu Val Asp Tyr Leu His Arg
      85                      90                      95
Asn Lys His Thr Phe Leu Gln His His Ser Asp Lys Arg Arg Pro Pro
      100                      105                      110
Ser Ala Glu Leu Tyr Ser Asn Ala Leu Pro Val Gly Leu Pro Leu Pro
      115                      120                      125
Ser His Val Ser Leu Thr Gly Glu Ser Asp Gly Gly Tyr Met Asp Met
      130                      135                      140
Ser Lys Asp Glu Ser Val Asp Tyr Val Pro Met Leu Asp Met Lys Gly
145                      150                      155                      160
Asp Val Lys Tyr Ala Asp Ile Glu Ser Ser Asn Tyr Met Ala Pro Tyr
      165                      170                      175
Asp Asn Tyr Val Pro Ser Ala Pro Glu Arg Thr Cys Arg Ala Thr Leu
      180                      185                      190
Ile Asn Glu Ser Pro Val Leu Ser Tyr Met Asp Leu Val Gly Phe Ser
      195                      200                      205
Tyr Gln Val Ala Asn Gly Met Glu Phe Leu Ala Ser Lys Asn Cys Val
      210                      215                      220
His Arg Asp Leu Ala Ala Arg Asn Val Leu Ile Cys Glu Gly Lys Leu
225                      230                      235                      240
Val Lys Ile Cys Asp Phe Gly Leu Ala Arg Asp Ile Met Arg Asp Ser
      245                      250                      255
Asn Tyr Ile Ser Lys Gly Ser Thr Phe Leu Pro Leu Lys Trp Met Ala
      260                      265                      270
Pro Glu Ser Ile Phe Asn Ser Leu Tyr Thr Thr Leu Ser Asp Val Trp
      275                      280                      285
Ser Phe Gly Ile Leu Leu Trp Glu Ile Phe Thr Leu Gly Gly Thr Pro
      290                      295                      300
Tyr Pro Glu Leu Pro Met Asn Glu Gln Phe Tyr Asn Ala Ile Lys Arg
305                      310                      315                      320
Gly Tyr Arg Met Ala Gln Pro Ala His Ala Ser Asp Glu Ile Tyr Glu
      325                      330                      335
Ile Met Gln Lys Cys Trp Glu Glu Lys Phe Glu Ile Arg Pro Pro Phe
      340                      345                      350
Ser Gln Leu Val Leu Leu Leu
      355

```

<210> 56

<211> 284

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 56

```

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

```

```

Ser Ser Leu Asp His Pro Asp Glu Arg Ala Leu Thr Met Gly Asp Leu
    115          120          125
Ile Ser Phe Ala Trp Gln Ile Ser Gln Gly Met Gln Tyr Leu Ala Glu
    130          135          140
Met Lys Leu Val His Arg Asp Leu Ala Ala Arg Asn Ile Leu Val Ala
    145          150          155          160
Glu Gly Arg Lys Met Lys Ile Ser Asp Phe Gly Leu Ser Arg Asp Val
    165          170          175
Tyr Glu Glu Asp Ser Tyr Val Lys Arg Ser Gln Gly Arg Ile Pro Val
    180          185          190
Lys Trp Met Ala Ile Glu Ser Leu Phe Asp His Ile Tyr Thr Thr Gln
    195          200          205
Ser Asp Val Trp Ser Phe Gly Val Leu Leu Trp Glu Ile Val Thr Leu
    210          215          220
Gly Gly Asn Pro Tyr Pro Gly Ile Pro Pro Glu Arg Leu Phe Asn Leu
    225          230          235          240
Leu Lys Thr Gly His Arg Met Glu Arg Pro Asp Asn Cys Ser Glu Glu
    245          250          255
Met Tyr Arg Leu Met Leu Gln Cys Trp Lys Gln Glu Pro Asp Lys Arg
    260          265          270
Pro Val Phe Ala Asp Ile Ser Lys Asp Leu Glu Lys
    275          280

```

<210> 57

<211> 267

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 57

```

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

```

```

Lys Gln Pro Phe Glu Asp Val Thr Asn Pro Leu Gln Ile Met Tyr Ser
  210                215                220
Val Ser Gln Gly His Arg Pro Val Ile Asn Glu Glu Ser Leu Pro Tyr
  225                230                235                240
Asp Ile Pro His Arg Ala Arg Met Ile Ser Leu Ile Glu Ser Gly Trp
                245                250                255
Ala Gln Asn Pro Asp Glu Arg Pro Ser Phe Leu
                260                265

```

```

<210> 58
<211> 249
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence:/note =
        synthetic construct

```

```

<400> 58
Leu Arg Leu Glu Val Lys Leu Gly Gln Gly Cys Phe Gly Glu Val Trp
  1                5                10                15
Met Gly Thr Trp Asn Gly Thr Thr Arg Val Ala Ile Lys Thr Leu Lys
                20                25                30
Pro Gly Thr Met Ser Pro Glu Ala Phe Leu Gln Glu Ala Gln Val Met
  35                40                45
Lys Lys Leu Arg His Glu Lys Leu Val Gln Leu Tyr Ala Val Val Ser
  50                55                60
Glu Glu Pro Ile Tyr Ile Val Thr Glu Tyr Met Ser Lys Gly Ser Leu
  65                70                75                80
Leu Asp Phe Leu Lys Gly Glu Thr Gly Lys Tyr Leu Arg Leu Pro Gln
                85                90                95
Leu Val Asp Met Ala Ala Gln Ile Ala Ser Gly Met Ala Tyr Val Glu
                100                105                110
Arg Met Asn Tyr Val His Arg Asp Leu Arg Ala Ala Asn Ile Leu Val
                115                120                125
Gly Glu Asn Leu Val Cys Lys Val Ala Asp Phe Gly Leu Ala Arg Leu
  130                135                140
Ile Glu Asp Asn Glu Tyr Thr Ala Arg Gln Gly Ala Lys Phe Pro Ile
  145                150                155                160
Lys Trp Thr Ala Pro Glu Ala Ala Leu Tyr Gly Arg Phe Thr Ile Lys
                165                170                175
Ser Asp Val Trp Ser Phe Gly Ile Leu Leu Thr Glu Leu Thr Thr Lys
                180                185                190
Gly Arg Val Pro Tyr Pro Gly Met Val Asn Arg Glu Val Leu Asp Gln
                195                200                205
Val Glu Arg Gly Tyr Arg Met Pro Cys Pro Pro Glu Cys Pro Glu Ser
  210                215                220
Leu His Asp Leu Met Cys Gln Cys Trp Arg Lys Glu Pro Glu Glu Arg
  225                230                235                240
Pro Thr Phe Glu Tyr Leu Gln Ala Phe
                245

```

```

<210> 59
<211> 249
<212> PRT
<213> Artificial Sequence

```

```

<220>
<223> Description of Artificial Sequence:/note =
        synthetic construct

```

<400> 59

```

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

```

<210> 60

<211> 250

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 60

```

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

```

```

Ser Glu Ala Gly Val Val Lys Val Ser Asp Phe Gly Met Ala Arg Tyr
  130                      135                      140
Phe Leu Asp Asp Gln Tyr Thr Ser Ser Ser Gly Ala Lys Phe Pro Val
 145                      150                      155                      160
Lys Trp Cys Pro Pro Glu Val Phe Asn Tyr Ser Arg Phe Ser Ser Lys
                      165                      170                      175
Ser Asp Val Trp Ser Phe Gly Val Leu Met Trp Glu Val Phe Thr Glu
                      180                      185                      190
Gly Arg Met Pro Phe Glu Lys Tyr Thr Asn Tyr Glu Val Val Thr Met
                      195                      200                      205
Val Thr Arg Gly His Arg Leu Tyr Gln Pro Lys Leu Ala Ser Asn Tyr
  210                      215                      220
Val Tyr Glu Val Met Leu Arg Cys Trp Gln Glu Lys Pro Glu Gly Arg
 225                      230                      235                      240
Pro Ser Phe Glu Asp Leu Leu Arg Thr Ile
                      245                      250

```

<210> 61

<211> 250

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 61

```

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

```

<210> 62

<211> 252
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 62

```

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

```

<210> 63
 <211> 258
 <212> PRT
 <213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
 synthetic construct

<400> 63

```

Ile Gln Leu Leu Lys Arg Leu Gly Ser Gly Gln Phe Gly Glu Val Trp
 1           5           10           15
Glu Gly Leu Trp Asn Asn Thr Thr Pro Val Ala Val Lys Thr Leu Lys
      20           25           30
Pro Gly Ser Met Asp Pro Asn Asp Phe Leu Arg Glu Ala Gln Ile Met
      35           40           45
Lys Asn Leu Arg His Pro Lys Leu Ile Gln Leu Tyr Ala Val Cys Thr
      50           55           60

```

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

<210> 64

<211> 259

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:/note =
synthetic construct

<400> 64

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

Leu	Glu	Met	Ala	Thr	Ser	Glu	Tyr	Pro	Tyr	Ser	Glu	Cys	Gln	Asn	Ala	
		195					200					205				
Ala	Gln	Ile	Tyr	Arg	Lys	Val	Thr	Ser	Gly	Ile	Lys	Pro	Ala	Ser	Phe	
	210					215					220					
Asn	Lys	Val	Thr	Asp	Pro	Glu	Val	Lys	Glu	Ile	Ile	Glu	Gly	Cys	Ile	
225					230					235					240	
Arg	Gln	Asn	Lys	Ser	Glu	Arg	Leu	Ser	Ile	Arg	Asp	Leu	Leu	Asn	His	
				245					250					255		
Ala	Phe	Phe														