(54) Title: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY AND DIAGNOSIS OF TUBERCULOSIS

(57) Abstract

Compounds and methods for inducing protective immunity against tuberculosis are disclosed. The compounds provided include polypeptides that contain at least one immunogenic portion of one or more \textit{M. tuberculosis} proteins and DNA molecules encoding such polypeptides. Such compounds may be formulated into vaccines and/or pharmaceutical compositions for immunization against \textit{M. tuberculosis} infection, or may be used for the diagnosis of tuberculosis.
FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<table>
<thead>
<tr>
<th>Code</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>AL</td>
<td>Albania</td>
</tr>
<tr>
<td>AM</td>
<td>Armenia</td>
</tr>
<tr>
<td>AT</td>
<td>Austria</td>
</tr>
<tr>
<td>AU</td>
<td>Australia</td>
</tr>
<tr>
<td>AZ</td>
<td>Azerbaijan</td>
</tr>
<tr>
<td>BA</td>
<td>Bosnia and Herzegovina</td>
</tr>
<tr>
<td>BB</td>
<td>Barbados</td>
</tr>
<tr>
<td>BE</td>
<td>Belgium</td>
</tr>
<tr>
<td>BF</td>
<td>Burkina Faso</td>
</tr>
<tr>
<td>BG</td>
<td>Bulgaria</td>
</tr>
<tr>
<td>BJ</td>
<td>Benin</td>
</tr>
<tr>
<td>BR</td>
<td>Brazil</td>
</tr>
<tr>
<td>BY</td>
<td>Belarus</td>
</tr>
<tr>
<td>CA</td>
<td>Canada</td>
</tr>
<tr>
<td>CF</td>
<td>Central African Republic</td>
</tr>
<tr>
<td>CG</td>
<td>Congo</td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
</tr>
<tr>
<td>CI</td>
<td>Côte d’Ivoire</td>
</tr>
<tr>
<td>CM</td>
<td>Cameroon</td>
</tr>
<tr>
<td>CN</td>
<td>China</td>
</tr>
<tr>
<td>CU</td>
<td>Cuba</td>
</tr>
<tr>
<td>CZ</td>
<td>Czech Republic</td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
</tr>
<tr>
<td>EE</td>
<td>Estonia</td>
</tr>
<tr>
<td>ES</td>
<td>Spain</td>
</tr>
<tr>
<td>FI</td>
<td>Finland</td>
</tr>
<tr>
<td>FR</td>
<td>France</td>
</tr>
<tr>
<td>GA</td>
<td>Gabon</td>
</tr>
<tr>
<td>GB</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>GE</td>
<td>Georgia</td>
</tr>
<tr>
<td>GH</td>
<td>Ghana</td>
</tr>
<tr>
<td>GN</td>
<td>Guinea</td>
</tr>
<tr>
<td>GR</td>
<td>Greece</td>
</tr>
<tr>
<td>HU</td>
<td>Hungary</td>
</tr>
<tr>
<td>IE</td>
<td>Ireland</td>
</tr>
<tr>
<td>IL</td>
<td>Israel</td>
</tr>
<tr>
<td>IS</td>
<td>Iceland</td>
</tr>
<tr>
<td>IT</td>
<td>Italy</td>
</tr>
<tr>
<td>JP</td>
<td>Japan</td>
</tr>
<tr>
<td>KE</td>
<td>Kenya</td>
</tr>
<tr>
<td>KG</td>
<td>Kyrgyzstan</td>
</tr>
<tr>
<td>KP</td>
<td>Democratic People’s</td>
</tr>
<tr>
<td>KR</td>
<td>Republic of Korea</td>
</tr>
<tr>
<td>LA</td>
<td>Latvia</td>
</tr>
<tr>
<td>MC</td>
<td>Monaco</td>
</tr>
<tr>
<td>MD</td>
<td>Republic of Moldova</td>
</tr>
<tr>
<td>MG</td>
<td>Madagascar</td>
</tr>
<tr>
<td>MK</td>
<td>The former Yugoslavia</td>
</tr>
<tr>
<td>ML</td>
<td>Mali</td>
</tr>
<tr>
<td>MN</td>
<td>Mongolia</td>
</tr>
<tr>
<td>MR</td>
<td>Mauritania</td>
</tr>
<tr>
<td>MW</td>
<td>Malawi</td>
</tr>
<tr>
<td>MX</td>
<td>Mexico</td>
</tr>
<tr>
<td>NE</td>
<td>Niger</td>
</tr>
<tr>
<td>NL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>NO</td>
<td>Norway</td>
</tr>
<tr>
<td>NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td>PL</td>
<td>Poland</td>
</tr>
<tr>
<td>PT</td>
<td>Portugal</td>
</tr>
<tr>
<td>RO</td>
<td>Romania</td>
</tr>
<tr>
<td>RU</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>SD</td>
<td>Sudan</td>
</tr>
<tr>
<td>SE</td>
<td>Sweden</td>
</tr>
<tr>
<td>SG</td>
<td>Singapore</td>
</tr>
<tr>
<td>SI</td>
<td>Slovenia</td>
</tr>
<tr>
<td>SK</td>
<td>Slovakia</td>
</tr>
<tr>
<td>SN</td>
<td>Senegal</td>
</tr>
<tr>
<td>SZ</td>
<td>Swaziland</td>
</tr>
<tr>
<td>TD</td>
<td>Chad</td>
</tr>
<tr>
<td>TG</td>
<td>Togo</td>
</tr>
<tr>
<td>TJ</td>
<td>Tajikistan</td>
</tr>
<tr>
<td>TM</td>
<td>Turkmenistan</td>
</tr>
<tr>
<td>TR</td>
<td>Turkey</td>
</tr>
<tr>
<td>TT</td>
<td>Trinidad and Tobago</td>
</tr>
<tr>
<td>UA</td>
<td>Ukraine</td>
</tr>
<tr>
<td>UG</td>
<td>Uganda</td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>UZ</td>
<td>Uzbekistan</td>
</tr>
<tr>
<td>VN</td>
<td>Vietnam</td>
</tr>
<tr>
<td>YU</td>
<td>Yugoslavia</td>
</tr>
<tr>
<td>ZW</td>
<td>Zimbabwe</td>
</tr>
</tbody>
</table>
CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Application No. 9/025,197, filed February 18, 1998; which is a continuation-in-part of U.S. Application No. 08/942,578, filed October 1, 1997; which is a continuation-in-part of U.S. Application No. 08/818,112, filed March 13, 1997; which is a continuation-in-part of U.S. Application No. 08/730,510, filed October 11, 1996; which claims priority from PCT Application No. PCT/US 96/14674, filed August 30, 1996; and is a continuation-in-part of U.S. Application No. 08/680,574, filed July 12, 1996; which is a continuation-in-part of U.S. Application No. 08/659,683, filed June 5, 1996; which is a continuation-in-part of U.S. Application No. 08/620,874, filed March 22, 1996, now abandoned; which is a continuation-in-part of U.S. Application No. 08/533,634, filed September 22, 1995, now abandoned; which is a continuation-in-part of U.S. Application No. 08/523,436, filed September 1, 1995, now abandoned.

TECHNICAL FIELD

The present invention relates generally to detecting, treating and preventing *Mycobacterium tuberculosis* infection. The invention is more particularly related to polypeptides comprising a *Mycobacterium tuberculosis* antigen, or a portion or other variant thereof, and the use of such polypeptides for diagnosing and vaccinating against *Mycobacterium tuberculosis* infection.

BACKGROUND OF THE INVENTION

Tuberculosis is a chronic, infectious disease, that is generally caused by infection with *Mycobacterium tuberculosis*. It is a major disease in developing countries, as well as an increasing problem in developed areas of the world, with about 8 million new cases and 3 million deaths each year. Although the infection may be asymptomatic for a considerable period of time, the disease is most commonly
manifested as an acute inflammation of the lungs, resulting in fever and a nonproductive cough. If left untreated, serious complications and death typically result.

Although tuberculosis can generally be controlled using extended antibiotic therapy, such treatment is not sufficient to prevent the spread of the disease. Infected individuals may be asymptomatic, but contagious, for some time. In addition, although compliance with the treatment regimen is critical, patient behavior is difficult to monitor. Some patients do not complete the course of treatment, which can lead to ineffective treatment and the development of drug resistance.

Inhibiting the spread of tuberculosis requires effective vaccination and accurate, early diagnosis of the disease. Currently, vaccination with live bacteria is the most efficient method for inducing protective immunity. The most common Mycobacterium employed for this purpose is Bacillus Calmette-Guerin (BCG), an avirulent strain of Mycobacterium bovis. However, the safety and efficacy of BCG is a source of controversy and some countries, such as the United States, do not vaccinate the general public. Diagnosis is commonly achieved using a skin test, which involves intradermal exposure to tuberculin PPD (protein-purified derivative). Antigen-specific T cell responses result in measurable induration at the injection site by 48-72 hours after injection, which indicates exposure to Mycobacterial antigens. Sensitivity and specificity have, however, been a problem with this test, and individuals vaccinated with BCG cannot be distinguished from infected individuals.

While macrophages have been shown to act as the principal effectors of M. tuberculosis immunity, T cells are the predominant inducers of such immunity. The essential role of T cells in protection against M. tuberculosis infection is illustrated by the frequent occurrence of M. tuberculosis in AIDS patients, due to the depletion of CD4 T cells associated with human immunodeficiency virus (HIV) infection. Mycobacterium-reactive CD4 T cells have been shown to be potent producers of gamma-interferon (IFN-γ), which, in turn, has been shown to trigger the antimycobacterial effects of macrophages in mice. While the role of IFN-γ in humans is less clear, studies have shown that 1,25-dihydroxy-vitamin D3, either alone or in combination with IFN-γ or tumor necrosis factor-alpha, activates human macrophages
to inhibit *M. tuberculosis* infection. Furthermore, it is known that IFN-γ stimulates human macrophages to make 1,25-dihydroxy-vitamin D3. Similarly, IL-12 has been shown to play a role in stimulating resistance to *M. tuberculosis* infection. For a review of the immunology of *M. tuberculosis* infection see Chan and Kaufmann in *Tuberculosis: Pathogenesis, Protection and Control*, Bloom (ed.), ASM Press, Washington, DC, 1994.

Accordingly, there is a need in the art for improved vaccines and methods for preventing, treating and detecting tuberculosis. The present invention fulfills these needs and further provides other related advantages.

**SUMMARY OF THE INVENTION**

Briefly stated, this invention provides compounds and methods for preventing and diagnosing tuberculosis. In one aspect, polypeptides are provided comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications. In one embodiment of this aspect, the soluble antigen has one of the following N-terminal sequences:

(a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)

(b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 121)

(c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)

(d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 123)

(e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 124)

(f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 125)

(g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID No. 126)
(h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 127)

(i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID No. 128)

(j) Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-Ser; (SEQ ID No. 134)

(k) Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-Asp; (SEQ ID No. 135) or

(l) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136)

wherein Xaa may be any amino acid.

In a related aspect, polypeptides are provided comprising an immunogenic portion of an M. tuberculosis antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications, the antigen having one of the following N-terminal sequences:

(m) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137) or

(n) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129)

wherein Xaa may be any amino acid.

In another embodiment, the soluble M. tuberculosis antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101 or a complement thereof under moderately stringent conditions.

In a related aspect, the polypeptides comprise an immunogenic portion of a M. tuberculosis antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications, wherein the antigen comprises an

In related aspects, DNA sequences encoding the above polypeptides, expression vectors comprising these DNA sequences and host cells transformed or transfected with such expression vectors are also provided.

In another aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, an inventive polypeptide and a known M. tuberculosis antigen.

Within other aspects, the present invention provides pharmaceutical compositions that comprise one or more of the above polypeptides, or a DNA molecule encoding such polypeptides, and a physiologically acceptable carrier. The invention also provides vaccines comprising one or more of the polypeptides as described above and a non-specific immune response enhancer, together with vaccines comprising one or more DNA sequences encoding such polypeptides and a non-specific immune response enhancer.

In yet another aspect, methods are provided for inducing protective immunity in a patient, comprising administering to a patient an effective amount of one or more of the above polypeptides.

In further aspects of this invention, methods and diagnostic kits are provided for detecting tuberculosis in a patient. The methods comprise contacting dermal cells of a patient with one or more of the above polypeptides and detecting an immune response on the patient's skin. The diagnostic kits comprise one or more of the above polypeptides in combination with an apparatus sufficient to contact the polypeptide with the dermal cells of a patient.

These and other aspects of the present invention will become apparent upon reference to the following detailed description and attached drawings. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

Figure 1A and B illustrate the stimulation of proliferation and interferon-γ production in T cells derived from a first and a second M. tuberculosis-immune donor, respectively, by the 14 Kd, 20 Kd and 26 Kd antigens described in Example 1.

Figure 2 illustrates the stimulation of proliferation and interferon-γ production in T cells derived from an M. tuberculosis-immune individual by the two representative polypeptides TbRa3 and TbRa9.

Figures 3A-D illustrate the reactivity of antisera raised against secretory M. tuberculosis proteins, the known M. tuberculosis antigen 85b and the inventive antigens Tb38-1 and TbH-9, respectively, with M. tuberculosis lysate (lane 2), M. tuberculosis secretory proteins (lane 3), recombinant Tb38-1 (lane 4), recombinant TbH-9 (lane 5) and recombinant 85b (lane 5).

Figure 4A illustrates the stimulation of proliferation in a TbH-9-specific T cell clone by secretory M. tuberculosis proteins, recombinant TbH-9 and a control antigen, TbRa11.
Figure 4B illustrates the stimulation of interferon-γ production in a TbH-9-specific T cell clone by secretory *M. tuberculosis* proteins, PPD and recombinant TbH-9.

Figures 5A and B illustrate the stimulation of proliferation and interferon-γ production in TbH9-specific T cells by the fusion protein TbH9-Tb38-1.

Figures 6A and B illustrate the stimulation of proliferation and interferon-γ production in Tb38-1-specific T cells by the fusion protein TbH9-Tb38-1.

Figures 7A and B illustrate the stimulation of proliferation and interferon-γ production in T cells previously shown to respond to both TbH-9 and Tb38-1 by the fusion protein TbH9-Tb38-1.

Figures 8A and B illustrate the stimulation of proliferation and interferon-γ production in T cells derived from a first *M. tuberculosis*-immune individual by the representative polypeptides XP-1, RDIF6, RDIF8, RDIF10 and RDIF11.

Figures 9A and B illustrate the stimulation of proliferation and interferon-γ production in T cells derived from a second *M. tuberculosis*-immune individual by the representative polypeptides XP-1, RDIF6, RDIF8, RDIF10 and RDIF11.

SEQ. ID NO. 1 is the DNA sequence of TbRa1.

SEQ. ID NO. 2 is the DNA sequence of TbRa10.

SEQ. ID NO. 3 is the DNA sequence of TbRa11.

SEQ. ID NO. 4 is the DNA sequence of TbRa12.

SEQ. ID NO. 5 is the DNA sequence of TbRa13.

SEQ. ID NO. 6 is the DNA sequence of TbRa16.

SEQ. ID NO. 7 is the DNA sequence of TbRa17.

SEQ. ID NO. 8 is the DNA sequence of TbRa18.

SEQ. ID NO. 9 is the DNA sequence of TbRa19.

SEQ. ID NO. 10 is the DNA sequence of TbRa24.

SEQ. ID NO. 11 is the DNA sequence of TbRa26.

SEQ. ID NO. 12 is the DNA sequence of TbRa28.
SEQ. ID NO. 13 is the DNA sequence of TbRa29.
SEQ. ID NO. 14 is the DNA sequence of TbRa2A.
SEQ. ID NO. 15 is the DNA sequence of TbRa3.
SEQ. ID NO. 16 is the DNA sequence of TbRa32.
SEQ. ID NO. 17 is the DNA sequence of TbRa35.
SEQ. ID NO. 18 is the DNA sequence of TbRa36.
SEQ. ID NO. 19 is the DNA sequence of TbRa4.
SEQ. ID NO. 20 is the DNA sequence of TbRa9.
SEQ. ID NO. 21 is the DNA sequence of TbRaB.
SEQ. ID NO. 22 is the DNA sequence of TbRaC.
SEQ. ID NO. 23 is the DNA sequence of TbRaD.
SEQ. ID NO. 24 is the DNA sequence of YYWCPG.
SEQ. ID NO. 25 is the DNA sequence of AAMK.
SEQ. ID NO. 26 is the DNA sequence of TbL-23.
SEQ. ID NO. 27 is the DNA sequence of TbL-24.
SEQ. ID NO. 28 is the DNA sequence of TbL-25.
SEQ. ID NO. 29 is the DNA sequence of TbL-28.
SEQ. ID NO. 30 is the DNA sequence of TbL-29.
SEQ. ID NO. 31 is the DNA sequence of TbH-5.
SEQ. ID NO. 32 is the DNA sequence of TbH-8.
SEQ. ID NO. 33 is the DNA sequence of TbH-9.
SEQ. ID NO. 34 is the DNA sequence of TbM-1.
SEQ. ID NO. 35 is the DNA sequence of TbM-3.
SEQ. ID NO. 36 is the DNA sequence of TbM-6.
SEQ. ID NO. 37 is the DNA sequence of TbM-7.
SEQ. ID NO. 38 is the DNA sequence of TbM-9.
SEQ. ID NO. 39 is the DNA sequence of TbM-12.
SEQ. ID NO. 40 is the DNA sequence of TbM-13.
SEQ. ID NO. 41 is the DNA sequence of TbM-14.
SEQ. ID NO. 42 is the DNA sequence of TbM-15.
SEQ. ID NO. 43 is the DNA sequence of TbH-4.
SEQ. ID NO. 44 is the DNA sequence of TbH-4-FWD.
SEQ. ID NO. 45 is the DNA sequence of TbH-12.
SEQ. ID NO. 46 is the DNA sequence of Tb38-1.
SEQ. ID NO. 47 is the DNA sequence of Tb38-4.
SEQ. ID NO. 48 is the DNA sequence of TbL-17.
SEQ. ID NO. 49 is the DNA sequence of TbL-20.
SEQ. ID NO. 50 is the DNA sequence of TbL-21.
SEQ. ID NO. 51 is the DNA sequence of TbH-16.
SEQ. ID NO. 52 is the DNA sequence of DPEP.
SEQ. ID NO. 53 is the deduced amino acid sequence of DPEP.
SEQ. ID NO. 54 is the protein sequence of DPV N-terminal Antigen.
SEQ. ID NO. 55 is the protein sequence of AVGS N-terminal Antigen.
SEQ. ID NO. 56 is the protein sequence of AAMK N-terminal Antigen.
SEQ. ID NO. 57 is the protein sequence of YYWC N-terminal Antigen.
SEQ. ID NO. 58 is the protein sequence of DIGS N-terminal Antigen.
SEQ. ID NO. 59 is the protein sequence of AEES N-terminal Antigen.
SEQ. ID NO. 60 is the protein sequence of DPEP N-terminal Antigen.
SEQ. ID NO. 61 is the protein sequence of APKT N-terminal Antigen.
SEQ. ID NO. 62 is the protein sequence of DPAS N-terminal Antigen.
SEQ. ID NO. 63 is the deduced amino acid sequence of TbRa1.
SEQ. ID NO. 64 is the deduced amino acid sequence of TbRa10.
SEQ. ID NO. 65 is the deduced amino acid sequence of TbRa11.
SEQ. ID NO. 66 is the deduced amino acid sequence of TbRa12.
SEQ. ID NO. 67 is the deduced amino acid sequence of TbRa13.
SEQ. ID NO. 68 is the deduced amino acid sequence of TbRa16.
SEQ. ID NO. 69 is the deduced amino acid sequence of TbRa17.
SEQ. ID NO. 70 is the deduced amino acid sequence of TbRa18.
SEQ. ID NO. 71 is the deduced amino acid sequence of TbRa19.
SEQ. ID NO. 72 is the deduced amino acid sequence of TbRa24.
SEQ. ID NO. 73 is the deduced amino acid sequence of TbRa26.
SEQ. ID NO. 74 is the deduced amino acid sequence of TbRa28.
SEQ. ID NO. 75 is the deduced amino acid sequence of TbRa29.
SEQ. ID NO. 76 is the deduced amino acid sequence of TbRa2A.
SEQ. ID NO. 77 is the deduced amino acid sequence of TbRa3.
SEQ. ID NO. 78 is the deduced amino acid sequence of TbRa32.
SEQ. ID NO. 79 is the deduced amino acid sequence of TbRa35.
SEQ. ID NO. 80 is the deduced amino acid sequence of TbRa36.
SEQ. ID NO. 81 is the deduced amino acid sequence of TbRa4.
SEQ. ID NO. 82 is the deduced amino acid sequence of TbRa9.
SEQ. ID NO. 83 is the deduced amino acid sequence of TbRaB.
SEQ. ID NO. 84 is the deduced amino acid sequence of TbRaC.
SEQ. ID NO. 85 is the deduced amino acid sequence of TbRaD.
SEQ. ID NO. 86 is the deduced amino acid sequence of YYWCPG.
SEQ. ID NO. 87 is the deduced amino acid sequence of TbAAMK.
SEQ. ID NO. 88 is the deduced amino acid sequence of Tb38-1.
SEQ. ID NO. 89 is the deduced amino acid sequence of TbH-4.
SEQ. ID NO. 90 is the deduced amino acid sequence of TbH-8.
SEQ. ID NO. 91 is the deduced amino acid sequence of TbH-9.
SEQ. ID NO. 92 is the deduced amino acid sequence of TbH-12.
SEQ. ID NO. 93 is the amino acid sequence of Tb38-1 Peptide 1.
SEQ. ID NO. 94 is the amino acid sequence of Tb38-1 Peptide 2.
SEQ. ID NO. 95 is the amino acid sequence of Tb38-1 Peptide 3.
SEQ. ID NO. 96 is the amino acid sequence of Tb38-1 Peptide 4.
SEQ. ID NO. 97 is the amino acid sequence of Tb38-1 Peptide 5.
SEQ. ID NO. 98 is the amino acid sequence of Tb38-1 Peptide 6.
SEQ. ID NO. 99 is the DNA sequence of DPAS.
SEQ. ID NO. 100 is the deduced amino acid sequence of DPAS.
SEQ. ID NO. 101 is the DNA sequence of DPV.
SEQ. ID NO. 102 is the deduced amino acid sequence of DPV.
SEQ. ID NO. 103 is the DNA sequence of ESAT-6.
SEQ. ID NO. 104 is the deduced amino acid sequence of ESAT-6.
SEQ. ID NO. 105 is the DNA sequence of TbH-8-2.
SEQ. ID NO. 106 is the DNA sequence of TbH-9FL.
SEQ. ID NO. 107 is the deduced amino acid sequence of TbH-9FL.
SEQ. ID NO. 108 is the DNA sequence of TbH-9-1.
SEQ. ID NO. 109 is the deduced amino acid sequence of TbH-9-1.
SEQ. ID NO. 110 is the DNA sequence of TbH-9-4.
SEQ. ID NO. 111 is the deduced amino acid sequence of TbH-9-4.
SEQ. ID NO. 112 is the DNA sequence of Tb38-1F2 IN.
SEQ. ID NO. 113 is the DNA sequence of Tb38-2F2 RP.
SEQ. ID NO. 114 is the deduced amino acid sequence of Tb37-FL.
SEQ. ID NO. 115 is the deduced amino acid sequence of Tb38-IN.
SEQ. ID NO. 116 is the DNA sequence of Tb38-1F3.
SEQ. ID NO. 117 is the deduced amino acid sequence of Tb38-1F3.
SEQ. ID NO. 118 is the DNA sequence of Tb38-1F5.
SEQ. ID NO. 119 is the DNA sequence of Tb38-1F6.
SEQ. ID NO. 120 is the deduced N-terminal amino acid sequence of DPV.
SEQ. ID NO. 121 is the deduced N-terminal amino acid sequence of AVGS.
SEQ. ID NO. 122 is the deduced N-terminal amino acid sequence of AAMK.
SEQ. ID NO. 123 is the deduced N-terminal amino acid sequence of YYWC.
SEQ. ID NO. 124 is the deduced N-terminal amino acid sequence of DIGS.
SEQ. ID NO. 125 is the deduced N-terminal amino acid sequence of AEES.
SEQ. ID NO. 126 is the deduced N-terminal amino acid sequence of DPEP.
SEQ. ID NO. 127 is the deduced N-terminal amino acid sequence of APKT.
SEQ. ID NO. 128 is the deduced amino acid sequence of DPAS.
SEQ. ID NO. 129 is the protein sequence of DPPD N-terminal Antigen.
SEQ ID NO. 130-133 are the protein sequences of four DPPD cyanogen bromide fragments.
SEQ ID NO. 134 is the N-terminal protein sequence of XDS antigen.
SEQ ID NO. 135 is the N-terminal protein sequence of AGD antigen.
SEQ ID NO. 136 is the N-terminal protein sequence of APE antigen.
SEQ ID NO. 137 is the N-terminal protein sequence of XYI antigen.
SEQ ID NO. 138 is the DNA sequence of TbH-29.
SEQ ID NO. 139 is the DNA sequence of TbH-30.
SEQ ID NO. 140 is the DNA sequence of TbH-32.
SEQ ID NO. 141 is the DNA sequence of TbH-33.
SEQ ID NO. 142 is the predicted amino acid sequence of TbH-29.
SEQ ID NO. 143 is the predicted amino acid sequence of TbH-30.
SEQ ID NO. 144 is the predicted amino acid sequence of TbH-32.
SEQ ID NO. 145 is the predicted amino acid sequence of TbH-33.
SEQ ID NO: 146-151 are PCR primers used in the preparation of a fusion protein containing TbRa3, 38 kD and Tb38-1.
SEQ ID NO: 152 is the DNA sequence of the fusion protein containing TbRa3, 38 kD and Tb38-1.
SEQ ID NO: 153 is the amino acid sequence of the fusion protein containing TbRa3, 38 kD and Tb38-1.
SEQ ID NO: 154 is the DNA sequence of the M. tuberculosis antigen 38 kD.
SEQ ID NO: 155 is the amino acid sequence of the M. tuberculosis antigen 38 kD.
SEQ ID NO: 156 is the DNA sequence of XP14.
SEQ ID NO: 157 is the DNA sequence of XP24.
SEQ ID NO: 158 is the DNA sequence of XP21.
SEQ ID NO: 159 is the 5' DNA sequence of XP22.
SEQ ID NO: 160 is the 3' DNA sequence of XP22.
SEQ ID NO: 161 is the predicted amino acid sequence of XP14.
SEQ ID NO: 162 is the predicted amino acid sequence encoded by the reverse complement of XP14.
SEQ ID NO: 163 is the DNA sequence of XP27.
SEQ ID NO: 164 is the DNA sequence of XP36.
SEQ ID NO: 165 is the 5' DNA sequence of XP4.
SEQ ID NO: 166 is the 5' DNA sequence of XP5.
SEQ ID NO: 167 is the 5' DNA sequence of XP17.
SEQ ID NO: 168 is the 5' DNA sequence of XP30.
SEQ ID NO: 169 is the 5' DNA sequence of XP2.
SEQ ID NO: 170 is the 3' DNA sequence of XP2.
SEQ ID NO: 171 is the 5' DNA sequence of XP3.
SEQ ID NO: 172 is the 3' DNA sequence of XP3.
SEQ ID NO: 173 is the 5' DNA sequence of XP6.
SEQ ID NO: 174 is the 3' DNA sequence of XP6.
SEQ ID NO: 175 is the 5' DNA sequence of XP18.
SEQ ID NO: 176 is the 3' DNA sequence of XP18.
SEQ ID NO: 177 is the 5' DNA sequence of XP19.
SEQ ID NO: 178 is the 3' DNA sequence of XP19.
SEQ ID NO: 179 is the 5' DNA sequence of XP22.
SEQ ID NO: 180 is the 3' DNA sequence of XP22.
SEQ ID NO: 181 is the 5' DNA sequence of XP25.
SEQ ID NO: 182 is the 3' DNA sequence of XP25.
SEQ ID NO: 183 is the full-length DNA sequence of TbH4-XP1.
SEQ ID NO: 184 is the predicted amino acid sequence of TbH4-XP1.
SEQ ID NO: 185 is the predicted amino acid sequence encoded by the reverse complement of TbH4-XP1.
SEQ ID NO: 186 is a first predicted amino acid sequence encoded by XP36.
SEQ ID NO: 187 is a second predicted amino acid sequence encoded by XP36.
SEQ ID NO: 188 is the predicted amino acid sequence encoded by the reverse complement of XP36.
SEQ ID NO: 189 is the DNA sequence of RDIF2.
SEQ ID NO: 190 is the DNA sequence of RDIF5.
SEQ ID NO: 191 is the DNA sequence of RDIF8.
SEQ ID NO: 192 is the DNA sequence of RDIF10.
SEQ ID NO: 193 is the DNA sequence of RDIF11.
SEQ ID NO: 194 is the predicted amino acid sequence of RDIF2.
SEQ ID NO: 195 is the predicted amino acid sequence of RDIF5.
SEQ ID NO: 196 is the predicted amino acid sequence of RDIF8.
SEQ ID NO: 197 is the predicted amino acid sequence of RDIF10.
SEQ ID NO: 198 is the predicted amino acid sequence of RDIF11.
SEQ ID NO: 199 is the 5' DNA sequence of RDIF12.
SEQ ID NO: 200 is the 3' DNA sequence of RDIF12.
SEQ ID NO: 201 is the DNA sequence of RDIF7.
SEQ ID NO: 202 is the predicted amino acid sequence of RDIF7.
SEQ ID NO: 203 is the DNA sequence of DIF2-1.
SEQ ID NO: 204 is the predicted amino acid sequence of DIF2-1.
SEQ ID NO: 205-212 are PCR primers used in the preparation of a fusion protein containing TbRa3, 38 kD, Tb38-1 and DPEP (hereinafter referred to as TbF-2).
SEQ ID NO: 213 is the DNA sequence of the fusion protein TbF-2.
SEQ ID NO: 214 is the amino acid sequence of the fusion protein TbF-2.
SEQ ID NO: 215 is the 5' DNA sequence of MO-1.
SEQ ID NO: 216 is the 5' DNA sequence for MO-2.
SEQ ID NO: 217 is the 5' DNA sequence for MO-4.
SEQ ID NO: 218 is the 5' DNA sequence for MO-8.
SEQ ID NO: 219 is the 5' DNA sequence for MO-9.
SEQ ID NO: 220 is the 5' DNA sequence for MO-26.
SEQ ID NO: 221 is the 5' DNA sequence for MO-28.
SEQ ID NO: 222 is the 5' DNA sequence for MO-29.
SEQ ID NO: 223 is the 5' DNA sequence for MO-30.
SEQ ID NO: 224 is the 5' DNA sequence for MO-34.
SEQ ID NO: 225 is the 5' DNA sequence for MO-35.
SEQ ID NO: 226 is the predicted amino acid sequence for MO-1.
SEQ ID NO: 227 is the predicted amino acid sequence for MO-2.
SEQ ID NO: 228 is the predicted amino acid sequence for MO-4.
SEQ ID NO: 229 is the predicted amino acid sequence for MO-8.
SEQ ID NO: 230 is the predicted amino acid sequence for MO-9.
SEQ ID NO: 231 is the predicted amino acid sequence for MO-26.
SEQ ID NO: 232 is the predicted amino acid sequence for MO-28.
SEQ ID NO: 233 is the predicted amino acid sequence for MO-29.
SEQ ID NO: 234 is the predicted amino acid sequence for MO-30.
SEQ ID NO: 235 is the predicted amino acid sequence for MO-34.
SEQ ID NO: 236 is the predicted amino acid sequence for MO-35.
SEQ ID NO: 237 is the determined DNA sequence for MO-10.
SEQ ID NO: 238 is the predicted amino acid sequence for MO-10.
SEQ ID NO: 239 is the 3' DNA sequence for MO-27.
SEQ ID NO: 240 is the full-length DNA sequence for DPPD.
SEQ ID NO: 241 is the predicted full-length amino acid sequence for DPPD.
SEQ ID NO: 242 is the determined 5' cDNA sequence for LSER-10
SEQ ID NO: 243 is the determined 5' cDNA sequence for LSER-11
SEQ ID NO: 244 is the determined 5' cDNA sequence for LSER-12
SEQ ID NO: 245 is the determined 5' cDNA sequence for LSER-13
SEQ ID NO: 246 is the determined 5' cDNA sequence for LSER-16
SEQ ID NO: 247 is the determined 5' cDNA sequence for LSER-25
SEQ ID NO: 248 is the predicted amino acid sequence for LSER-10
SEQ ID NO: 249 is the predicted amino acid sequence for LSER-12
SEQ ID NO: 250 is the predicted amino acid sequence for LSER-13
SEQ ID NO: 251 is the predicted amino acid sequence for LSER-16
SEQ ID NO: 252 is the predicted amino acid sequence for LSER-25
SEQ ID NO: 253 is the determined cDNA sequence for LSER-18
SEQ ID NO: 254 is the determined cDNA sequence for LSER-23
SEQ ID NO: 255 is the determined cDNA sequence for LSER-24
SEQ ID NO: 256 is the determined cDNA sequence for LSER-27
SEQ ID NO: 257 is the predicted amino acid sequence for LSER-18
SEQ ID NO: 258 is the predicted amino acid sequence for LSER-23
SEQ ID NO: 259 is the predicted amino acid sequence for LSER-24
SEQ ID NO: 260 is the predicted amino acid sequence for LSER-27
SEQ ID NO: 261 is the determined 5' cDNA sequence for LSER-1
SEQ ID NO: 262 is the determined 5' cDNA sequence for LSER-3
SEQ ID NO: 263 is the determined 5' cDNA sequence for LSER-4
SEQ ID NO: 264 is the determined 5' cDNA sequence for LSER-5
SEQ ID NO: 265 is the determined 5' cDNA sequence for LSER-6
SEQ ID NO: 266 is the determined 5' cDNA sequence for LSER-8
SEQ ID NO: 267 is the determined 5' cDNA sequence for LSER-14
SEQ ID NO: 268 is the determined 5' cDNA sequence for LSER-15
SEQ ID NO: 269 is the determined 5' cDNA sequence for LSER-17
SEQ ID NO: 270 is the determined 5' cDNA sequence for LSER-19
SEQ ID NO: 271 is the determined 5' cDNA sequence for LSER-20
SEQ ID NO: 272 is the determined 5' cDNA sequence for LSER-22
SEQ ID NO: 273 is the determined 5' cDNA sequence for LSER-26
SEQ ID NO: 274 is the determined 5' cDNA sequence for LSER-28
SEQ ID NO: 275 is the determined 5' cDNA sequence for LSER-29
SEQ ID NO: 276 is the determined 5' cDNA sequence for LSER-30
SEQ ID NO: 277 is the predicted amino acid sequence for LSER-1
SEQ ID NO: 278 is the predicted amino acid sequence for LSER-3
SEQ ID NO: 279 is the predicted amino acid sequence for LSER-5
SEQ ID NO: 280 is the predicted amino acid sequence for LSER-6
SEQ ID NO: 281 is the predicted amino acid sequence for LSER-8
SEQ ID NO: 282 is the predicted amino acid sequence for LSER-14
SEQ ID NO: 283 is the predicted amino acid sequence for LSER-15
SEQ ID NO: 284 is the predicted amino acid sequence for LSER-17
SEQ ID NO: 285 is the predicted amino acid sequence for LSER-19
SEQ ID NO: 286 is the predicted amino acid sequence for LSER-20
SEQ ID NO: 287 is the predicted amino acid sequence for LSER-22
SEQ ID NO: 288 is the predicted amino acid sequence for LSER-26
SEQ ID NO: 289 is the predicted amino acid sequence for LSER-28
SEQ ID NO: 290 is the predicted amino acid sequence for LSER-29
SEQ ID NO: 291 is the predicted amino acid sequence for LSER-30
SEQ ID NO: 292 is the determined cDNA sequence for LSER-9
SEQ ID NO: 293 is the determined cDNA sequence for the reverse complement of LSER-6
SEQ ID NO: 294 is the predicted amino acid sequence for the reverse complement of LSER-6
SEQ ID NO: 295 is the determined 5’ cDNA sequence for MO-12
SEQ ID NO: 296 is the determined 5’ cDNA sequence for MO-13
SEQ ID NO: 297 is the determined 5’ cDNA sequence for MO-19
SEQ ID NO: 298 is the determined 5’ cDNA sequence for MO-39
SEQ ID NO: 299 is the predicted amino acid sequence for MO-12
SEQ ID NO: 300 is the predicted amino acid sequence for MO-13
SEQ ID NO: 301 is the predicted amino acid sequence for MO-19
SEQ ID NO: 302 is the predicted amino acid sequence for MO-39
SEQ ID NO: 303 is the determined 5’ cDNA sequence for Erdsn-1
SEQ ID NO: 304 is the determined 5’ cDNA sequence for Erdsn-2
SEQ ID NO: 305 is the determined 5’ cDNA sequence for Erdsn-4
SEQ ID NO: 306 is the determined 5’ cDNA sequence for Erdsn-5
SEQ ID NO: 307 is the determined 5’ cDNA sequence for Erdsn-6
SEQ ID NO: 308 is the determined 5’ cDNA sequence for Erdsn-7
SEQ ID NO: 309 is the determined 5’ cDNA sequence for Erdsn-8
SEQ ID NO: 310 is the determined 5’ cDNA sequence for Erdsn-9
SEQ ID NO: 311 is the determined 5’ cDNA sequence for Erdsn-10
SEQ ID NO: 312 is the determined 5’ cDNA sequence for Erdsn-12
SEQ ID NO: 313 is the determined 5’ cDNA sequence for Erdsn-13
SEQ ID NO: 314 is the determined 5’ cDNA sequence for Erdsn-14
SEQ ID NO: 315 is the determined 5’ cDNA sequence for Erdsn-15
SEQ ID NO: 316 is the determined 5' cDNA sequence for Erdsn-16
SEQ ID NO: 317 is the determined 5' cDNA sequence for Erdsn-17
SEQ ID NO: 318 is the determined 5' cDNA sequence for Erdsn-18
SEQ ID NO: 319 is the determined 5' cDNA sequence for Erdsn-21
SEQ ID NO: 320 is the determined 5' cDNA sequence for Erdsn-22
SEQ ID NO: 321 is the determined 5' cDNA sequence for Erdsn-23
SEQ ID NO: 322 is the determined 5' cDNA sequence for Erdsn-25
SEQ ID NO: 323 is the determined 3' cDNA sequence for Erdsn-1
SEQ ID NO: 324 is the determined 3' cDNA sequence for Erdsn-2
SEQ ID NO: 325 is the determined 3' cDNA sequence for Erdsn-4
SEQ ID NO: 326 is the determined 3' cDNA sequence for Erdsn-5
SEQ ID NO: 327 is the determined 3' cDNA sequence for Erdsn-7
SEQ ID NO: 328 is the determined 3' cDNA sequence for Erdsn-8
SEQ ID NO: 329 is the determined 3' cDNA sequence for Erdsn-9
SEQ ID NO: 330 is the determined 3' cDNA sequence for Erdsn-10
SEQ ID NO: 331 is the determined 3' cDNA sequence for Erdsn-12
SEQ ID NO: 332 is the determined 3' cDNA sequence for Erdsn-13
SEQ ID NO: 333 is the determined 3' cDNA sequence for Erdsn-14
SEQ ID NO: 334 is the determined 3' cDNA sequence for Erdsn-15
SEQ ID NO: 335 is the determined 3' cDNA sequence for Erdsn-16
SEQ ID NO: 336 is the determined 3' cDNA sequence for Erdsn-17
SEQ ID NO: 337 is the determined 3' cDNA sequence for Erdsn-18
SEQ ID NO: 338 is the determined 3' cDNA sequence for Erdsn-21
SEQ ID NO: 339 is the determined 3' cDNA sequence for Erdsn-22
SEQ ID NO: 340 is the determined 3' cDNA sequence for Erdsn-23
SEQ ID NO: 341 is the determined 3' cDNA sequence for Erdsn-25
SEQ ID NO: 342 is the determined cDNA sequence for Erdsn-24
SEQ ID NO: 343 is the determined amino acid sequence for a *M. tuberculosis*
85b precursor homolog
SEQ ID NO: 344 is the determined amino acid sequence for spot 1
SEQ ID NO: 345 is a determined amino acid sequence for spot 2
SEQ ID NO: 346 is a determined amino acid sequence for spot 2
SEQ ID NO: 347 is the determined amino acid seq for spot 4
SEQ ID NO: 348 is the sequence of primer PDM-157
SEQ ID NO: 349 is the sequence of primer PDM-160
SEQ ID NO: 350 is the DNA sequence of the fusion protein TbF-6
SEQ ID NO: 351 is the amino acid sequence of fusion protein TbF-6
SEQ ID NO: 352 is the sequence of primer PDM-176
SEQ ID NO: 353 is the sequence of primer PDM-175
SEQ ID NO: 354 is the DNA sequence of the fusion protein TbF-8
SEQ ID NO: 355 is the amino acid sequence of the fusion protein TbF-8

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for preventing, treating and diagnosing tuberculosis. The compositions of the subject invention include polypeptides that comprise at least one immunogenic portion of a *M. tuberculosis* antigen, or a variant of such an antigen that differs only in conservative substitutions and/or modifications. Polypeptides within the scope of the present invention include, but are not limited to, immunogenic soluble *M. tuberculosis* antigens. A "soluble *M. tuberculosis* antigen" is a protein of *M. tuberculosis* origin that is present in *M. tuberculosis* culture filtrate. As used herein, the term "polypeptide" encompasses amino acid chains of any length, including full length proteins (i.e., antigens), wherein the amino acid residues are linked by covalent peptide bonds. Thus, a polypeptide comprising an immunogenic portion of one of the above antigens may consist entirely of the immunogenic portion, or may contain additional sequences. The additional sequences may be derived from the native *M. tuberculosis* antigen or may be heterologous, and such sequences may (but need not) be immunogenic.

"Immunogenic," as used herein, refers to the ability to elicit an immune response (e.g., cellular) in a patient, such as a human, and/or in a biological sample. In particular, antigens that are immunogenic (and immunogenic portions or other variants
of such antigens) are capable of stimulating cell proliferation, interleukin-12 production and/or interferon-γ production in biological samples comprising one or more cells selected from the group of T cells, NK cells, B cells and macrophages, where the cells are derived from an *M. tuberculosis*-immune individual. Polypeptides comprising at least an immunogenic portion of one or more *M. tuberculosis* antigens may generally be used to detect tuberculosis or to induce protective immunity against tuberculosis in a patient.

The compositions and methods of the present invention also encompass variants of the above polypeptides and DNA molecules. A polypeptide "variant," as used herein, is a polypeptide that differs from the recited polypeptide only in conservative substitutions and/or modifications, such that the therapeutic, antigenic and/or immunogenic properties of the polypeptide are retained. Polypeptide variants preferably exhibit at least about 70%, more preferably at least about 90% and most preferably at least about 95% identity to the identified polypeptides. For polypeptides with immunoreactive properties, variants may, alternatively, be identified by modifying the amino acid sequence of one of the above polypeptides, and evaluating the immunoreactivity of the modified polypeptide. For polypeptides useful for the generation of diagnostic binding agents, a variant may be identified by evaluating a modified polypeptide for the ability to generate antibodies that detect the presence or absence of tuberculosis. Such modified sequences may be prepared and tested using, for example, the representative procedures described herein.

As used herein, a "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydrophobic nature of the polypeptide to be substantially unchanged. In general, the following groups of amino acids represent conservative changes: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his.

Variants may also, or alternatively, contain other modifications, including the deletion or addition of amino acids that have minimal influence on the
antigenic properties, secondary structure and hydropathic nature of the polypeptide. For example, a polypeptide may be conjugated to a signal (or leader) sequence at the N-terminal end of the protein which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (e.g., poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

A nucleotide “variant” is a sequence that differs from the recited nucleotide sequence in having one or more nucleotide deletions, substitutions or additions. Such modifications may be readily introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis as taught, for example, by Adelman et al. (DNA, 2:183, 1983). Nucleotide variants may be naturally occurring allelic variants, or non-naturally occurring variants. Variant nucleotide sequences preferably exhibit at least about 70%, more preferably at least about 80% and most preferably at least about 90% identity to the recited sequence. Such variant nucleotide sequences will generally hybridize to the recited nucleotide sequence under stringent conditions. As used herein, “stringent conditions” refers to prewashing in a solution of 6X SSC, 0.2% SDS; hybridizing at 65 °C. 6X SSC, 0.2% SDS overnight; followed by two washes of 30 minutes each in 1X SSC, 0.1% SDS at 65 °C and two washes of 30 minutes each in 0.2X SSC, 0.1% SDS at 65 °C.

In a related aspect, combination polypeptides are disclosed. A "combination polypeptide" is a polypeptide comprising at least one of the above immunogenic portions and one or more additional immunogenic M. tuberculosis sequences, which are joined via a peptide linkage into a single amino acid chain. The sequences may be joined directly (i.e., with no intervening amino acids) or may be joined by way of a linker sequence (e.g., Gly-Cys-Gly) that does not significantly diminish the immunogenic properties of the component polypeptides.

In general, M. tuberculosis antigens, and DNA sequences encoding such antigens, may be prepared using any of a variety of procedures. For example, soluble antigens may be isolated from M. tuberculosis culture filtrate by procedures known to
those of ordinary skill in the art, including anion-exchange and reverse phase chromatography. Purified antigens are then evaluated for their ability to elicit an appropriate immune response (e.g., cellular) using, for example, the representative methods described herein. Immunogenic antigens may then be partially sequenced using techniques such as traditional Edman chemistry. See Edman and Berg, Eur. J. Biochem. 80:116-132, 1967.

Immunogenic antigens may also be produced recombinantly using a DNA sequence that encodes the antigen, which has been inserted into an expression vector and expressed in an appropriate host. DNA molecules encoding soluble antigens may be isolated by screening an appropriate *M. tuberculosis* expression library with anti-sera (e.g., rabbit) raised specifically against soluble *M. tuberculosis* antigens. DNA sequences encoding antigens that may or may not be soluble may be identified by screening an appropriate *M. tuberculosis* genomic or cDNA expression library with sera obtained from patients infected with *M. tuberculosis*. Such screens may generally be performed using techniques well known to those of ordinary skill in the art, such as those described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989.

DNA sequences encoding soluble antigens may also be obtained by screening an appropriate *M. tuberculosis* cDNA or genomic DNA library for DNA sequences that hybridize to degenerate oligonucleotides derived from partial amino acid sequences of isolated soluble antigens. Degenerate oligonucleotide sequences for use in such a screen may be designed and synthesized, and the screen may be performed, as described (for example) in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989 (and references cited therein). Polymerase chain reaction (PCR) may also be employed, using the above oligonucleotides in methods well known in the art, to isolate a nucleic acid probe from a cDNA or genomic library. The library screen may then be performed using the isolated probe.

Alternatively, genomic or cDNA libraries derived from *M. tuberculosis* may be screened directly using peripheral blood mononuclear cells (PBMCs) or T cell
lines or clones derived from one or more *M. tuberculosis*-immune individuals. In general, PBMCs and/or T cells for use in such screens may be prepared as described below. Direct library screens may generally be performed by assaying pools of expressed recombinant proteins for the ability to induce proliferation and/or interferon-γ production in T cells derived from an *M. tuberculosis*-immune individual. Alternatively, potential T cell antigens may be first selected based on antibody reactivity, as described above.

Regardless of the method of preparation, the antigens (and immunogenic portions thereof) described herein (which may or may not be soluble) have the ability to induce an immunogenic response. More specifically, the antigens have the ability to induce proliferation and/or cytokine production (*i.e.*, interferon-γ and/or interleukin-12 production) in T cells, NK cells, B cells and/or macrophages derived from an *M. tuberculosis*-immune individual. The selection of cell type for use in evaluating an immunogenic response to a antigen will, of course, depend on the desired response. For example, interleukin-12 production is most readily evaluated using preparations containing B cells and/or macrophages. An *M. tuberculosis*-immune individual is one who is considered to be resistant to the development of tuberculosis by virtue of having mounted an effective T cell response to *M. tuberculosis* (*i.e.*, substantially free of disease symptoms). Such individuals may be identified based on a strongly positive (*i.e.*, greater than about 10 mm diameter induration) intradermal skin test response to tuberculosis proteins (PPD) and an absence of any signs or symptoms of tuberculosis disease. T cells, NK cells. B cells and macrophages derived from *M. tuberculosis*-immune individuals may be prepared using methods known to those of ordinary skill in the art. For example, a preparation of PBMCs (*i.e.*, peripheral blood mononuclear cells) may be employed without further separation of component cells. PBMCs may generally be prepared, for example, using density centrifugation through Ficoll™ (Winthrop Laboratories, NY). T cells for use in the assays described herein may also be purified directly from PBMCs. Alternatively, an enriched T cell line reactive against mycobacterial proteins, or T cell clones reactive to individual mycobacterial proteins, may be employed. Such T cell clones may be generated by, for example, culturing
PBMCs from *M. tuberculosis*-immune individuals with mycobacterial proteins for a period of 2-4 weeks. This allows expansion of only the mycobacterial protein-specific T cells, resulting in a line composed solely of such cells. These cells may then be cloned and tested with individual proteins, using methods known to those of ordinary skill in the art, to more accurately define individual T cell specificity. In general, antigens that test positive in assays for proliferation and/or cytokine production (i.e., interferon-γ and/or interleukin-12 production) performed using T cells, NK cells, B cells and/or macrophages derived from an *M. tuberculosis*-immune individual are considered immunogenic. Such assays may be performed, for example, using the representative procedures described below. Immunogenic portions of such antigens may be identified using similar assays, and may be present within the polypeptides described herein.

The ability of a polypeptide (e.g., an immunogenic antigen, or a portion or other variant thereof) to induce cell proliferation is evaluated by contacting the cells (e.g., T cells and/or NK cells) with the polypeptide and measuring the proliferation of the cells. In general, the amount of polypeptide that is sufficient for evaluation of about $10^5$ cells ranges from about 10 ng/mL to about 100 μg/mL and preferably is about 10 μg/mL. The incubation of polypeptide with cells is typically performed at 37°C for about six days. Following incubation with polypeptide, the cells are assayed for a proliferative response, which may be evaluated by methods known to those of ordinary skill in the art, such as exposing cells to a pulse of radiolabeled thymidine and measuring the incorporation of label into cellular DNA. In general, a polypeptide that results in at least a three fold increase in proliferation above background (i.e., the proliferation observed for cells cultured without polypeptide) is considered to be able to induce proliferation.

The ability of a polypeptide to stimulate the production of interferon-γ and/or interleukin-12 in cells may be evaluated by contacting the cells with the polypeptide and measuring the level of interferon-γ or interleukin-12 produced by the cells. In general, the amount of polypeptide that is sufficient for the evaluation of about $10^5$ cells ranges from about 10 ng/mL to about 100 μg/mL and preferably is about 10 μg/mL. The polypeptide may, but need not, be immobilized on a solid support, such as a
bead or a biodegradable microsphere, such as those described in U.S. Patent Nos. 4,897,268 and 5,075,109. The incubation of polypeptide with the cells is typically performed at 37°C for about six days. Following incubation with polypeptide, the cells are assayed for interferon-γ and/or interleukin-12 (or one or more subunits thereof), which may be evaluated by methods known to those of ordinary skill in the art, such as an enzyme-linked immunosorbent assay (ELISA) or, in the case of IL-12 P70 subunit, a bioassay such as an assay measuring proliferation of T cells. In general, a polypeptide that results in the production of at least 50 pg of interferon-γ per mL of cultured supernatant (containing $10^4$-$10^5$ T cells per mL) is considered able to stimulate the production of interferon-γ. A polypeptide that stimulates the production of at least 10 pg/mL of IL-12 P70 subunit, and/or at least 100 pg/mL of IL-12 P40 subunit, per $10^5$ macrophages or B cells (or per $3 \times 10^5$ PBMC) is considered able to stimulate the production of IL-12.

In general, immunogenic antigens are those antigens that stimulate proliferation and/or cytokine production (i.e., interferon-γ and/or interleukin-12 production) in T cells, NK cells, B cells and/or macrophages derived from at least about 25% of M. tuberculosis-immune individuals. Among these immunogenic antigens, polypeptides having superior therapeutic properties may be distinguished based on the magnitude of the responses in the above assays and based on the percentage of individuals for which a response is observed. In addition, antigens having superior therapeutic properties will not stimulate proliferation and/or cytokine production in vitro in cells derived from more than about 25% of individuals that are not M. tuberculosis-immune, thereby eliminating responses that are not specifically due to M. tuberculosis-responsive cells. Those antigens that induce a response in a high percentage of T cell, NK cell, B cell and/or macrophage preparations from M. tuberculosis-immune individuals (with a low incidence of responses in cell preparations from other individuals) have superior therapeutic properties.

Antigens with superior therapeutic properties may also be identified based on their ability to diminish the severity of M. tuberculosis infection in experimental animals, when administered as a vaccine. Suitable vaccine preparations
for use on experimental animals are described in detail below. Efficacy may be
determined based on the ability of the antigen to provide at least about a 50% reduction
in bacterial numbers and/or at least about a 40% decrease in mortality following
experimental infection. Suitable experimental animals include mice, guinea pigs and
primates.

Antigens having superior diagnostic properties may generally be
identified based on the ability to elicit a response in an intradermal skin test performed
on an individual with active tuberculosis, but not in a test performed on an individual
who is not infected with *M. tuberculosis*. Skin tests may generally be performed as
described below, with a response of at least 5 mm induration considered positive.

Immunogenic portions of the antigens described herein may be prepared
and identified using well known techniques, such as those summarized in Paul,
*Fundamental Immunology*, 3d ed., Raven Press, 1993, pp. 243-247 and references cited
therein. Such techniques include screening polypeptide portions of the native antigen
for immunogenic properties. The representative proliferation and cytokine production
assays described herein may generally be employed in these screens. An immunogenic
portion of a polypeptide is a portion that, within such representative assays, generates
an immune response (*e.g.*, proliferation, interferon-γ production and/or interleukin-12
production) that is substantially similar to that generated by the full length antigen. In
other words, an immunogenic portion of an antigen may generate at least about 20%,
and preferably about 100%, of the proliferation induced by the full length antigen in the
model proliferation assay described herein. An immunogenic portion may also, or
alternatively, stimulate the production of at least about 20%, and preferably about
100%, of the interferon-γ and/or interleukin-12 induced by the full length antigen in the
model assay described herein.

Portions and other variants of *M. tuberculosis* antigens may be generated
by synthetic or recombinant means. Synthetic polypeptides having fewer than about
100 amino acids, and generally fewer than about 50 amino acids, may be generated
using techniques well known to those of ordinary skill in the art. For example, such
polypeptides may be synthesized using any of the commercially available solid-phase
techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. See Merrifield, *J. Am. Chem. Soc.* 85:2149-2146, 1963. Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Applied BioSystems, Inc., Foster City, CA, and may be operated according to the manufacturer's instructions. Variants of a native antigen may generally be prepared using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis. Sections of the DNA sequence may also be removed using standard techniques to permit preparation of truncated polypeptides.

Recombinant polypeptides containing portions and/or variants of a native antigen may be readily prepared from a DNA sequence encoding the polypeptide using a variety of techniques well known to those of ordinary skill in the art. For example, supernatants from suitable host/vector systems which secrete recombinant protein into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant protein.

Any of a variety of expression vectors known to those of ordinary skill in the art may be employed to express recombinant polypeptides of this invention. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast and higher eukaryotic cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. The DNA sequences expressed in this manner may encode naturally occurring antigens, portions of naturally occurring antigens, or other variants thereof.

In general, regardless of the method of preparation, the polypeptides disclosed herein are prepared in substantially pure form. Preferably, the polypeptides are at least about 80% pure, more preferably at least about 90% pure and most
preferably at least about 99% pure. In certain preferred embodiments, described in
detail below, the substantially pure polypeptides are incorporated into pharmaceutical
compositions or vaccines for use in one or more of the methods disclosed herein.

In certain specific embodiments, the subject invention discloses polypeptides comprising at least an immunogenic portion of a soluble *M. tuberculosis*
antigen having one of the following N-terminal sequences, or a variant thereof that
differs only in conservative substitutions and/or modifications:

(a)  Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-
     Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)

(b)  Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-
     Ser; (SEQ ID No. 121)

(c)  Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-
     Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)

(d)  Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-
     Pro; (SEQ ID No. 123)

(e)  Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val;
     (SEQ ID No. 124)

(f)  Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID
     No. 125)

(g)  Asp-Pro-Glu-Pro-Ala-Pro-Val-Pro-Thr-Ala-Ala-Ala-Ser-
     Pro-Pro-Ser; (SEQ ID No. 126)

(h)  Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-
     Gly; (SEQ ID No. 127)

(i)  Asp-Pro-Ala-Ser-Ala-Pro-Ala-Pro-Thr-Ala-Ala-Gln-Leu-
     Thr-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-
     Ala-Asn; (SEQ ID No. 128)

(j)  Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-
     Ser; (SEQ ID No. 134)

(k)  Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-
     Asp; (SEQ ID No. 135) or
wherein Xaa may be any amino acid, preferably a cysteine residue. A DNA sequence encoding the antigen identified as (g) above is provided in SEQ ID No. 52, and the polypeptide encoded by SEQ ID No. 52 is provided in SEQ ID No. 53. A DNA sequence encoding the antigen defined as (a) above is provided in SEQ ID No. 101; its deduced amino acid sequence is provided in SEQ ID No. 102. A DNA sequence corresponding to antigen (d) above is provided in SEQ ID No. 24 a DNA sequence corresponding to antigen (c) is provided in SEQ ID No. 25 and a DNA sequence corresponding to antigen (i) is provided in SEQ ID No. 99; its deduced amino acid sequence is provided in SEQ ID No. 100.

In a further specific embodiment, the subject invention discloses polypeptides comprising at least an immunogenic portion of an M. tuberculosis antigen having one of the following N-terminal sequences, or a variant thereof that differs only in conservative substitutions and/or modifications:

(m) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No 137) or

(n) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129)

wherein Xaa may be any amino acid, preferably a cysteine residue.

In other specific embodiments, the subject invention discloses polypeptides comprising at least an immunogenic portion of a soluble M. tuberculosis antigen (or a variant of such an antigen) that comprises one or more of the amino acid sequences encoded by (a) the DNA sequences of SEQ ID Nos.: 1, 2, 4-10, 13-25 and 52; (b) the complements of such DNA sequences, or (c) DNA sequences substantially homologous to a sequence in (a) or (b).

In further specific embodiments, the subject invention discloses polypeptides comprising at least an immunogenic portion of a M. tuberculosis antigen (or a variant of such an antigen), which may or may not be soluble, that comprises one or more of the amino acid sequences encoded by (a) the DNA sequences of SEQ ID
Nos.: 26-51, 138, 139, 163-183, 189-193, 199, 200, 201, 203, 215-225, 239, 240, 242-247, 253-256, 261-276, 292, 293, 295-298 and 303-342, (b) the complements of such DNA sequences or (c) DNA sequences substantially homologous to a sequence in (a) or (b).

In the specific embodiments discussed above, the *M. tuberculosis* antigens include variants that are encoded by DNA sequences which are substantially homologous to one or more of DNA sequences specifically recited herein. "Substantial homology," as used herein, refers to DNA sequences that are capable of hybridizing under moderately stringent conditions. Suitable moderately stringent conditions include prewashing in a solution of 5X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5X SSC, overnight or, in the case of cross-species homology at 45°C, 0.5X SSC; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS). Such hybridizing DNA sequences are also within the scope of this invention, as are nucleotide sequences that, due to code degeneracy, encode an immunogenic polypeptide that is encoded by a hybridizing DNA sequence.

In a related aspect, the present invention provides fusion proteins comprising a first and a second inventive polypeptide or, alternatively, a polypeptide of the present invention and a known *M. tuberculosis* antigen, such as the 38 kD antigen described in Andersen and Hansen, *Infect. Immun.* 57:2481-2488, 1989. (Genbank Accession No. M30046) or ESAT-6 (SEQ ID Nos. 103 and 104), together with variants of such fusion proteins. The fusion proteins of the present invention may also include a linker peptide between the first and second polypeptides.

A DNA sequence encoding a fusion protein of the present invention is constructed using known recombinant DNA techniques to assemble separate DNA sequences encoding the first and second polypeptides into an appropriate expression vector. The 3' end of a DNA sequence encoding the first polypeptide is ligated, with or without a peptide linker, to the 5'-end of a DNA sequence encoding the second polypeptide so that the reading frames of the sequences are in phase to permit mRNA
translation of the two DNA sequences into a single fusion protein that retains the biological activity of both the first and the second polypeptides.

A peptide linker sequence may be employed to separate the first and the second polypeptides by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea et al., *Gene* 40:39-46, 1985; Murphy et al., *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may be from 1 to about 50 amino acids in length. Peptide sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons require to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

In another aspect, the present invention provides methods for using one or more of the above polypeptides or fusion proteins (or DNA molecules encoding such polypeptides) to induce protective immunity against tuberculosis in a patient. As used herein, a "patient" refers to any warm-blooded animal, preferably a human. A patient
may be afflicted with a disease, or may be free of detectable disease and/or infection. In other words, protective immunity may be induced to prevent or treat tuberculosis.

In this aspect, the polypeptide, fusion protein or DNA molecule is generally present within a pharmaceutical composition and/or a vaccine. Pharmaceutical compositions may comprise one or more polypeptides, each of which may contain one or more of the above sequences (or variants thereof), and a physiologically acceptable carrier. Vaccines may comprise one or more of the above polypeptides and a non-specific immune response enhancer, such as an adjuvant or a liposome (into which the polypeptide is incorporated). Such pharmaceutical compositions and vaccines may also contain other \textit{M. tuberculosis} antigens, either incorporated into a combination polypeptide or present within a separate polypeptide.

Alternatively, a vaccine may contain DNA encoding one or more polypeptides as described above, such that the polypeptide is generated \textit{in situ}. In such vaccines, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacterial and viral expression systems. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as \textit{Bacillus-Calmette-Guerrin}) that expresses an immunogenic portion of the polypeptide on its cell surface. In a preferred embodiment, the DNA may be introduced using a viral expression system (\textit{e.g.}, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer et al., \textit{Science} 259:1745-1749, 1993 and reviewed by Cohen, \textit{Science} 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells.

In a related aspect, a DNA vaccine as described above may be administered simultaneously with or sequentially to either a polypeptide of the present
invention or a known *M. tuberculosis* antigen, such as the 38 kD antigen described above. For example, administration of DNA encoding a polypeptide of the present invention, either "naked" or in a delivery system as described above, may be followed by administration of an antigen in order to enhance the protective immune effect of the vaccine.

Routes and frequency of administration, as well as dosage, will vary from individual to individual and may parallel those currently being used in immunization using BCG. In general, the pharmaceutical compositions and vaccines may be administered by injection (e.g., intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (e.g., by aspiration) or orally. Between 1 and 3 doses may be administered for a 1-36 week period. Preferably, 3 doses are administered, at intervals of 3-4 months, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of polypeptide or DNA that, when administered as described above, is capable of raising an immune response in an immunized patient sufficient to protect the patient from *M. tuberculosis* infection for at least 1-2 years. In general, the amount of polypeptide present in a dose (or produced *in situ* by the DNA in a dose) ranges from about 1 pg to about 100 mg per kg of host, typically from about 10 pg to about 1 mg, and preferably from about 100 pg to about 1 μg. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

While any suitable carrier known to those of ordinary skill in the art may be employed in the pharmaceutical compositions of this invention, the type of carrier will vary depending on the mode of administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactic galactide) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268 and 5,075,109.
Any of a variety of adjuvants may be employed in the vaccines of this invention to nonspecifically enhance the immune response. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a nonspecific stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis*. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Freund's Complete Adjuvant (Difco Laboratories) and Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ). Other suitable adjuvants include alum, biodegradable microspheres, monophosphoryl lipid A and quill A.

In another aspect, this invention provides methods for using one or more of the polypeptides described above to diagnose tuberculosis using a skin test. As used herein, a "skin test" is any assay performed directly on a patient in which a delayed-type hypersensitivity (DTH) reaction (such as swelling, reddening or dermatitis) is measured following intradermal injection of one or more polypeptides as described above. Such injection may be achieved using any suitable device sufficient to contact the polypeptide or polypeptides with dermal cells of the patient, such as a tuberculin syringe or 1 mL syringe. Preferably, the reaction is measured at least 48 hours after injection, more preferably 48-72 hours.

The DTH reaction is a cell-mediated immune response, which is greater in patients that have been exposed previously to the test antigen (*i.e.*, the immunogenic portion of the polypeptide employed, or a variant thereof). The response may be measured visually, using a ruler. In general, a response that is greater than about 0.5 cm in diameter, preferably greater than about 1.0 cm in diameter, is a positive response, indicative of tuberculosis infection, which may or may not be manifested as an active disease.

The polypeptides of this invention are preferably formulated, for use in a skin test, as pharmaceutical compositions containing a polypeptide and a physiologically acceptable carrier, as described above. Such compositions typically contain one or more of the above polypeptides in an amount ranging from about 1 µg to about 100 µg, preferably from about 10 µg to about 50 µg in a volume of 0.1 mL.
Preferably, the carrier employed in such pharmaceutical compositions is a saline solution with appropriate preservatives, such as phenol and/or Tween 80™.

In a preferred embodiment, a polypeptide employed in a skin test is of sufficient size such that it remains at the site of injection for the duration of the reaction period. In general, a polypeptide that is at least 9 amino acids in length is sufficient. The polypeptide is also preferably broken down by macrophages within hours of injection to allow presentation to T-cells. Such polypeptides may contain repeats of one or more of the above sequences and/or other immunogenic or nonimmunogenic sequences.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLES

EXAMPLE 1

PURIFICATION AND CHARACTERIZATION OF POLYPEPTIDES FROM M. TUBERCULOSIS CULTURE FILTRATE

This example illustrates the preparation of M. tuberculosis soluble polypeptides from culture filtrate. Unless otherwise noted, all percentages in the following example are weight per volume.

M. tuberculosis (either H37Ra, ATCC No. 25177, or H37Rv, ATCC No. 25618) was cultured in sterile GAS media at 37°C for fourteen days. The media was then vacuum filtered (leaving the bulk of the cells) through a 0.45 µ filter into a sterile 2.5 L bottle. The media was next filtered through a 0.2 µ filter into a sterile 4 L bottle and NaN₃ was added to the culture filtrate to a concentration of 0.04%. The bottles were then placed in a 4°C cold room.

The culture filtrate was concentrated by placing the filtrate in a 12 L reservoir that had been autoclaved and feeding the filtrate into a 400 ml Amicon stir cell
which had been rinsed with ethanol and contained a 10,000 kDa MWCO membrane. The pressure was maintained at 60 psi using nitrogen gas. This procedure reduced the 12 L volume to approximately 50 ml.

The culture filtrate was dialyzed into 0.1% ammonium bicarbonate using a 8,000 kDa MWCO cellulose ester membrane, with two changes of ammonium bicarbonate solution. Protein concentration was then determined by a commercially available BCA assay (Pierce, Rockford, IL).

The dialyzed culture filtrate was then lyophilized, and the polypeptides resuspended in distilled water. The polypeptides were dialyzed against 0.01 mM 1.3 bis[tris(hydroxymethyl)-methylamino]propane. pH 7.5 (Bis-Tris propane buffer), the initial conditions for anion exchange chromatography. Fractionation was performed using gel perfusion chromatography on a POROS 146 II Q/M anion exchange column 4.6 mm x 100 mm (Perseptive BioSystems, Framingham, MA) equilibrated in 0.01 mM Bis-Tris propane buffer pH 7.5. Polypeptides were eluted with a linear 0-0.5 M NaCl gradient in the above buffer system. The column eluent was monitored at a wavelength of 220 nm.

The pools of polypeptides eluting from the ion exchange column were dialyzed against distilled water and lyophilized. The resulting material was dissolved in 0.1% trifluoroacetic acid (TFA) pH 1.9 in water. and the polypeptides were purified on a Delta-Pak C18 column (Waters, Milford, MA) 300 Angstrom pore size, 5 micron particle size (3.9 x 150 mm). The polypeptides were eluted from the column with a linear gradient from 0-60% dilution buffer (0.1% TFA in acetonitrile). The flow rate was 0.75 ml/minute and the HPLC eluent was monitored at 214 nm. Fractions containing the eluted polypeptides were collected to maximize the purity of the individual samples. Approximately 200 purified polypeptides were obtained.

The purified polypeptides were then screened for the ability to induce T-cell proliferation in PBMC preparations. The PBMCs from donors known to be PPD skin test positive and whose T-cells were shown to proliferate in response to PPD and crude soluble proteins from MTB were cultured in medium comprising RPMI 1640 supplemented with 10% pooled human serum and 50 μg/ml gentamicin. Purified
polypeptides were added in duplicate at concentrations of 0.5 to 10 μg/mL. After six days of culture in 96-well round-bottom plates in a volume of 200 μl, 50 μl of medium was removed from each well for determination of IFN-γ levels, as described below. The plates were then pulsed with 1 μCi/well of tritiated thymidine for a further 18 hours, harvested and tritium uptake determined using a gas scintillation counter. Fractions that resulted in proliferation in both replicates three fold greater than the proliferation observed in cells cultured in medium alone were considered positive.

IFN-γ was measured using an enzyme-linked immunosorbent assay (ELISA). ELISA plates were coated with a mouse monoclonal antibody directed to human IFN-γ (PharMingen, San Diego, CA) in PBS for four hours at room temperature. Wells were then blocked with PBS containing 5% (W/V) non-fat dried milk for 1 hour at room temperature. The plates were then washed six times in PBS/0.2% TWEEN-20 and samples diluted 1:2 in culture medium in the ELISA plates were incubated overnight at room temperature. The plates were again washed and a polyclonal rabbit anti-human IFN-γ serum diluted 1:3000 in PBS/10% normal goat serum was added to each well. The plates were then incubated for two hours at room temperature, washed and horseradish peroxidase-coupled anti-rabbit IgG (Sigma Chemical So., St. Louis, MO) was added at a 1:2000 dilution in PBS/5% non-fat dried milk. After a further two hour incubation at room temperature, the plates were washed and TMB substrate added. The reaction was stopped after 20 min with 1 N sulfuric acid. Optical density was determined at 450 nm using 570 nm as a reference wavelength. Fractions that resulted in both replicates giving an OD two fold greater than the mean OD from cells cultured in medium alone, plus 3 standard deviations, were considered positive.

For sequencing, the polypeptides were individually dried onto Biobrene™ (Perkin Elmer/Applied BioSystems Division, Foster City, CA) treated glass fiber filters. The filters with polypeptide were loaded onto a Perkin Elmer/Applied BioSystems Division Procise 492 protein sequencer. The polypeptides were sequenced from the amino terminal and using traditional Edman chemistry. The amino acid sequence was determined for each polypeptide by comparing the retention time of the PTH amino acid derivative to the appropriate PTH derivative standards.
Using the procedure described above, antigens having the following N-terminal sequences were isolated:

(a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Xaa-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 54)
(b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 55)
(c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 56)
(d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 57)
(e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 58)
(f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 59)
(g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Ala-Ala-Ala-Ala-Pro-Pro-Ala; (SEQ ID No. 60) and
(h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 61)

wherein Xaa may be any amino acid.

An additional antigen was isolated employing a microbore HPLC purification step in addition to the procedure described above. Specifically, 20 μl of a fraction comprising a mixture of antigens from the chromatographic purification step previously described, was purified on an Aquapore C18 column (Perkin Elmer/Applied Biosystems Division, Foster City, CA) with a 7 micron pore size, column size 1 mm x 100 mm, in a Perkin Elmer/Applied Biosystems Division Model 172 HPLC. Fractions were eluted from the column with a linear gradient of 1%/minute of acetonitrile (containing 0.05% TFA) in water (0.05% TFA) at a flow rate of 80 μl/minute. The eluent was monitored at 250 nm. The original fraction was separated into 4 major peaks plus other smaller components and a polypeptide was obtained which was shown to
have a molecular weight of 12.054 Kd (by mass spectrometry) and the following N-terminal sequence:

(i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Gln-Thr-Ser-Leu-Leu-Asn-Asn-Leu-Ala-Asp-Pro-Asp-Val-Ser-Phe-Ala-Asp (SEQ ID No. 62).

This polypeptide was shown to induce proliferation and IFN-γ production in PBMC preparations using the assays described above.

Additional soluble antigens were isolated from *M. tuberculosis* culture filtrate as follows. *M. tuberculosis* culture filtrate was prepared as described above.

Following dialysis against Bis-Tris propane buffer, at pH 5.5, fractionation was performed using anion exchange chromatography on a Poros QE column 4.6 x 100 mm (Perseptive Biosystems) equilibrated in Bis-Tris propane buffer pH 5.5. Polypeptides were eluted with a linear 0-1.5 M NaCl gradient in the above buffer system at a flow rate of 10 ml/min. The column eluent was monitored at a wavelength of 214 nm.

The fractions eluting from the ion exchange column were pooled and subjected to reverse phase chromatography using a Poros R2 column 4.6 x 100 mm (Perseptive Biosystems). Polypeptides were eluted from the column with a linear gradient from 0-100% acetonitrile (0.1% TFA) at a flow rate of 5 ml/min. The eluent was monitored at 214 nm.

Fractions containing the eluted polypeptides were lyophilized and resuspended in 80 µl of aqueous 0.1% TFA and further subjected to reverse phase chromatography on a Vydac C4 column 4.6 x 150 mm (Western Analytical, Temecula, CA) with a linear gradient of 0-100% acetonitrile (0.1% TFA) at a flow rate of 2 ml/min. Eluent was monitored at 214 nm.

The fraction with biological activity was separated into one major peak plus other smaller components. Western blot of this peak onto PVDF membrane revealed three major bands of molecular weights 14 Kd, 20 Kd and 26 Kd. These polypeptides were determined to have the following N-terminal sequences, respectively:

(j) Xaa-Asp-Ser-Glu-Lys-Ser-Ala-Thr-Ile-Lys-Val-Thr-Asp-Ala-Ser; (SEQ ID No. 134)
(k) Ala-Gly-Asp-Thr-Xaa-Ile-Tyr-Ile-Val-Gly-Asn-Leu-Thr-Ala-Asp; (SEQ ID No. 135) and
(l) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136), wherein Xaa may be any amino acid.

Using the assays described above, these polypeptides were shown to induce proliferation and IFN-γ production in PBMC preparations. Figs. 1A and B show the results of such assays using PBMC preparations from a first and a second donor, respectively.

DNA sequences that encode the antigens designated as (a), (c), (d) and (g) above were obtained by screening a genomic *M. tuberculosis* library using 32P end labeled degenerate oligonucleotides corresponding to the N-terminal sequence and containing *M. tuberculosis* codon bias. The screen performed using a probe corresponding to antigen (a) above identified a clone having the sequence provided in SEQ ID No. 101. The polypeptide encoded by SEQ ID No. 101 is provided in SEQ ID No. 102. The screen performed using a probe corresponding to antigen (g) above identified a clone having the sequence provided in SEQ ID No. 52. The polypeptide encoded by SEQ ID No. 52 is provided in SEQ ID No. 53. The screen performed using a probe corresponding to antigen (d) above identified a clone having the sequence provided in SEQ ID No. 24, and the screen performed with a probe corresponding to antigen (c) identified a clone having the sequence provided in SEQ ID No. 25.

The above amino acid sequences were compared to known amino acid sequences in the gene bank using the DNA STAR system. The database searched contains some 173,000 proteins and is a combination of the Swiss, PIR databases along with translated protein sequences (Version 87). No significant homologies to the amino acid sequences for antigens (a)-(h) and (l) were detected.

The amino acid sequence for antigen (i) was found to be homologous to a sequence from *M. leprae*. The full length *M. leprae* sequence was amplified from genomic DNA using the sequence obtained from GENBANK. This sequence was then used to screen the *M. tuberculosis* library described below in Example 2 and a full length copy of the *M. tuberculosis* homologue was obtained (SEQ ID No. 99).
The amino acid sequence for antigen (j) was found to be homologous to a known *M. tuberculosis* protein translated from a DNA sequence. To the best of the inventors’ knowledge, this protein has not been previously shown to possess T-cell stimulatory activity. The amino acid sequence for antigen (k) was found to be related to a sequence from *M. leprae*.

In the proliferation and IFN-γ assays described above, using three PPD positive donors, the results for representative antigens provided above are presented in Table 1:

**TABLE 1**

**RESULTS OF PBMC PROLIFERATION AND IFN-γ ASSAYS**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Proliferation</th>
<th>IFN-γ</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a)</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>(c)</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>(d)</td>
<td>++</td>
<td>++</td>
</tr>
<tr>
<td>(g)</td>
<td>+++</td>
<td>+++</td>
</tr>
<tr>
<td>(h)</td>
<td>+++</td>
<td>+++</td>
</tr>
</tbody>
</table>

In Table 1, responses that gave a stimulation index (SI) of between 2 and 4 (compared to cells cultured in medium alone) were scored as +, an SI of 4-8 or 2-4 at a concentration of 1 µg or less was scored as ++ and an SI of greater than 8 was scored as ++++. The antigen of sequence (i) was found to have a high SI (++++) for one donor and lower SI (++ and +) for the two other donors in both proliferation and IFN-γ assays. These results indicate that these antigens are capable of inducing proliferation and/or interferon-γ production.
EXAMPLE 2

USE OF PATIENT SERA TO ISOLATE M. TUBERCULOSIS ANTIGENS

This example illustrates the isolation of antigens from *M. tuberculosis* lysate by screening with serum from *M. tuberculosis*-infected individuals.

Dessicated *M. tuberculosis* H37Ra (Difco Laboratories) was added to a 2% NP40 solution, and alternately homogenized and sonicated three times. The resulting suspension was centrifuged at 13,000 rpm in microfuge tubes and the supernatant put through a 0.2 micron syringe filter. The filtrate was bound to Macro Prep DEAE beads (BioRad, Hercules, CA). The beads were extensively washed with 20 mM Tris pH 7.5 and bound proteins eluted with 1M NaCl. The 1M NaCl elute was dialyzed overnight against 10 mM Tris, pH 7.5. Dialyzed solution was treated with DNase and RNase at 0.05 mg/ml for 30 min. at room temperature and then with α-D-mannnosidase, 0.5 U/mg at pH 4.5 for 3-4 hours at room temperature. After returning to pH 7.5, the material was fractionated via FPLC over a Bio Scale-Q-20 column (BioRad). Fractions were combined into nine pools, concentrated in a Centriprep 10 (Amicon, Beverley, MA) and then screened by Western blot for serological activity using a serum pool from *M. tuberculosis*-infected patients which was not immunoreactive with other antigens of the present invention.

The most reactive fraction was run in SDS-PAGE and transferred to PVDF. A band at approximately 85 Kd was cut out yielding the sequence:

\[(m) \quad \text{Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137), wherein Xaa may be any amino acid.}\]

Comparison of this sequence with those in the gene bank as described above revealed no significant homologies to known sequences.

A DNA sequence that encodes the antigen designated as (m) above was obtained by screening a genomic *M. tuberculosis* Erdman strain library using labeled degenerate oligonucleotides corresponding to the N-terminal sequence of SEQ ID NO:137. A clone was identified having the DNA sequence provided in SEQ ID NO:
This sequence was found to encode the amino acid sequence provided in SEQ ID NO: 204. Comparison of these sequences with those in the genebank revealed some similarity to sequences previously identified in *M. tuberculosis* and *M. bovis*.

**EXAMPLE 3**

**PREPARATION OF DNA SEQUENCES ENCODING *M. TUBERCULOSIS* ANTIGENS**

This example illustrates the preparation of DNA sequences encoding *M. tuberculosis* antigens by screening a *M. tuberculosis* expression library with sera obtained from patients infected with *M. tuberculosis*, or with anti-sera raised against soluble *M. tuberculosis* antigens.

**A. PREPARATION OF *M. TUBERCULOSIS* SOLUBLE ANTIGENS USING RABBIT ANTI SERA RAISED AGAINST *M. TUBERCULOSIS* SUPERNATANT**

Genomic DNA was isolated from the *M. tuberculosis* strain H37Ra. The DNA was randomly sheared and used to construct an expression library using the Lambda ZAP expression system (Stratagene, La Jolla, CA). Rabbit anti-sera was generated against secretory proteins of the *M. tuberculosis* strains H37Ra, H37Rv and Erdman by immunizing a rabbit with concentrated supernatant of the *M. tuberculosis* cultures. Specifically, the rabbit was first immunized subcutaneously with 200 μg of protein antigen in a total volume of 2 ml containing 10 μg muramyl dipeptide (Calbiochem, La Jolla, CA) and 1 ml of incomplete Freund's adjuvant. Four weeks later the rabbit was boosted subcutaneously with 100 μg antigen in incomplete Freund's adjuvant. Finally, the rabbit was immunized intravenously four weeks later with 50 μg protein antigen. The anti-sera were used to screen the expression library as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. Bacteriophage plaques expressing immunoreactive antigens were purified. Phagemid from the plaques was rescued and the nucleotide sequences of the *M. tuberculosis* clones deduced.

Thirty two clones were purified. Of these, 25 represent sequences that have not been previously identified in human *M. tuberculosis*. Recombinant antigens
were expressed and purified antigens used in the immunological analysis described in Example 1. Proteins were induced by IPTG and purified by gel elution, as described in Skeiky et al., J. Exp. Med. 181:1527-1537, 1995. Representative sequences of DNA molecules identified in this screen are provided in SEQ ID Nos.: 1-25. The corresponding predicted amino acid sequences are shown in SEQ ID Nos. 63-87.

On comparison of these sequences with known sequences in the gene bank using the databases described above, it was found that the clones referred to hereinafter as TbRA2A, TbRA16, TbRA18, and TbRA29 (SEQ ID Nos. 76, 68, 70, 75) show some homology to sequences previously identified in *Mycobacterium leprae* but not in *M. tuberculosis*. TbRA2A was found to be a lipoprotein, with a six residue lipidation sequence being located adjacent to a hydrophobic secretory sequence. TbRA11, TbRA26, TbRA28 and TbDPEP (SEQ ID Nos.: 65, 73, 74, 53) have been previously identified in *M. tuberculosis*. No significant homologies were found to TbRA1, TbRA3, TbRA4, TbRA9, TbRA10, TbRA13, TbRA17, TbRa19, TbRA29, TbRA32, TbRA36 and the overlapping clones TbRA35 and TbRA12 (SEQ ID Nos. 63, 77, 81, 82, 64, 67, 69, 71, 75, 78, 80, 79, 66). The clone TbRa24 is overlapping with clone TbRa29.

The results of PBMC proliferation and interferon-γ assays performed on representative recombinant antigens, and using T-cell preparations from several different *M. tuberculosis*-immune patients, are presented in Tables 2 and 3, respectively.
### TABLE 2

**RESULTS OF PBMC PROLIFERATION TO REPRESENTATIVE SOLUBLE ANTIGENS**

<table>
<thead>
<tr>
<th>Antigen</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbRa1</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>++</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>±</td>
<td>-</td>
</tr>
<tr>
<td>TbRa3</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>++</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>±</td>
<td>-</td>
</tr>
<tr>
<td>TbRa9</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>++</td>
<td>++</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa10</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>nt</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>TbRa11</td>
<td>±</td>
<td>±</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>nt</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>±</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa12</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TbRa16</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa24</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa26</td>
<td>-</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa29</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa35</td>
<td>++</td>
<td>nt</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>nt</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>++</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaB</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaC</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaD</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>AAMK</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>±</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>YY</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>+</td>
<td>nt</td>
</tr>
<tr>
<td>DPEP</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>++</td>
<td>±</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>±</td>
<td>nt</td>
</tr>
<tr>
<td>Control</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

nt = not tested
### TABLE 3

**RESULTS OF PBMC INTERFERON-Y PRODUCTION TO REPRESENTATIVE SOLUBLE ANTIGENS**

<table>
<thead>
<tr>
<th>Antigen</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbRa1</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>±</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TbRa3</td>
<td>-</td>
<td>+</td>
<td>++</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TbRa9</td>
<td>++</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>++</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa10</td>
<td>+</td>
<td>-</td>
<td>1</td>
<td>±</td>
<td>1</td>
<td>±</td>
<td>+</td>
<td>nt</td>
<td>1</td>
<td>±</td>
<td>-</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>TbRa11</td>
<td>±</td>
<td>1</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>nt</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>±</td>
<td>nt</td>
<td>-</td>
</tr>
<tr>
<td>TbRa12</td>
<td>-</td>
<td>-</td>
<td>+</td>
<td>+</td>
<td>±</td>
<td>+++</td>
<td>+</td>
<td>±</td>
<td>1</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>TbRa16</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>1</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa24</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa26</td>
<td>++</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa29</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>-</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRa35</td>
<td>++</td>
<td>nt</td>
<td>++</td>
<td>+++</td>
<td>+++</td>
<td>nt</td>
<td>++</td>
<td>1</td>
<td>+++</td>
<td>+++</td>
<td>++</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaB</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>++</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaC</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>TbRaD</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>+</td>
<td>+</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
<td>nt</td>
</tr>
<tr>
<td>AAMK</td>
<td>-</td>
<td>-</td>
<td>±</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>±</td>
<td>nt</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>YY</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>-</td>
<td>-</td>
<td>nt</td>
<td>+</td>
<td>nt</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DPEP</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+++</td>
<td>+</td>
<td>-</td>
<td>nt</td>
<td>+++</td>
<td>±</td>
<td>+</td>
<td>±</td>
<td>±</td>
<td>nt</td>
</tr>
<tr>
<td>Control</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
In Tables 2 and 3, responses that gave a stimulation index (SI) of between 1.2 and 2 (compared to cells cultured in medium alone) were scored as ±, a SI of 2-4 was scored as +, as SI of 4-8 or 2-4 at a concentration of 1 µg or less was scored as ++ and an SI of greater than 8 was scored as ++++. In addition, the effect of concentration on proliferation and interferon-γ production is shown for two of the above antigens in the attached Figure. For both proliferation and interferon-γ production, TbRa3 was scored as ++ and TbRa9 as +.

These results indicate that these soluble antigens can induce proliferation and/or interferon-γ production in T-cells derived from an *M. tuberculosis*-immune individual.

**B. USE OF SERA FROM PATIENTS HAVING PULMONARY OR PLEURAL TUBERCULOSIS TO IDENTIFY DNA SEQUENCES ENCODING M. TUBERCULOSIS ANTIGENS**

The genomic DNA library described above, and an additional H37Rv library, were screened using pools of sera obtained from patients with active tuberculosis. To prepare the H37Rv library, *M. tuberculosis* strain H37Rv genomic DNA was isolated, subjected to partial *Sau3A* digestion and used to construct an expression library using the Lambda Zap expression system (Stratagene, La Jolla, Ca). Three different pools of sera, each containing sera obtained from three individuals with active pulmonary or pleural disease, were used in the expression screening. The pools were designated TbL, TbM and TbH, referring to relative reactivity with H37Ra lysate (*i.e.*, TbL = low reactivity, TbM = medium reactivity and TbH = high reactivity) in both ELISA and immunoblot format. A fourth pool of sera from seven patients with active pulmonary tuberculosis was also employed. All of the sera lacked increased reactivity with the recombinant 38 kD *M. tuberculosis* H37Ra phosphate-binding protein.

All pools were pre-adsorbed with *E. coli* lysate and used to screen the H37Ra and H37Rv expression libraries, as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. Bacteriophage plaques expressing immunoreactive antigens were purified. Phagemid from the plaques was rescued and the nucleotide sequences of the *M. tuberculosis* clones deduced.
Thirty two clones were purified. Of these, 31 represented sequences that had not been previously identified in human *M. tuberculosis*. Representative sequences of the DNA molecules identified are provided in SEQ ID Nos.: 26-51 and 105. Of these, TbH-8-2 (SEQ. ID NO. 105) is a partial clone of TbH-8, and TbH-4 (SEQ. ID NO. 43) and TbH-4-FWD (SEQ. ID NO. 44) are non-contiguous sequences from the same clone. Amino acid sequences for the antigens hereinafter identified as Tb38-1, TbH-4, TbH-8, TbH-9, and TbH-12 are shown in SEQ ID Nos.: 88-92. Comparison of these sequences with known sequences in the gene bank using the databases identified above revealed no significant homologies to TbH-4, TbH-8, TbH-9 and TbM-3, although weak homologies were found to TbH-9. TbH-12 was found to be homologous to a 34 kD antigenic protein previously identified in *M. paratuberculosis* (Acc. No. S28515). Tb38-1 was found to be located 34 base pairs upstream of the open reading frame for the antigen ESAT-6 previously identified in *M. bovis* (Acc. No. U34848) and in *M. tuberculosis* (Sorensen et al., *Infect. Immun.* 63:1710-1717, 1995).

Probes derived from Tb38-1 and TbH-9, both isolated from an H37Ra library, were used to identify clones in an H37Rv library. Tb38-1 hybridized to Tb38-1F2, Tb38-1F3, Tb38-1F5 and Tb38-1F6 (SEQ. ID NOS. 112, 113, 116, 118, and 119). (SEQ ID NOS. 112 and 113 are non-contiguous sequences from clone Tb38-1F2.) Two open reading frames were deduced in Tb38-1F2; one corresponds to Tb37FL (SEQ. ID. NO. 114), the second, a partial sequence, may be the homologue of Tb38-1 and is called Tb38-IN (SEQ. ID. NO. 115). The deduced amino acid sequence of Tb38-1F3 is presented in SEQ. ID. NO. 117. A TbH-9 probe identified three clones in the H37Rv library: TbH-9-FL (SEQ. ID. NO. 106), which may be the homologue of TbH-9 (R37Ra), TbH-9-1 (SEQ. ID. NO. 108), and TbH-9-4 (SEQ. ID. NO. 110), all of which are highly related sequences to TbH-9. The deduced amino acid sequences for these three clones are presented in SEQ ID NOS. 107, 109 and 111.

Further screening of the *M. tuberculosis* genomic DNA library, as described above, resulted in the recovery of ten additional reactive clones, representing seven different genes. One of these genes was identified as the 38 Kd antigen discussed
above, one was determined to be identical to the 14Kd alpha crystallin heat shock protein previously shown to be present in *M. tuberculosis*, and a third was determined to be identical to the antigen TbH-8 described above. The determined DNA sequences for the remaining five clones (hereinafter referred to as TbH-29, TbH-30, TbH-32 and TbH-33) are provided in SEQ ID NO: 138-141, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 142-145, respectively. The DNA and amino acid sequences for these antigens were compared with those in the gene bank as described above. No homologies were found to the 5' end of TbH-29 (which contains the reactive open reading frame), although the 3' end of TbH-29 was found to be identical to the *M. tuberculosis* cosmid Y227. TbH-32 and TbH-33 were found to be identical to the previously identified *M. tuberculosis* insertion element IS6110 and to the *M. tuberculosis* cosmid Y50, respectively. No significant homologies to TbH-30 were found.

Positive phagemid from this additional screening were used to infect *E. coli* XL-1 Blue MRF', as described in Sambrook et al., supra. Induction of recombinant protein was accomplished by the addition of IPTG. Induced and uninduced lysates were run in duplicate on SDS-PAGE and transferred to nitrocellulose filters. Filters were reacted with human *M. tuberculosis* sera (1:200 dilution) reactive with TbH and a rabbit sera (1:200 or 1:250 dilution) reactive with the N-terminal 4 Kd portion of lacZ. Sera incubations were performed for 2 hours at room temperature. Bound antibody was detected by addition of 125I-labeled Protein A and subsequent exposure to film for variable times ranging from 16 hours to 11 days. The results of the immunoblots are summarized in Table 4.
TABLE 4

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Human M. tb Sera</th>
<th>Anti-lacZ Sera</th>
</tr>
</thead>
<tbody>
<tr>
<td>TbH-29</td>
<td>45 Kd</td>
<td>45 Kd</td>
</tr>
<tr>
<td>TbH-30</td>
<td>No reactivity</td>
<td>29 Kd</td>
</tr>
<tr>
<td>TbH-32</td>
<td>12 Kd</td>
<td>12 Kd</td>
</tr>
<tr>
<td>TbH-33</td>
<td>16 Kd</td>
<td>16 Kd</td>
</tr>
</tbody>
</table>

Positive reaction of the recombinant human *M. tuberculosis* antigens with both the human *M. tuberculosis* sera and anti-lacZ sera indicate that reactivity of the human *M. tuberculosis* sera is directed towards the fusion protein. Antigens reactive with the anti-lacZ sera but not with the human *M. tuberculosis* sera may be the result of the human *M. tuberculosis* sera recognizing conformational epitopes, or the antigen-antibody binding kinetics may be such that the 2 hour sera exposure in the immunoblot is not sufficient.

The results of T-cell assays performed on Tb38-1, ESAT-6 and other representative recombinant antigens are presented in Tables 5A, B and 6, respectively, below:

TABLE 5A

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Donor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Tb38.1</td>
<td>+</td>
</tr>
<tr>
<td>ESAT-6</td>
<td>+</td>
</tr>
<tr>
<td>TbH-9</td>
<td>-</td>
</tr>
</tbody>
</table>
TABLE 5B

RESULTS OF PBMC INTERFERON-γ PRODUCTION TO REPRESENTATIVE ANTIGENS

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Donor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9 10 11</td>
</tr>
<tr>
<td>Tb38-1</td>
<td>+++ ++ - + - ++ - ++ - +++</td>
</tr>
<tr>
<td>ESAT-6</td>
<td>+++ + + ++ - + + + + +++</td>
</tr>
<tr>
<td>TbH-9</td>
<td>++ ++ - +++ ± ± +++ +++ + +</td>
</tr>
</tbody>
</table>

TABLE 6

SUMMARY OF T-CELL RESPONSES TO REPRESENTATIVE ANTIGENS

<table>
<thead>
<tr>
<th>Antigen</th>
<th>Proliferation</th>
<th>Interferon-γ</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>patient 4 5 6</td>
<td>patient 4 5 6</td>
<td></td>
</tr>
<tr>
<td>TbH9</td>
<td>++ ++ ++</td>
<td>+++ ++ ++</td>
<td>13</td>
</tr>
<tr>
<td>TbM7</td>
<td>- + -</td>
<td>++ + -</td>
<td>4</td>
</tr>
<tr>
<td>TbH5</td>
<td>- + +</td>
<td>++ ++ ++</td>
<td>8</td>
</tr>
<tr>
<td>TbL23</td>
<td>- + ±</td>
<td>++ ++ +</td>
<td>7.5</td>
</tr>
<tr>
<td>TbH4</td>
<td>- ++ ±</td>
<td>++ ++ ±</td>
<td>7</td>
</tr>
<tr>
<td>control</td>
<td>- - -</td>
<td>- - -</td>
<td>0</td>
</tr>
</tbody>
</table>

These results indicate that both the inventive *M. tuberculosis* antigens and ESAT-6 can induce proliferation and/or interferon-γ production in T-cells derived from an *M. tuberculosis*-immune individual. To the best of the inventors' knowledge, ESAT-6 has not been previously shown to stimulate human immune responses.

A set of six overlapping peptides covering the amino acid sequence of the antigen Tb38-1 was constructed using the method described in Example 6. The sequences of these peptides, hereinafter referred to as pep1-6, are provided in SEQ ID Nos. 93-98, respectively. The results of T-cell assays using these peptides are shown in Tables 7 and 8. These results confirm the existence, and help to localize T-cell epitopes within Tb38-1 capable of inducing proliferation and interferon-γ production in T-cells derived from an *M. tuberculosis* immune individual.
<table>
<thead>
<tr>
<th>Peptide</th>
<th>Patient</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>pep1</td>
<td>-</td>
</tr>
<tr>
<td>pep2</td>
<td>±</td>
</tr>
<tr>
<td>pep3</td>
<td>-</td>
</tr>
<tr>
<td>pep4</td>
<td>++</td>
</tr>
<tr>
<td>pep5</td>
<td>++</td>
</tr>
<tr>
<td>pep6</td>
<td>-</td>
</tr>
<tr>
<td>Control</td>
<td>-</td>
</tr>
<tr>
<td>Peptide</td>
<td>1</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
</tr>
<tr>
<td>pep1</td>
<td>+</td>
</tr>
<tr>
<td>pep2</td>
<td>-</td>
</tr>
<tr>
<td>pep3</td>
<td>-</td>
</tr>
<tr>
<td>pep4</td>
<td>-</td>
</tr>
<tr>
<td>pep5</td>
<td>±</td>
</tr>
<tr>
<td>pep6</td>
<td>+</td>
</tr>
<tr>
<td>Control</td>
<td>-</td>
</tr>
</tbody>
</table>

**TABLE 8**

RESULTS OF PBMC INTERFERON-γ PRODUCTION TO 1b38-1 PEPTIDES
Studies were undertaken to determine whether the antigens TbH-9 and Tb38-1 represent cellular proteins or are secreted into *M. tuberculosis* culture media. In the first study, rabbit sera were raised against A) secretory proteins of *M. tuberculosis*, B) the known secretory recombinant *M. tuberculosis* antigen 85b, C) recombinant Tb38-1 and D) recombinant TbH-9, using protocols substantially the same as that as described in Example 3A. Total *M. tuberculosis* lysate, concentrated supernatant of *M. tuberculosis* cultures and the recombinant antigens 85b, TbH-9 and Tb38-1 were resolved on denaturing gels, immobilized on nitrocellulose membranes and duplicate blots were probed using the rabbit sera described above.

The results of this analysis using control sera (panel I) and antisera (panel II) against secretory proteins, recombinant 85b, recombinant Tb38-1 and recombinant TbH-9 are shown in Figures 3A-D, respectively, wherein the lane designations are as follows: 1) molecular weight protein standards; 2) 5 µg of *M. tuberculosis* lysate; 3) 5 µg secretory proteins; 4) 50 ng recombinant Tb38-1; 5) 50 ng recombinant TbH-9; and 6) 50 ng recombinant 85b. The recombinant antigens were engineered with six terminal histidine residues and would therefore be expected to migrate with a mobility approximately 1 kDa larger that the native protein. In Figure 3D, recombinant TbH-9 is lacking approximately 10 kDa of the full-length 42 kDa antigen, hence the significant difference in the size of the immunoreactive native TbH-9 antigen in the lysate lane (indicated by an arrow). These results demonstrate that Tb38-1 and TbH-9 are intracellular antigens and are not actively secreted by *M. tuberculosis*.

The finding that TbH-9 is an intracellular antigen was confirmed by determining the reactivity of TbH-9-specific human T cell clones to recombinant TbH-9, secretory *M. tuberculosis* proteins and PPD. A TbH-9-specific T cell clone (designated 131TbH-9) was generated from PBMC of a healthy PPD-positive donor. The proliferative response of 131TbH-9 to secretory proteins, recombinant TbH-9 and a control *M. tuberculosis* antigen, TbRa11, was determined by measuring uptake of tritiated thymidine, as described in Example 1. As shown in Figure 4A, the clone 131TbH-9 responds specifically to TbH-9, showing that TbH-9 is not a significant component of *M. tuberculosis* secretory proteins. Figure 4B shows the production of IFN-γ by a second TbH-9-specific T cell clone.
(designated PPD 800-10) prepared from PBMC from a healthy PPD-positive donor,
following stimulation of the T cell clone with secretory proteins, PPD or recombinant TbH-9.
These results further confirm that TbH-9 is not secreted by *M. tuberculosis*.

C. **USE OF SERA FROM PATIENTS HAVING EXTRAPULMONARY TUBERCULOSIS TO IDENTIFY DNA SEQUENCES ENCODING M. TUBERCULOSIS ANTIGENS**

Genomic DNA was isolated from *M. tuberculosis* Erdman strain, randomly sheared and used to construct an expression library employing the Lambda ZAP expression system (Stratagene, La Jolla, CA). The resulting library was screened using pools of sera obtained from individuals with extrapulmonary tuberculosis, as described above in Example 3B, with the secondary antibody being goat anti-human IgG + A + M (H+L) conjugated with alkaline phosphatase.

Eighteen clones were purified. Of these, 4 clones (hereinafter referred to as XP14, XP24, XP31 and XP32) were found to bear some similarity to known sequences. The determined DNA sequences for XP14, XP24 and XP31 are provided in SEQ ID Nos.: 156-158, respectively, with the 5' and 3' DNA sequences for XP32 being provided in SEQ ID Nos.: 159 and 160, respectively. The predicted amino acid sequence for XP14 is provided in SEQ ID No: 161. The reverse complement of XP14 was found to encode the amino acid sequence provided in SEQ ID No.: 162.

Comparison of the sequences for the remaining 14 clones (hereinafter referred to as XP1-XP6, XP17-XP19, XP22, XP25, XP27, XP30 and XP36) with those in the genebank as described above, revealed no homologies with the exception of the 3' ends of XP2 and XP6 which were found to bear some homology to known *M. tuberculosis* cosmids. The DNA sequences for XP27 and XP36 are shown in SEQ ID Nos.: 163 and 164, respectively, with the 5' sequences for XP4, XP5, XP17 and XP30 being shown in SEQ ID Nos: 165-168, respectively, and the 5' and 3' sequences for XP2, XP3, XP6, XP18, XP19, XP22 and XP25 being shown in SEQ ID Nos: 169 and 170; 171 and 172; 173 and 174; 175 and 176; 177 and 178; 179 and 180; and 181 and 182, respectively. XP1 was found to overlap with the DNA sequences for TbH4, disclosed above. The full-length DNA sequence for TbH4-XP1 is provided in SEQ ID No.: 183. This DNA sequence was found to contain an
open reading frame encoding the amino acid sequence shown in SEQ ID No: 184. The reverse complement of TbH4-XP1 was found to contain an open reading frame encoding the amino acid sequence shown in SEQ ID No.: 185. The DNA sequence for XP36 was found to contain two open reading frames encoding the amino acid sequence shown in SEQ ID Nos.: 186 and 187, with the reverse complement containing an open reading frame encoding the amino acid sequence shown in SEQ ID No.: 188.

Recombinant XP1 protein was prepared as described above in Example 3B, with a metal ion affinity chromatography column being employed for purification. As illustrated in Figures 8A-B and 9A-B, using the assays described herein, recombinant XP1 was found to stimulate cell proliferation and IFN-γ production in T cells isolated from an M. tuberculosis-immune donors.

D. USE OF A LYSATE POSITIVE SERUM POOL FROM PATIENTS HAVING TUBERCULOSIS TO IDENTIFY DNA SEQUENCES ENCODING M. TUBERCULOSIS ANTIGENS

Genomic DNA was isolated from M. tuberculosis Erdman strain, randomly sheared and used to construct an expression library employing the Lambda Screen expression system (Novagen, Madison, WI), as described below in Example 6. Pooled serum obtained from M. tuberculosis-infected patients and that was shown to react with M. tuberculosis lysate but not with the previously expressed proteins 38kD, Tb38-1, TbRa3, TbH4, DPEP and TbRa11, was used to screen the expression library as described above in Example 3B, with the secondary antibody being goat anti-human IgG + A + M (H+L) conjugated with alkaline phosphatase.

Twenty-seven clones were purified. Comparison of the determined cDNA sequences for these clones revealed no significant homologies to 10 of the clones (hereinafter referred to as LSER-10, LSER-11, LSER-12, LSER-13, LSER-16, LSER-18, LSER-23, LSER-24, LSER-25 and LSER-27). The determined 5' cDNA sequences for LSER-10, LSER-11, LSER-12, LSER-13, LSER-16 and LSER-25 are provided in SEQ ID NO: 242-247, respectively, with the corresponding predicted amino acid sequences for LSER-10, LSER-12, LSER-13, LSER-16 and LSER-25 being provided in SEQ ID NO: 248-252, respectively. The determined full-length cDNA sequences for LSER-18, LSER-23, LSER-24
and LSER-27 are shown in SEQ ID NO: 253-256, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 257-260. The remaining seventeen clones were found to show similarities to unknown sequences previously identified in *M. tuberculosis*. The determined 5' cDNA sequences for sixteen of these clones (hereinafter referred to as LSER-1, LSER-3, LSER-4, LSER-5, LSER-6, LSER-8, LSER-14, LSER-15, LSER-17, LSER-19, LSER-20, LSER-22, LSER-26, LSER-28, LSER-29 and LSER-30) are provided in SEQ ID NO: 261-276, respectively, with the corresponding predicted amino acid sequences for LSER-1, LSER-3, LSER-5, LSER-6, LSER-8, LSER-14, LSER-15, LSER-17, LSER-19, LSER-20, LSER-22, LSER-26, LSER-28, LSER-29 and LSER-30 being provided in SEQ ID NO: 277-291, respectively. The determined full-length cDNA sequence for the clone LSER-9 is provided in SEQ ID NO: 292. The reverse complement of LSER-6 (SEQ ID NO: 293) was found to encode the predicted amino acid sequence of SEQ ID NO: 294.

E. PREPARATION OF *M. TUBERCULOSIS* SOLUBLE ANTIGENS USING RABBIT ANTI-ERA RAISED AGAINST *M. TUBERCULOSIS* FRACTIONATED PROTEINS

*M. tuberculosis* lysate was prepared as described above in Example 2. The resulting material was fractionated by HPLC and the fractions screened by Western blot for serological activity with a serum pool from *M. tuberculosis*-infected patients which showed little or no immunoreactivity with other antigens of the present invention. Rabbit anti-era was generated against the most reactive fraction using the method described in Example 3A. The anti- sera was used to screen an *M. tuberculosis* Erdman strain genomic DNA expression library prepared as described above. Bacteriophage plaques expressing immunoreactive antigens were purified. Phagelmid from the plaques was rescued and the nucleotide sequences of the *M. tuberculosis* clones determined.

Ten different clones were purified. Of these, one was found to be TbRa35, described above, and one was found to be the previously identified *M. tuberculosis* antigen, HSP60. Of the remaining eight clones, seven (hereinafter referred to as RDIF2, RDIF3, RDIF8, RDIF10, RDIF11 and RDIF12) were found to bear some similarity to previously identified *M. tuberculosis* sequences. The determined DNA sequences for RDIF2, RDIF5, RDIF8, RDIF10 and RDIF11 are provided in SEQ ID Nos.: 189-193, respectively, with the
corresponding predicted amino acid sequences being provided in SEQ ID Nos: 194-198, respectively. The 5' and 3' DNA sequences for RDIF12 are provided in SEQ ID Nos.: 199 and 200, respectively. No significant homologies were found to the antigen RDIF-7. The determined DNA and predicted amino acid sequences for RDIF7 are provided in SEQ ID Nos.: 201 and 202, respectively. One additional clone, referred to as RDIF6 was isolated, however, this was found to be identical to RDIF5.

Recombinant RDIF6, RDIF8, RDIF10 and RDIF11 were prepared as described above. As shown in Figures 8A-B and 9A-B, these antigens were found to stimulate cell proliferation and IFN-γ production in T cells isolated from M. tuberculosis-immune donors.

EXAMPLE 4

PURIFICATION AND CHARACTERIZATION OF A POLYPEPTIDE FROM TUBERCULIN PURIFIED PROTEIN DERIVATIVE

An M. tuberculosis polypeptide was isolated from tuberculin purified protein derivative (PPD) as follows.

PPD was prepared as published with some modification (Seibert, F. et al., Tuberculin purified protein derivative. Preparation and analyses of a large quantity for standard. The American Review of Tuberculosis 44:9-25, 1941).

M. tuberculosis Rv strain was grown for 6 weeks in synthetic medium in roller bottles at 37°C. Bottles containing the bacterial growth were then heated to 100°C in water vapor for 3 hours. Cultures were sterile filtered using a 0.22 μ filter and the liquid phase was concentrated 20 times using a 3 kD cut-off membrane. Proteins were precipitated once with 50% ammonium sulfate solution and eight times with 25% ammonium sulfate solution. The resulting proteins (PPD) were fractionated by reverse phase liquid chromatography (RP-HPLC) using a C18 column (7.8 x 300 mM; Waters, Milford, MA) in a Biocad HPLC system (Perseptive Biosystems, Framingham, MA). Fractions were eluted from the column with a linear gradient from 0-100% buffer (0.1% TFA in acetonitrile). The flow rate was 10 ml/minute and eluent was monitored at 214 nm and 280 nm.
Six fractions were collected, dried, suspended in PBS and tested individually in *M. tuberculosis*-infected guinea pigs for induction of delayed type hypersensitivity (DTH) reaction. One fraction was found to induce a strong DTH reaction and was subsequently fractionated further by RP-HPLC on a microbore Vydac C18 column (Cat. No. 218TP5115) in a Perkin Elmer/Applied Biosystems Division Model 172 HPLC. Fractions were eluted with a linear gradient from 5-100% buffer (0.05% TFA in acetonitrile) with a flow rate of 80 μl/minute. Eluent was monitored at 215 nm. Eight fractions were collected and tested for induction of DTH in *M. tuberculosis*-infected guinea pigs. One fraction was found to induce strong DTH of about 16 mm induration. The other fractions did not induce detectable DTH. The positive fraction was submitted to SDS-PAGE gel electrophoresis and found to contain a single protein band of approximately 12 kD molecular weight.

This polypeptide, herein after referred to as DPPD, was sequenced from the amino terminal using a Perkin Elmer/Applied Biosystems Division Procise 492 protein sequencer as described above and found to have the N-terminal sequence shown in SEQ ID No.: 129. Comparison of this sequence with known sequences in the gene bank as described above revealed no known homologies. Four cyanogen bromide fragments of DPPD were isolated and found to have the sequences shown in SEQ ID Nos.: 130-133. A subsequent search of the *M. tuberculosis* genome database released by the Institute for Genomic Research revealed a match of the DPPD partial amino acid sequence with a sequence present within the *M. tuberculosis* cosmid MTY21C12. An open reading frame of 336 bp was identified. The full-length DNA sequence for DPPD is provided in SEQ ID NO: 240, with the corresponding full-length amino acid sequence being provided in SEQ ID NO: 241.

The ability of the antigen DPPD to stimulate human PBMC to proliferate and to produce IFN-γ was assayed as described in Example 1. As shown in Table 9, DPPD was found to stimulate proliferation and elicit production of large quantities of IFN-γ; more than that elicited by commercial PPD.
TABLE 9

RESULTS OF PROLIFERATION AND INTERFERON-γ ASSAYS TO DPPD

<table>
<thead>
<tr>
<th>PBMC Donor</th>
<th>Stimulator</th>
<th>Proliferation (CPM)</th>
<th>IFN-γ (OD_{580})</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Medium</td>
<td>1,089</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>PPD (commercial)</td>
<td>8,394</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>DPPD</td>
<td>13,451</td>
<td>2.21</td>
</tr>
<tr>
<td>B</td>
<td>Medium</td>
<td>450</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>PPD (commercial)</td>
<td>3,929</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>DPPD</td>
<td>6,184</td>
<td>1.49</td>
</tr>
<tr>
<td>C</td>
<td>Medium</td>
<td>541</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>PPD (commercial)</td>
<td>8,907</td>
<td>0.76</td>
</tr>
<tr>
<td></td>
<td>DPPD</td>
<td>23,024</td>
<td>&gt;2.70</td>
</tr>
</tbody>
</table>

EXAMPLE 5

USE OF SERA FROM TUBERCULOSIS-INFECTED MONKEYS TO IDENTIFY DNA SEQUENCES ENCODING M. TUBERCULOSIS ANTIGENS

Genomic DNA was isolated from M. tuberculosis Erdman strain, randomly sheared and used to construct an expression library employing the Lambda ZAP expression system (Stratagene, La Jolla, CA). Serum samples were obtained from a cynomolgous monkey 18, 33, 51 and 56 days following infection with M. tuberculosis Erdman strain. These samples were pooled and used to screen the M. tuberculosis genomic DNA expression library using the procedure described above in Example 3C.

Twenty clones were purified. The determined 5' DNA sequences for the clones referred to as MO-1, MO-2, MO-4, MO-8, MO-9, MO-26, MO-28, MO-29, MO-30, MO-34 and MO-35 are provided in SEQ ID NO: 215-225, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 226-236. The full-length DNA sequence of the clone MO-10 is provided in SEQ ID NO: 237, with the corresponding
predicted amino acid sequence being provided in SEQ ID NO: 238. The 3’ DNA sequence for the clone MO-27 is provided in SEQ ID NO: 239.

Clones MO-1, MO-30 and MO-35 were found to show a high degree of relatedness and showed some homology to a previously identified unknown *M. tuberculosis* sequence and to cosmid MTC1237. MO-2 was found to show some homology to aspartokinase from *M. tuberculosis*. Clones MO-3, MO-7 and MO-27 were found to be identical and to show a high degree of relatedness to MO-5. All four of these clones showed some homology to *M. tuberculosis* heat shock protein 70. MO-27 was found to show some homology to *M. tuberculosis* cosmid MTCY339. MO-4 and MO-34 were found to show some homology to cosmid SCY21B4 and *M. smegmatis* integration host factor, and were both found to show some homology to a previously identified, unknown *M. tuberculosis* sequence. MO-6 was found to show some homology to *M. tuberculosis* heat shock protein 65. MO-8, MO-9, MO-10, MO-26 and MO-29 were found to be highly related to each other and to show some homology to *M. tuberculosis* dihydrolipamide succinyltransferase. MO-28, MO-31 and MO-32 were found to be identical and to show some homology to a previously identified *M. tuberculosis* protein. MO-33 was found to show some homology to a previously identified 14 kDa *M. tuberculosis* heat shock protein.

Further studies using the above protocol resulted in the isolation of an additional four clones, hereinafter referred to as MO-12, MO-13, MO-19 and MO-39. The determined 5’ cDNA sequences for these clones are provided in SEQ ID NO: 295-298, respectively, with the corresponding predicted protein sequences being provided in SEQ ID NO: 299-302, respectively. Comparison of these sequences with those in the gene bank as described above revealed no significant homologies to MO-39. MO-12, MO-13 and MO-19 were found to show some homologies to unknown sequences previously isolated from *M. tuberculosis*.

**EXAMPLE 6**

**ISOLATION OF DNA SEQUENCES ENCODING *M. TUBERCULOSIS* ANTIGENS**

**BY SCREENING OF A NOVEL EXPRESSION LIBRARY**
This example illustrates isolation of DNA sequences encoding *M. tuberculosis* antigens by screening of a novel expression library with sera from *M. tuberculosis*-infected patients that were shown to be unreactive with a panel of the recombinant *M. tuberculosis* antigens TbRa11, TbRa3, Tb38-1, TbH4, TbF and 38 kD.

Genomic DNA from *M. tuberculosis* Erdman strain was randomly sheared to an average size of 2 kb, and blunt ended with Klenow polymerase, followed by the addition of EcoRI adaptors. The insert was subsequently ligated into the Screen phage vector (Novagen, Madison, WI) and packaged *in vitro* using the PhageMaker extract (Novagen). The resulting library was screened with sera from several *M. tuberculosis* donors that had been shown to be negative on a panel of previously identified *M. tuberculosis* antigens as described above in Example 3B.

A total of 22 different clones were isolated. By comparison, screening of the λ Zap library described above using the same sera did not result in any positive hits. One of the clones was found to represent TbRa11, described above. The determined 5' cDNA sequences for 19 of the remaining 21 clones (hereinafter referred to as Erdsn1, Erdsn2, Erdsn4-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25) are provided in SEQ ID NO: 303-322, respectively, with the determined 3' cDNA sequences for Erdsn1, Erdsn2, Erdsn4, Erdsn5, Erdsn7-Erdsn10, Erdsn12-18, Erdsn21-Erdsn23 and Erdsn25 being provided in SEQ ID NO: 323-341, respectively. The complete cDNA insert sequence for the clone Erdsn24 is provided in SEQ ID NO: 342. Comparison of the determined cDNA sequences with those in the gene bank revealed no significant homologies to the sequences provided in SEQ ID NO: 309, 316, 318-320, 322, 324, 328, 329, 333, 335, 337, 339 and 341. The sequences of SEQ ID NO: 303-308, 310-315, 317, 321, 323, 325-327, 330-332, 334, 336, 338, 340 and 342 were found to show some homology to unknown sequences previously identified in *M. tuberculosis*.

**EXAMPLE 7**

**ISOLATION OF SOLUBLE *M. TUBERCULOSIS* ANTIGENS USING MASS SPECTROMETRY**
This example illustrates the use of mass spectrometry to identify soluble \textit{M. tuberculosis} antigens.

In a first approach, \textit{M. tuberculosis} culture filtrate was screened by Western analysis using serum from a tuberculosis-infected individual. The reactive bands were excised from a silver stained gel and the amino acid sequences determined by mass spectrometry. The determined amino acid sequence for one of the isolated antigens is provided in SEQ ID NO: 343. Comparison of this sequence with those in the gene bank revealed homology to the 85b precursor antigen previously identified in \textit{M. tuberculosis}.

In a second approach, the high molecular weight region of \textit{M. tuberculosis} culture supernatant was studied. This area may contain immunodominant antigens which may be useful in the diagnosis of \textit{M. tuberculosis} infection. Two known monoclonal antibodies, IT42 and IT57 (available from the Center for Disease Control, Atlanta, GA), show reactivity by Western analysis to antigens in this vicinity, although the identity of the antigens remains unknown. In addition, unknown high-molecular weight proteins have been described as containing a surrogate marker for \textit{M. tuberculosis} infection in HIV-positive individuals (\textit{Jnl. Infect. Dis.}, 176:133-143, 1997). To determine the identity of these antigens, two-dimensional gel electrophoresis and two-dimensional Western analysis were performed using the antibodies IT57 and IT42. Five protein spots in the high molecular weight region were identified, individually excised, enzymatically digested and subjected to mass spectrometric analysis.

The determined amino acid sequences for three of these spots (referred to as spots 1, 2 and 4) are provided in SEQ ID NO: 344, 345-346 and 347, respectively. Comparison of these sequences with those in the gene bank revealed that spot 1 is the previously identified PcK-1, a phosphoenolpyruvate kinase. The two sequences isolated from spot 2 were determined to be from two DNAks, previously identified in \textit{M. tuberculosis} as heat shock proteins. Spot 4 was determined to be the previously identified \textit{M. tuberculosis} protein Kat G. To the best of the inventors’ knowledge, neither PcK-1 nor the two DNAks have previously been shown to have utility in the diagnosis of \textit{M. tuberculosis} infection.

\textbf{EXAMPLE 8}
USE OF REPRESENTATIVE ANTIGENS FOR DIAGNOSIS OF TUBERCULOSIS

This example illustrates the effectiveness of several representative polypeptides in skin tests for the diagnosis of *M. tuberculosis* infection.

Individuals were injected intradermally with 100 μl of either PBS or PBS plus Tween 20™ containing either 0.1 μg of protein (for TbH-9 and TbRa35) or 1.0 μg of protein (for TbRa38-1). Induration was measured between 5-7 days after injection, with a response of 5 mm or greater being considered positive. Of the 20 individuals tested, 2 were PPD negative and 18 were PPD positive. Of the PPD positive individuals, 3 had active tuberculosis, 3 had been previously infected with tuberculosis and 9 were healthy. In a second study, 13 PPD positive individuals were tested with 0.1 μg TbRa11 in either PBS or PBS plus Tween 20™ as described above. The results of both studies are shown in Table 10.

### TABLE 10
RESULTS OF DTH TESTING WITH REPRESENTATIVE ANTIGENS

<table>
<thead>
<tr>
<th></th>
<th>TbH-9 Pos/Total</th>
<th>Tb38-1 Pos/Total</th>
<th>TbRa35 Pos/Total</th>
<th>Cumulative Pos/Total</th>
<th>TbRa11 Pos/Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPD negative</td>
<td>0/2</td>
<td>0/2</td>
<td>0/2</td>
<td>0/2</td>
<td></td>
</tr>
<tr>
<td>PPD positive</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>healthy</td>
<td>5/9</td>
<td>4/9</td>
<td>4/9</td>
<td>6/9</td>
<td>1/4</td>
</tr>
<tr>
<td>prior TB</td>
<td>3/5</td>
<td>2/5</td>
<td>2/5</td>
<td>4/5</td>
<td>3/5</td>
</tr>
<tr>
<td>active</td>
<td>3/4</td>
<td>3/4</td>
<td>0/4</td>
<td>4/4</td>
<td>1/4</td>
</tr>
<tr>
<td>TOTAL</td>
<td>11/18</td>
<td>9/18</td>
<td>6/18</td>
<td>14/18</td>
<td>5/13</td>
</tr>
</tbody>
</table>

EXAMPLE 9
SYNTHESIS OF SYNTHETIC POLYPEPTIDES

Polypeptides may be synthesized on a Millipore 9050 peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium
hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanediethiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray mass spectrometry and by amino acid analysis.

**EXAMPLE 10**

**PREPARATION AND CHARACTERIZATION OF *M. TUBERCULOSIS* FUSION PROTEINS**

A fusion protein containing TbRa3, the 38 kD antigen and Tb38-1 was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR in order to facilitate their fusion and the subsequent expression of the fusion protein TbRa3-38 kD-Tb38-1. TbRa3, 38 kD and Tb38-1 DNA was used to perform PCR using the primers PDM-64 and PDM-65 (SEQ ID NO: 146 and 147), PDM-57 and PDM-58 (SEQ ID NO: 148 and 149), and PDM-69 and PDM-60 (SEQ ID NO: 150 and 151), respectively. In each case, the DNA amplification was performed using 10 μl 10X Pfu buffer, 2 μl 10 mM dNTPs, 2 μl each of the PCR primers at 10 μM concentration, 81.5 μl water, 1.5 μl Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 μl DNA at either 70 ng/μl (for TbRa3) or 50 ng/μl (for 38 kD and Tb38-1). For TbRa3, denaturation at 94°C was performed for 2 min, followed by 40 cycles of 96°C for 15 sec and 72°C for 1 min, and lastly by 72°C for 4 min. For 38 kD, denaturation at 96°C was performed for 2 min, followed by 40 cycles of 96°C for 30 sec, 68°C for 15 sec and 72°C for 3 min, and finally by 72°C for 4 min. For Tb38-1 denaturation at 94°C for 2 min was followed by 10 cycles of 96°C for 15 sec, 68°C for 15 sec and 72°C for
1.5 min, 30 cycles of 96°C for 15 sec, 64°C for 15 sec and 72°C for 1.5, and finally by 72°C for 4 min.

The TbRa3 PCR fragment was digested with NdeI and EcoRI and cloned directly into pT7^L2 IL 1 vector using NdeI and EcoRI sites. The 38 kD PCR fragment was digested with Sse8387I, treated with T4 DNA polymerase to make blunt ends and then digested with EcoRI for direct cloning into the pT7^L2Ra3-1 vector which was digested with StuI and EcoRI. The 38-1 PCR fragment was digested with Eco47III and EcoRI and directly subcloned into pT7^L2Ra3/38kD-17 digested with the same enzymes. The whole fusion was then transferred to pET28b – using NdeI and EcoRI sites. The fusion construct was confirmed by DNA sequencing.

The expression construct was transformed into BLR pLys S E. coli (Novagen, Madison, WI) and grown overnight in LB broth with kanamycin (30 µg/ml) and chloramphenicol (34 µg/ml). This culture (12 ml) was used to inoculate 500 ml 2XYS with the same antibiotics and the culture was induced with IPTG at an OD560 of 0.44 to a final concentration of 1.2 mM. Four hours post-induction, the bacteria were harvested and sonicated in 20 mM Tris (8.0), 100 mM NaCl, 0.1% DOC, 20 µg/ml Leupeptin, 20 mM PMSF followed by centrifugation at 26,000 X g. The resulting pellet was resuspended in 8 M urea, 20 mM Tris (8.0), 100 mM NaCl and bound to Pro-bond nickel resin (Invitrogen, Carlsbad, CA). The column was washed several times with the above buffer then eluted with an imidazole gradient (50 mM, 100 mM, 500 mM imidazole was added to 8 M urea, 20 mM Tris (8.0), 100 mM NaCl). The eluates containing the protein of interest were then dialyzed against 10 mM Tris (8.0).

The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbRa3-38 kD-Tb38-1) are provided in SEQ ID NO: 152 and 153, respectively.

A fusion protein containing the two antigens TbH-9 and Tb38-1 (hereinafter referred to as TbH9-Tb38-1) without a hinge sequence, was prepared using a similar procedure to that described above. The DNA sequence for the TbH9-Tb38-1 fusion protein is provided in SEQ ID NO: 156.
The ability of the fusion protein TbH9-Tb38-1 to induce T cell proliferation and IFN-γ production in PBMC preparations was examined using the protocol described above in Example 1. PBMC from three donors were employed: one who had been previously shown to respond to TbH9 but not Tb38-1 (donor 131); one who had been shown to respond to Tb38-1 but not TbH9 (donor 184); and one who had been shown to respond to both antigens (donor 201). The results of these studies (Figs. 5-7, respectively) demonstrate the functional activity of both the antigens in the fusion protein.

A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and DPEP was prepared as follows.

Each of the DNA constructs TbRa3, 38 kD and Tb38-1 were modified by PCR and cloned into vectors essentially as described above, with the primers PDM-69 (SEQ ID NO:150 and PDM-83 (SEQ ID NO: 205) being used for amplification of the Tb38-1A fragment. Tb38-1A differs from Tb38-1 by a Dral site at the 3' end of the coding region that keeps the final amino acid intact while creating a blunt restriction site that is in frame. The TbRa3/38kD/Tb38-1A fusion was then transferred to pET28b using NdeI and EcoRI sites.

DPEP DNA was used to perform PCR using the primers PDM-84 and PDM-85 (SEQ ID NO: 206 and 207, respectively) and 1 µl DNA at 50 ng/µl. Denaturation at 94 °C was performed for 2 min, followed by 10 cycles of 96 °C for 15 sec, 68 °C for 15 sec and 72 °C for 1.5 min; 30 cycles of 96 °C for 15 sec, 64 °C for 15 sec and 72 °C for 1.5 min; and finally by 72 °C for 4 min. The DPEP PCR fragment was digested with EcoRI and Eco72I and clones directly into the pET28Ra3/38kD/38-1A construct which was digested with Dral and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-2) are provided in SEQ ID NO: 208 and 209, respectively.

The reactivity of the fusion protein TbF-2 with sera from M. tuberculosis-infected patients was examined by ELISA using the protocol described above. The results of these studies (Table 11) demonstrate that all four antigens function independently in the fusion protein.
# Table 11

**Reactivity of Tbf-2 Fusion Recombinant with TB and Normal Sera**

<table>
<thead>
<tr>
<th>Serum ID</th>
<th>Status</th>
<th>Tbf OD450</th>
<th>Status</th>
<th>Tbf-2 OD450</th>
<th>Status</th>
<th>ELISA Reactivity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38 kD TbRa3 Tb38-1 DPEP</td>
</tr>
<tr>
<td>B931-40</td>
<td>TB</td>
<td>0.57</td>
<td>+</td>
<td>0.321</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>B931-41</td>
<td>TB</td>
<td>0.601</td>
<td>-</td>
<td>0.396</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>B931-109</td>
<td>TB</td>
<td>0.494</td>
<td>+</td>
<td>0.404</td>
<td>+</td>
<td>-</td>
</tr>
<tr>
<td>B931-132</td>
<td>TB</td>
<td>1.502</td>
<td>-</td>
<td>1.292</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>5004</td>
<td>TB</td>
<td>1.806</td>
<td>+</td>
<td>1.666</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>15004</td>
<td>TB</td>
<td>2.862</td>
<td>-</td>
<td>2.468</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>39004</td>
<td>TB</td>
<td>2.443</td>
<td>-</td>
<td>1.722</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>68004</td>
<td>TB</td>
<td>2.871</td>
<td>-</td>
<td>2.575</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>99004</td>
<td>TB</td>
<td>0.691</td>
<td>-</td>
<td>0.971</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>107004</td>
<td>TB</td>
<td>0.875</td>
<td>-</td>
<td>0.732</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>92004</td>
<td>TB</td>
<td>1.932</td>
<td>-</td>
<td>1.394</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>97004</td>
<td>TB</td>
<td>1.491</td>
<td>-</td>
<td>1.979</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>118004</td>
<td>TB</td>
<td>3.182</td>
<td>+</td>
<td>3.045</td>
<td>±</td>
<td>±</td>
</tr>
<tr>
<td>173004</td>
<td>TB</td>
<td>3.644</td>
<td>-</td>
<td>3.578</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>175004</td>
<td>TB</td>
<td>3.322</td>
<td>-</td>
<td>2.916</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>274004</td>
<td>TB</td>
<td>3.696</td>
<td>-</td>
<td>3.716</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>276004</td>
<td>TB</td>
<td>3.245</td>
<td>-</td>
<td>2.56</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>282004</td>
<td>TB</td>
<td>1.249</td>
<td>-</td>
<td>1.254</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>289004</td>
<td>TB</td>
<td>1.373</td>
<td>-</td>
<td>1.17</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>308004</td>
<td>TB</td>
<td>3.708</td>
<td>-</td>
<td>3.355</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>314004</td>
<td>TB</td>
<td>1.665</td>
<td>-</td>
<td>1.399</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>317004</td>
<td>TB</td>
<td>1.165</td>
<td>-</td>
<td>0.92</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>312004</td>
<td>TB</td>
<td>1.709</td>
<td>-</td>
<td>1.453</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>380004</td>
<td>TB</td>
<td>0.238</td>
<td>-</td>
<td>0.461</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>451004</td>
<td>TB</td>
<td>0.18</td>
<td>-</td>
<td>0.2</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>478004</td>
<td>TB</td>
<td>0.188</td>
<td>-</td>
<td>0.469</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>410004</td>
<td>TB</td>
<td>0.384</td>
<td>-</td>
<td>2.392</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>411004</td>
<td>TB</td>
<td>0.306</td>
<td>-</td>
<td>0.874</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>421004</td>
<td>TB</td>
<td>0.357</td>
<td>-</td>
<td>1.456</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>523004</td>
<td>TB</td>
<td>0.047</td>
<td>-</td>
<td>0.196</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-87</td>
<td>Normal</td>
<td>0.094</td>
<td>-</td>
<td>0.063</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-88</td>
<td>Normal</td>
<td>0.214</td>
<td>-</td>
<td>0.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-89</td>
<td>Normal</td>
<td>0.248</td>
<td>-</td>
<td>0.125</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-90</td>
<td>Normal</td>
<td>0.179</td>
<td>-</td>
<td>0.206</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-91</td>
<td>Normal</td>
<td>0.135</td>
<td>-</td>
<td>0.151</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-92</td>
<td>Normal</td>
<td>0.064</td>
<td>-</td>
<td>0.097</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-93</td>
<td>Normal</td>
<td>0.072</td>
<td>-</td>
<td>0.098</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-94</td>
<td>Normal</td>
<td>0.072</td>
<td>-</td>
<td>0.064</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-95</td>
<td>Normal</td>
<td>0.125</td>
<td>-</td>
<td>0.159</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>A6-96</td>
<td>Normal</td>
<td>0.121</td>
<td>-</td>
<td>0.12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Cut-off</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.284</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.266</td>
</tr>
</tbody>
</table>
A fusion protein containing TbRa3, the antigen 38kD, Tb38-1 and TbH4 was prepared as follows.

Genomic *M. tuberculosis* DNA was used to PCR full-length TbH4 (FL TbH4) with the primers PDM-157 and PDM-160 (SEQ ID NO: 348 and 349, respectively) and 2 µl DNA at 100 ng/µl. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 61 °C for 20 sec and 72 °C for 5 min; and finally by annealing at 72 °C for 10 min. The FL TbH4 PCR fragment was digested with EcoRI and Sca I (New England Biolabs.) and cloned directly into the pET28Ra3/38kD/38-1A construct described above which was digested with Dral and EcoRI. The fusion construct was confirmed to be correct by DNA sequencing. Recombinant protein was prepared as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-6) are provided in SEQ ID NO: 350 and 351, respectively.

A fusion protein containing the antigen 38kD and DPEP separated by a linker was prepared as follows.

38 kD DNA was used to perform PCR using the primers PDM-176 and PDM-175 (SEQ ID NO: 352 and 353, respectively), and 1 µl PET28Ra3/38kD/38-1/Ra2A-12 DNA at 110 ng/µl. Denaturation at 96 °C was performed for 2 min, followed by 40 cycles of 96 °C for 30 sec, 71 °C for 15 sec and 72 °C for 5 min and 40 sec; and finally by annealing at 72 °C for 4 min. The two sets of primers PDM-171, PDM-172, and PDM-173. PDM-174 were annealed by heating to 95 °C for 2 min and then ramping down to 25 °C slowly at 0.1 °C/sec. DPEP DNA was used to perform PCR as described above. The 38 kD fragment was digested with Eco RI (New England Biolabs) and cloned into a modified pT7ΔL2 vector which was cut with Eco 72 I (Promega) and Eco RI. The modified pT7ΔL2 construct was designed to have a MGHHHHHHH amino acid coding region in frame just 5' of the Eco 72 I site. The construct was digested with Kpn 2I (Gibco, BRL) and Pst I (New England Biolabs) and the annealed sets of phosphorylated primers (PDM-171, PDM-172 and PDM-173, PDM-174) were cloned in. The DPEP PCR fragment was digested with Eco RI and Eco 72 I and cloned into this second construct which was digested with Eco 47 III (New England Biolabs) and Eco RI. Ligations were done with a ligation kit from Panvera (Madison, WI). The resulting construct was digested with NdeI (New England Biolabs) and Eco RI, and transferred to a
modified pET28 vector. The fusion construct was confirmed to be correct by DNA sequencing.

Recombinant protein was prepared essentially as described above. The DNA and amino acid sequences for the resulting fusion protein (hereinafter referred to as TbF-8) are provided in SEQ ID NO: 354 and 355, respectively.

One of skill in the art will appreciate that the order of the individual antigens within the fusion protein may be changed and that comparable activity would be expected provided each of the epitopes is still functionally available. In addition, truncated forms of the proteins containing active epitopes may be used in the construction of fusion proteins.

From the foregoing, it will be appreciated that, although specific embodiments of the invention have been described herein for the purpose of illustration, various modifications may be made without deviating from the spirit and scope of the invention.
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Reed, Steven G.
Skeiky, Yasir A.W.
Dillon, David
Campus-Neto, Antonio
Houghton, Raymond
Vedvick, Thomas S.
Twardzik, Daniel R.
Lodes, Michael J.
Hendrickson, Ronald

(ii) TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS

(iii) NUMBER OF SEQUENCES: 355

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: SEED and BERRY LLP
(B) STREET: 6300 Columbia Center, 701 Fifth Avenue
(C) CITY: Seattle
(D) STATE: Washington
(E) COUNTRY: USA
(F) ZIP: 98104-7092

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 05-MAY-1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Maki, David J.
(B) REGISTRATION NUMBER: 31,392
(C) REFERENCE/DOCKET NUMBER: 210124.111C9

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (206) 622-4900
(B) TELEFAX: (206) 682-6031

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 766 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| CGAGGCACCG | GTATTTTGAA | CCAAGGACAC | AATCGACGCG | CAACAGAAGC | GAAGAAGACA | 60 |
| ACCATGGAGA | TGGTAGAAC | GATCCCGCAC | AATCGGACGC | GCCGGATTGC | AATCGGCGGC | 120 |
| CCTGCGCCGC | GTGCTGCCCTTC | GATCATGAACT | GCCGGCGCGG | GCCTGATACC | GATGCGGCGC | 180 |
| GTGCTGTCGG | CCGGCGGACT | CCGTGTGGAC | CCGGACCATC | AGACGAGCGT | GCATGGACGAC | 240 |
| GCCCAGGTGA | CCAAGCTGCTG | CAAAGGGATT | GCCCAGGGACG | AGCTGTCGTT | TGCCGACAGG | 300 |
| GCCAGTCTGG | TCGAGGGCGG | CAGAGGGCGG | ACCAGGGCGG | GATCGGAGAA | CATCCAGCGG | 360 |
| GCGGCCCGCG | GTGTCGCCCA | CCGGGACCGT | GCCGAGCTTT | TGCCGAGCTT | CTGCTGCGGC | 420 |
| GTTCGCGCTG | ACGTCAAAGG | CCGGGAATCA | GATCGCGGGA | TGCTGACGTC | GACCTGGGGA | 480 |
| ATGAGGATTG | TGCAGGACGC | AAGGNAATG | ATGAGGCGGG | CGCTAAAGGC | CGCTAAAGGC | 540 |
| GAGTACGGCC | CCGCTCTCGT | ACGACTCCAT | CGGACTAAAT | CGGGCCTGTT | GCCGGTTGGC | 600 |
| GTGTCGCGGC | GCCAGCGGCT | GCCGGCGGCT | GCCGGGGCGT | CCTGAGATAG | GTGTTGCTCTC | 660 |
| GNCACACGAG | ANCACCCCGC | NTTCTGCTC | TCTCTGNTGNT | GNATGA | 720 |

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 752 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| ATGCATCACC | ATCCACCATCAG | CGATGGAAGTC | AGCCTAGAGA | CGACCTCGCGT | CTCTCGCGCAG | 60 |
| GACTCCCTCA | GCGAGCTGGA | CCGTCTCAGG | CGAGGAGTTA | CGGAGGCGGT | GCTCGCTCGG | 120 |
| GTCGAGGGGC | TCCCGCGGCG | CTGCGCTCTG | CTGCTAGTCA | AACAGGGCGG | CGCCCGCGG | 180 |
| TCCCGGTGCC | TACCTCGAAC | AGCCTAGGATC | CGCCTGCTGCT | GCCGGCAAAGA | GCAGGAGAT | 240 |
| TTCTCTGCAG | AGCTGACGCT | GAGGCGGTGTC | CATGCTGATC | TCGCTGTTGAC | ACAAAGGAGA | 300 |
| TCACATTGGC | TCGATCTGCG | GAGTCTCATG | GCACCTGAC | TCAAGGCTGC | GCCGGGTGGA | 360 |
| CGTCGCGTCG | TGGCGCAAGG | CCCAGAGGTC | CAGATGAGGA | AGCTGACGCT | GTTCTGCTG | 420 |
| ACCGAGCCAA | AGCAAGGCGA | GAGTACGGGG | AGTACCCGGG | GCCGGTGACG | GACCGGATA | 480 |
| GCCCGCGGTG | GCTGGCGGCG | TCGATTGGG | CGTCTGCGG | ATGCTGACG | ACCGATTTT | 540 |
| CCGGAGTTCG | AACACATCTCG | AAGATTCGAT | TCTTGAGGAG | CTGAGGGGTG | AGGGTGGGC | 600 |
| CCCCCGGCTG | CTCATCGGCG | GGTNTCCTCGN | GTTTCCACCC | GCTACCCAG | GCCGGCCCGN | 660 |
| TTGTCAATCC | NTTCTGCNCT | GCCCNAAAGG | GAAACNTTAN | GTTGCGCGTGN | GAANNGTGN | 720 |
| TCNGGGCCCG | NTCTCTGGAAN | CNTCTCCGCC | CT | 752 |

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 813 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| CATACTGATC | ACCATCACC | TCACACTTCT | AACCAGCGCG | CCGGAGGCGG | GCCTGAGCGA | 60 |
| CCAGAGCCCA | CCGGCGCGGA | TCAGACTTTG | AGCTGAGTTC | TGGCTGCGCA | TGGCGCTCGG | 120 |
| CAGCCGAGTG | CCGTATGTTT | GTGCTGAGCT | CAGATATGCC | GCCGATGGC | GTCCTGCGCTG | 180 |
(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 447 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQUENCE ID NO:4:

CGGTATGAAAC AGGGCGGCAG CGGATAACTT CCAAGCTGCC CGGATGGGCC AGGGATTCGC 60
CATATTGATCT GGGGAGGCTA TGCTGATGCC GCAGGGAGTG GGAGGCTGAC 120
CAGCTTTGCAA ATCCGACGCT CGCGGAGGAGT TGGCTGAGCA ACAAGGGCAA 180
CSSCACTAACA GTCCGACGCT CCGCTGATCC CTGCTGATCC GCAGGGAGTG 240
CGGGAGGAGATG TGGCCGAGGCT CCGCTGATCC CTGCTGATCC GCAGGGAGTG 300
ATACCGGAGG CGGCTGAGGCT CCGCTGATCC CTGCTGATCC GCAGGGAGTG 360

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 604 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQUENCE ID NO:5:

GTCCGACGCT GTTGGCCGGAG TATGTCGCCG ACAGAATGTC CCAACGGAAAT 60
CGGTGAGTTCC GAGTGTGGCC GTTGGCCGGAG TGGCCGAGGCT CCGCTGATCC 120
AGCGGACGCG CCGGAGGAGG CGGAATGGGA CGAGTGAGGA GCAGGGAGTG 180
CGGGAGGAGATG TGGCCGAGGCT CCGCTGATCC CTGCTGATCC GCAGGGAGTG 240
ATACCGGAGG CGGCTGAGGCT CCGCTGATCC CTGCTGATCC GCAGGGAGTG 300
TGAGAATGGAG AAAACGGGNG NGAGCAGCAGT GTTGTTCTGTTTG TGGNTAGTG 360
NGTNGGNGNNT ATCCAGGATGC TCTCGCNCA AANCTGATGN CGAGAAACGG 420
NNANNCNAN NNGTGTCCNAN CCNANNTCCN TCGNATGNCANAN NNNANNGNNG 480
NAAAGGGTG GANCAGNNNN AANTNGNNGN CCHAANANNC NNNANNGNNG NNNANNGNNT 540
NNNNTNNNNC NNNNNNNNTG NNGGNGNNCN NNNCAANNNCN NNNNNGNNA NNGGNTNTNT 600
NAAT 604
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 633 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TTGCAMGTGC
AACACCATCA
CTAAGGGAGA
CAAAAGCTING
AGCTCCACCC
CGGTTGGC
CG
CGGCTCTAGAA
CTATGGAA
TYCCKGCCTG
CATGAATYCG
GTYACGGACAT
TGGGACCT
120
TACGTCTGGTCTG
TGGGATACCC
GACGATCC
TGGCAGCAC
CG
CGGCTGGAGA
CCCTCACCCT
CACCGCGGC
CAGTCGGCTCYAT
AGCGGCTTCG
GGCCGCTCT
180
CGGATGGCTT
TTTGGCGGY
GGGGYGGAC
GCCGAGGGC
AGCAGACAT
CG
ATCTCACCAG
GACCCGATCC
GGTTGTCTGC
ACCGACTG
AGCTGGCCG
300
GCCAAGCCCG
CTGACCGACA
CCTCAGCGCG
GTGGGCACGC
ATAGCCAGA
CG
CGGACTGAC
GGCCGCGCTG
TCGCCAGCTG
GCTCGAGACTG
CGACCTCCT
420
GATGGCCTCC
GACACGGCC
CCTGGCGNGA
CCACACCCG
CGGTTGGGCG
TCGTCGGC
540
CTGGGAGGCT
AGTGGTTGCTT
CTCAGCAGAAA
GGCCGGCTAC
GCCGCTGCGC
GGTGGATGAG
600
CCTGACCCGGC
GACATTCTCTG
CGGTGAGCAGA
633

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 1362 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGACGACGAC
GGCCGGCGGAG
AGCCGGCGCG
AAGCCGGATCT
GACGCGGGCC
TGGCAGAGT
CG
CGGACCATCC
CAGGGGCGAG
TGGGAATCTG
AAATTGTGCA
ACRATATTGA
GGCCGGCTGG
120
CCGCCGGAG
CCGGGGCCCG
GTGCGGCAG
GTTACTGCG
AGCCGGCCGC
180
CGGCTGGCCC
AGCCGGCGTC
CATGGTCTCC
CCGCACAGGG
GACTGCTCA
CGCCGGCTGG
240
GCCGCGGGCG
TGGCGGCCAG
CCACGCGCC
CTCGTCGGGC
CCGGCCGGGC
160
CTGCTACCGG
CAGCACCCAC
GGCCCAACAC
ACCCAGCCG
420
GCCGGGACG
GAACGCAAGGC
GTATTGGCG
GGGCGAGCTT
AGCGCGGGAG
540
CAGTTCTCGCC
GGTTGCTCAG
CGGCAAGACTC
GGGACGAGA
480
ACATCTGCGT
CAGCGTCGGG
CTGGTCTGAG
GTTAGGAAGAA
CTTTCCTCGGC
600
CGCGGGCCAC
AGCTATACG
CGGGGGCGGT
GAGCTGTGGT
GCCCGCGCA
GGGCGGC
660
GAGACCCTGG
GGGGCCCGCT
CACCGCGCGC
GACGGGGGGC
GACGACGGG
720
GCACTGGGCA
CAGCGTCCGA
GCGATACAG
ACCGGCTGGC
CCGGCCGGAC
780
GAGACCCCGC
GACATCGGC
GACGCGACTT
GTGACGTGCG
CCGGCCGGGC
840
TGGCCAGGCC
AGCCCAATCCG
GATGACAGCT
GGCTGAGAG
CGGAGCACAC
900
CCCACCCAGCC
TTCAGCGCCGC
GGCCCTGTCA
CCGGCCCTCA
960
GTCACGCCAG
ACGACGTCTGC
CGGGCCGCGA
TTCTCTGCTG
AGCCCGATGC
GGGCGTGGTT
1020
GGCCCTCTGG
CTGGCGCGGC
CTTCAGCGGC
GGCCGGCCGA
TGGCGTACGT
1080
GCCCGCGAG
GCCAGTGCTG
GGGCGAAAAC
GCAGCTGTGGT
GATCGCTGAT
1140
TAGGTGGCTA
CTGGCTGGCCCC
GAGGATCGCC
GGGGCGCGCA
ACGGGATGCG
1200
GGAAGTGCGG
CCCAGCTGGCT
TGGCGGGCA
GCCGCTGGGG
GGGTTGGGCT
TGCGCGAGCT
1260
GGCCGATCGG
CTGGCGCGGC
GCCCTTGGCC
GAAACGTGTG
CTCAGCCTG
CGTCACGCGA
1320
GGACCGGACG
GTGACCCGCG
GCTACCTCGCC
GGCGCGAAGC
1362
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1458 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

<table>
<thead>
<tr>
<th>GCGACGACCC</th>
<th>CGATATCCGG</th>
<th>GGCACCCTAG</th>
<th>CGAAAGCGGT</th>
<th>CGCCGACGCC</th>
<th>CTCCGGCGGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GTATCGCTCC</td>
<td>CTTGAGAGAC</td>
<td>ATTCAGGACT</td>
<td>CGGATGGAGG</td>
<td>CGGCGTGGG</td>
<td>GAACCCGCTC</td>
</tr>
<tr>
<td>TGATGAGCCT</td>
<td>TACGGCGGTT</td>
<td>TATATCGACT</td>
<td>ACGGCGAGG</td>
<td>GGCGCGGGGC</td>
<td>GGCGGCTGAG</td>
</tr>
<tr>
<td>CTAGGAGCCTT</td>
<td>GCTCGGGTG</td>
<td>CCGGAGAGGT</td>
<td>TAAAGCGTGA</td>
<td>CTTGCGCGGC</td>
<td>GTGACCTGAC</td>
</tr>
<tr>
<td>TGCGCCAGGC</td>
<td>CATCTGCTCTG</td>
<td>CACGAGGAGC</td>
<td>AGGGCGGCCC</td>
<td>GCGGCAGGG</td>
<td>ACGGGGCGC</td>
</tr>
<tr>
<td>TGATGGAGCC</td>
<td>ATCGCGGGGC</td>
<td>TGGTCGCGGG</td>
<td>CGGCCAGAG</td>
<td>CAGTATGAG</td>
<td>CGGCGCTGCT</td>
</tr>
<tr>
<td>CGAGGCGGCTG</td>
<td>GGGCGAGGGG</td>
<td>TTCGCGACCG</td>
<td>TATCTACGAA</td>
<td>CTCGGAATTC</td>
<td>CTCCGGCAAT</td>
</tr>
<tr>
<td>CGCGCCACTT</td>
<td>GATGAAACTT</td>
<td>GCAGCCGCGC</td>
<td>TGGAAGCTGT</td>
<td>GGCGCGTTGT</td>
<td>TTTTGCTGG</td>
</tr>
<tr>
<td>CGATGAGGCA</td>
<td>TCTCGCTCGA</td>
<td>CTTGATTTTG</td>
<td>CGACTTTGG</td>
<td>AGACGCGGCC</td>
<td>GACTGCGAGC</td>
</tr>
<tr>
<td>GGGCTGAGGG</td>
<td>CGCCACGCGG</td>
<td>TATCGGTTCA</td>
<td>TGGCGCGGG</td>
<td>AGCGCGGAG</td>
<td>GCGCTGGG</td>
</tr>
<tr>
<td>CTCGACCGG</td>
<td>CGCAGCGGGC</td>
<td>ACGCGAGCGC</td>
<td>TGTCGTTTCTT</td>
<td>AGCGGCTGT</td>
<td>GACAGTCGCC</td>
</tr>
<tr>
<td>CGGGGTTGCT</td>
<td>CTCCTAGGCC</td>
<td>GTGCGCGGCG</td>
<td>GTGCGCCCGT</td>
<td>TATGGCTGTG</td>
<td>CTTGAATGTC</td>
</tr>
<tr>
<td>CGACCCGCTG</td>
<td>TATCTTGATT</td>
<td>TCGCTCGACG</td>
<td>CCAAGCGCGA</td>
<td>ATCCCCGCGA</td>
<td>GAGCTCCGG</td>
</tr>
<tr>
<td>ATCGCTACCT</td>
<td>ATCGTGTTGT</td>
<td>GTGACCCAG</td>
<td>GTCTCCGCGG</td>
<td>GCACCTGGAA</td>
<td>CGCAACGGCC</td>
</tr>
<tr>
<td>TACACCGGCT</td>
<td>GTGCAACTCCG</td>
<td>CGCAAGCGCC</td>
<td>AGATGCGTG</td>
<td>GCGCTGGCC</td>
<td>GGGCCCGAAC</td>
</tr>
<tr>
<td>TGTTCCCGGC</td>
<td>CACCTGGCAA</td>
<td>GCGCGCGAGG</td>
<td>CGGGGCGGGA</td>
<td>CGGCGGGTC</td>
<td>TCCGGCTGG</td>
</tr>
<tr>
<td>ACAGGATCAA</td>
<td>TACGGGACAAC</td>
<td>CGGCGCGGGC</td>
<td>GAGAGCGCGG</td>
<td>TATCGGAGGC</td>
<td>ACCAAGCGG</td>
</tr>
<tr>
<td>CGCGGAGGAT</td>
<td>CCCAGCCTGGT</td>
<td>CTTGACTCAT</td>
<td>CAATGATATC</td>
<td>CGGCTGGACC</td>
<td>AACTCCGCC</td>
</tr>
<tr>
<td>GAGATCTCGG</td>
<td>CGACGCTGGG</td>
<td>GTCGACTGGG</td>
<td>ACGCGCGGCG</td>
<td>GAGGCGCCG</td>
<td>GTCGCGCGG</td>
</tr>
<tr>
<td>TGGCTGTTCC</td>
<td>TGATGAGCTG</td>
<td>ATCGATGCTA</td>
<td>GCCGCTCAGG</td>
<td>CTCGGTCCCA</td>
<td>CTGGCGGAG</td>
</tr>
<tr>
<td>CGCGCCGCAA</td>
<td>CACCCCGGAC</td>
<td>ATCGGCGCTG</td>
<td>GAATCGATGG</td>
<td>TTGGCGGCA</td>
<td>CTGCTGCGG</td>
</tr>
<tr>
<td>CACTGTTGTAT</td>
<td>TCGTACGAC</td>
<td>AGTGAAGAGA</td>
<td>CGCTGCGGTT</td>
<td>AGCCCGCCG</td>
<td>TCTAGCGGT</td>
</tr>
<tr>
<td>GCTACACGA</td>
<td>GCCGCCTGGG</td>
<td>ACACCGTCCC</td>
<td>GGAGGCTGCG</td>
<td>CAGAAGCGG</td>
<td>GGCCCACCGC</td>
</tr>
<tr>
<td>CGCGTTCCTG</td>
<td>CGATAGGGGG</td>
<td>TTCGCGGCGT</td>
<td>CGGGGGCCCG</td>
<td>GCGGACGCCA</td>
<td>CAGTACCGT</td>
</tr>
<tr>
<td>CCGTCGCTCC</td>
<td>GACGGGCA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 862 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

<table>
<thead>
<tr>
<th>ACGTTGTAAT</th>
<th>CGTGGCTGAT</th>
<th>CTGGAACCCG</th>
<th>GTGGCCGCGCT</th>
<th>ACCTACCGAG</th>
<th>ATCTACGTCG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GCGCGAGGG</td>
<td>GCTGCGGCTCC</td>
<td>GCGGAGCGGT</td>
<td>TGCGCTATGT</td>
<td>GCGGATCGG</td>
<td>TGGCGGATCG</td>
</tr>
<tr>
<td>TCACTGCCCT</td>
<td>GCTGCAACGC</td>
<td>AGGCCGGGTG</td>
<td>CAACACCGGT</td>
<td>CGGCGCGGCC</td>
<td>AAGCGGCGCT</td>
</tr>
<tr>
<td>CCGCGCAAG</td>
<td>CCATCGGGGC</td>
<td>TGCCCGCGAC</td>
<td>CCAAGGACC</td>
<td>CAGGCGGGCG</td>
<td>GGCGCAGCG</td>
</tr>
<tr>
<td>AAGTTAACGC</td>
<td>CGCGCGGGCC</td>
<td>CGCGCGCGAG</td>
<td>GCACAAACC</td>
<td>CGGACAGACC</td>
<td>ACGGCCACCG</td>
</tr>
<tr>
<td>CGGCGGCTCA</td>
<td>GCGCGCGGCC</td>
<td>GTGCTCAAGG</td>
<td>AAGGGGACGA</td>
<td>TTGGCCGATG</td>
<td>TCGAGCGCTG</td>
</tr>
<tr>
<td>CGGTCAGAG</td>
<td>TTTGACGAGC</td>
<td>GCGGCGCAAG</td>
<td>ACTAGCGCTG</td>
<td>CGACGAGCG</td>
<td>AAGTTGACCA</td>
</tr>
<tr>
<td>TGCTGCACTC</td>
<td>CAAACTCCG</td>
<td>CTGCGTCTTC</td>
<td>GTAAAGCGGA</td>
<td>CTGCGGGGCG</td>
<td>GCGGCTGTTG</td>
</tr>
<tr>
<td>CCGCTACTG</td>
<td>TTTACTGCTG</td>
<td>GACACACAGG</td>
<td>GGTGGTGTGT</td>
<td>CAACCTGGAC</td>
<td>TGCGCCCTCT</td>
</tr>
<tr>
<td>CGGATGAGC</td>
<td>GCTGCTCAAG</td>
<td>AGCTTTCGCC</td>
<td>CGGTCGAGCA</td>
<td>GTGAAAGCAG</td>
<td>GCCGCGATCT</td>
</tr>
</tbody>
</table>

ACGGTGGAGA 60
ACGTTGTAAT 120
CGGCGGCTCC 180
GCTGCGGCTCC 240
CGGCGGCTCC 300
CGGCGGCTCC 360
CGGCGGCTCC 420
CGGCGGCTCC 480
CGGCGGCTCC 540
CGGCGGCTCC 600
GGACCGGGAT GGGATCGCCG CGCCGCTGCC CATTCGCCGG GCGCGGATGC GGGCGGCGCA 660
CCCTACATCT CGTGCTGACG CGTGGGAAAT CGTGGGCACT GATGGTCCGT CGGATCTCTG 720
ATTCGCGCC CGCCGCGCG CGGACCGTG TCGGCGCGG GCGGATCGGC GGGCTGCGCG 780
CGGAGTGCTTG CGCCGCAAGGC GCTAACATTG TGTACGCTGA TGTTGCTAGTC CGGCGAGCTG 840
GACAGACCTG CGCCTCGTGCG CG 862

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 622 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:10:

TTGATGACGAC CGCCGCAAGGC GTTACATGCG TTCTTTGGTG TGGAGGTTGC GTAATGCAA 60
GACACCCCGG GGCAGCCAGAT CGTGGCAAGTA GTGGCCGGTG GTGGCGCGGG GAGCGGCTGA 120
GTCCGCGGAG CGCCCTGCTTG CAAAGCCTGG CAGGCCGAGG GCGCCGCGGG GTTGGCTTAGC 180
CTTGCAGCCG CACGTGGCAAC AGTGGCAAGTC ACCCTGCGCA ACGGGCAGCA GTGATGAAAG 240
TGGCGGTCTG CGCCTGCTGTG CTTACGTGTA CGCTGCGCAG CATGGCAGG GCAGTGCGATG 300
ACACCAGCGT ACCGGTTTCTG ACCAGCTGCGT TGTCAGGGAG GCCGAGCTGAG GCCGTTGACCG 360
TGGGCTCCGG GTGCCGCGGC AGCCGCGCTG TCCGACTGGA ATCCGCGCGG GCCGAGCTGAG 420
GGGCGGTTGG TGCTCTGCGG GCCGCGCGCG GAGCGGCGGC TCCGAGCAGA CAGAGGCTGG 480
CGGGGTGGTG CGCTGTGCGG GCGGGCGGCG GCAGGCGCTG ACGGCTGCGG CAGTTGGTGG 540
CGGAAGCGAC CGCCCGAGCT CT

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1200 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GGCGCAGGCA TAAGCTTGTG GGCACCAGGG AAGCTTGTTG TACACCTGTA GCGGGCGTGGC 60
ACCAACACGT CTGGTGTCCG CGCAACGCGA AGCTTGGTGT CGTGGCACTG CGGCGGCAAG 120
AAGGAGCTCC ACTCCAGCAGG ACTGGAGCGA CAGAAATATG CAGTGGACGA GTTGCCTCATAT 180
GCTACGGTG GAGTGGGTAC GGCTAATACG TTGGACTACA AGGCCAAGG GTCGGGGCGC 240
GGGGTGACCC AGTTTTCTCAA CAAAGAAGC GATTTCGGCG GTCTGAGAAT GTCTGGATGTG 300
CGGCTGCGGG GTCACATCTG CGGGCGGCGG GAGCGGCGGC TTCCGCGCGG ATGGGCCTGC 360
CGGAGGCGGT TCGGCAGCGT CGGCGATCAG TACAATATCA AGGGCGGTAG CAGCTGTGAA 420
TGTACGACG ACCTATTGGG CAGAAATG AACCGCAGCA TACACCTGTA GAAATGATCCA 480
CAGATCTCCG CCATTCTGCG CAGCACGGAG CTGGGGCAAA CAGCATGGAG CTGGTTGCTTC 540
CGCGCCGCCA AGTTCCTGATT GCTGGCAAGC TTCCAGAAAT ACTGCAGGCG TGTATCCCAA 600
GGGGTGGGAG GCAAAGGCAG CAGCGAACGG TCTAGGGC AGTCTGGGCG GCGGGCCAGC 660
GGGACGACAG GAAAGCGTG CACCTGCGTG ACGGCAAGG GGTGGTCATC CTACAGCAAG 720
GATCCCTGCA ACTGACGCGC CGGACGCTGTC GGTAAAGCAAC TGCGCGGCGG CGAGCTGCGG 780
GGCAAGACGGA AGACCGCTGG CAGACAGCGG CAGGTGGGAA ATGGCCCGAG GTAATGACGC 840
TCCCTACCGA TGCTGGCGGG GAACAGCTATG AGCTTGTTGTT ACAGACCCAC CAAGCTGGGC 900
ACGCTGCTTG CGCTATAGAG AGTGCTGCTG CGAAAAACCC GGAACGCGCG 960
ACGCTGCTTG CGCTATAGAG AGTGCTGCTG CGAAAAACCC GGAACGCGCG 1020
GACCAATACG GCTCCATTCC GTTGCCTAAA TCAGTCGAGC CAAAATTTGC GCCTCCGCGGT 1080
AACATTTTTC TCTGACCTAG TGAAGAGGATG TGTGCCTGAG GCAGATGCGGT TCCGCGTACGTA 1140
GGTGGAGTAA TGGGCGCTGA TCGCTGTATT CGGCTGCTCG GCAGGAGGGG GAGTGGCGGAG 1200

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1155 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

GCAAGGAGCT CGAGTCGGTG CTGTTCCAGC GACTGGGCAAT CACTGGGCAAG AACTGGGCAAC CCGACAGGCC AAACAGGACA 60
AGCAGCAACT TACAGCCGAC GCGCCGCAAG TGGAGCTGATG GTGGGCAAG ACAGGGCTTC 120
CCTTTTAATCA AGCTCGGTCC GCGCGCGGCG GCAGTCAAGT TCAGACGACG TCTCAGGTTC 180
GCTGATGCTA AGCTGACGCA AGCTGAGGAG TGCAGCGGAC CGGAGCGGAC AGCTGTCGGG 240
ACGCGACGGG GGCTGACGGG ATGCTGACGGG ATGCTGACGGG ATGCTGACGGG 300
CGGGGGCTA ACGAGCTCGG ATGCTGACGGG ATGCTGACGGG ATGCTGACGGG 360
TCATCGAGCC GTGGACCAAC GGGAGGACG GGTTATTCCG GCTGCGGTCC GTGGGCTTCG 420
GCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC 480
GGTGGAGTAA TGGGCGCTGA TCGCTGTATT CGGCTGCTCG GCAGGAGGGG GAGTGGCGGAG 540
AAGCGCAGAA CACAGTGAGC TCTGACGAGT TCTGACGAGT TCTGACGAGT TCTGACGAGT 600
GCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC 660
GGCGCAATCG GACAGCTCGG GACAGCTCGG GACAGCTCGG GACAGCTCGG GACAGCTCGG 720
GCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC 780
TCGACAGGCC GACAGCTCGG GACAGCTCGG GACAGCTCGG GACAGCTCGG GACAGCTCGG 840
AGCTGACGCG ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC ATGCTGACCGTC 900
AGATGCGCCAG CTCGGATCTC CTGGCGACGC TGGCGATCTC CTGGCGACGC TGGCGATCTC 960
GCGCGAGCTC TCAGATAGCT TGGCGATCTC CTGGCGACGC TGGCGATCTC CTGGCGACGC 1020
TTTTCGCGCTA GAGTCCAGCG TGGGCGCAAT CGGACAGGAG TTTGCACACG GTTGACAGGT 1080
CGAGTCAGCT CGTCA

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1771 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

GAGCCGGTTC TGGGGTTTGA ACGGTTTATC CGAGTGGCAT CGGACGCGGC GTGGCCGCTG 60
TCCGGCGCTG GGTTGGGCGG CGTCAACAGG TGGCGTCTCA ACCAGCGCGG ATGGTGGCGG 120
ATCGAAGAAG CCGACACGGG CCGGACACGG CGGTGAAAGT GGATTTACGG GTGGCGCCCG 180
GGCGGCTGAG TGCGGATGCG GTTGGGTGACG TGGTCGCGCC GGAGTTGCGCC 240
ATCGAAGAAG CCGACACGGG CCGGACACGG CGGTGAAAGT GGATTTACGG GTGGCGCCCG 300
ACCTAGTTGT GCAGTTGACT TCAGAAAGCA CACCATGGTG AGTCCAGCGG TGCGGCAAGTT 360
GGGGCGGGAT GTGGGGAATG TCGAGAGGAA GACAGCTCGG ATGCTGACCGTC ATGCTGACCGTC 420
GTAGCTGCA CCGCCGGCGG ACGCGCGACG CCGCCTGAAT GCGAGGCGGG ACCAGCAGAAG 480
GTAGCTGCA CCGCCGGCGG ACGCGCGACG CCGCCTGAAT GCGAGGCGGG ACCAGCAGAAG 540
CCGCTAACCC TACGGGCGGT TGGGTCGTA CCGCGGCGGT CTGTGACTCTG GCTGATCTCC 600
### INFORMATION FOR SEQ ID NO: 14:

#### (i) SEQUENCE CHARACTERISTICS:
- **LENGTH:** 1058 base pairs
- **TYPE:** nucleic acid
- **STRANDEDNESS:** single
- **TOPOLOGY:** linear

#### (ii) SEQUENCE DESCRIPTION: SEQ ID NO: 14:
```
CTCCACCGCG GTGGCGGCGG CGTCTAGAAC CTCTAGAAGCT CTCTAGATCG CTCTAGATCC
AGAGGAGATCC GAGATGCGGCA GTCTAGAAC CGCCGCGGCG GCAGCCGCGG GGAAATATCG CTCTAGAAGCT
ACCCGGCGGA CGCCGCGGCG GCAGCCGCGG GTCTAGAAC CTCTAGATCG CTCTAGATCC
AAATGCGGCA GGATGCGGCA GTCTAGAAC CTCTAGATCG CTCTAGATCC
```

### INFORMATION FOR SEQ ID NO: 15:

#### (i) SEQUENCE CHARACTERISTICS:
- **LENGTH:** 542 base pairs
- **TYPE:** nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAATTCGCGCA CGGAGGGTGA TCGACATCAT CGGACCACGC CCCCATCCTG GGGAAACAGGC 60
GGCGCGCGAG GCGTTCGAGC GGGCGCGCGG TAGGCTGCTAT GACATCCCGG TCGCTCGCGGT 120
CATTGACGAG CAGATGGCGC TGGACAGGCC CGGCAAGATC ACCATCAGCA TGCAAGTCTGA 180
AGTGTCTGTT AACAGTGAAGG CCGGCGAACCC GGGCTAGCAC GGGCGCGGCA GCAAAAGCA 240
AAATCGCAGG GTTACGGTTT GATTCTGAGC ATTTTGTGTC TGCTGCGCGA GGGCTTACCAG 300
GGCGCGCGCA GCGTTCGAGC CGTCCGCTATC CAGGGCTGAC TGGCGATTCG GGGCGCGCGC 360
CCGAGATTTA TGTCTCGGCT CGGACCAGAC TGGGCGATTC TGGCGGAGAG TGATCGTGTA 420
CCGTGCGCGC CGGCGTGAGC CCGAGTGGGC CCGAGGAAAC GTGGTGCCCG GCGCTTGAAGA 480
AGGCTCGGTA GGGCGCGGGT CGTACGCGCT CTGACGGCTG CCTACGTCGCG GCCAGGCGAC 540
GG 542

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 913 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGGTGCAAGCC CGGCGCTCCGG TTGCCCCAT TGGCGCGCGG CGGCGTCGAG TGGCGATCGG 60
CACCACACGC GCCCTTGCGC CGCGACCACG CGGGCGGCCG CGGTCTGGCC ACCGGCGCGC 120
TGACCAAGCT CGGCGCGCGC CGCCATTGCC ATACAGCAGC CGGGCGGCGG ACCGGTTACC 180
GCCGTCCCGA CGCTCGGCGG CGGCTGCTT TTGCGCCTCT TGGCGCGGGT GAGGCGCAAT 240
CAAGCGCGCG CCGCGCAAGCG TTGCGCGCTT TTGCGCGCGC CCGCGCGCGG CCGCGAATTG 300
CGCGAACAGG AMGACGCGTT GCGCGCGAGG CGCGCGAATG GACGGCGGAT CGGCGCGGAC 360
GCCGGGACG GCCGATATTGC CGCGGTGCCG GTGGCGTGCG CGGCGTTAC CGGCGGCGCC 420
GTCTGGCGAC AAATCCGCGC GGGCGCTCCGG AGACCGCGG CCGCGCGCGG TGGCGCGCGG 480
CACGGAAAC CAAGCGCGAC GGGCGCCCGG GGGCGCGCTT TGGCGCGATC CACCGCGCAG 540
TACCGCGCA CGCGCGACGG CCGACGATAG AATGGTAAAG CCTACGCGGCGG CCGCGTTATT 600
CGGGCGGCGC GAGNCGCGTG AAGCTGTAAG AAGGGCGCTG CCGACGGCGC CCGCGTTATT 660
CGCGCGCGCG AAGCCGCCAG GTCCGCGCGT TGCCGCGGCT CCGCGCGCGT CCGCGCGCT 720
TGTCGCTGCT GAAACGCGTG GCCTGGCGTG CGGCGTGGTT GCNTGGCGCG GCGCGTCGCG 780
CGCGCGCGCG GGCGGCGTAC ACAGCGCGGC CGGTGCGCGT TTGGCGCGCA TTGGCGCGAT 840
TGCGCGCGTCG GCGCGCTTGC CCGGCGCGCC CCGCGGTTGG CGCGCGCGCG 900
CGCGCGCGCG CCG 913

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1872 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GACTACGGTG GTGTAAGAAAA ATCTCTGGCG CGGACCGCGT TAAACGCTGG GAAATTTCTGA 60
TAGCTACCCCG GACACCGGAG GTACGGGAGT GAGAACATCG CGCCGCGCCT CACTACGGTG 120
80

GTCATGTTTG CTCAGCCGCTG CGGACTGCGGT TGGCTGGCGGT CAGCAGTGGCC CTCGACAGGCG 180
GCGCGCCCGG CCGCCTGGTGT CGACAGACCG TGGTCCGACG TTCCCGCCTG TGCCGCTCGA 240
CCCGTCGCCCG ATGCGCCGCC AAGTGGCGCC ACAGGCTGTC ACCATCAACA CCAAACTCGG 300
CTCAACACAC GCCGCCGGCGG CGGACGACGG CATCGTCATG GATCCCCAGC GTTCGCTGCT 360
GACGCAACAC CACTGTCAGC CGGCGGCCCAC CGACATAAAT CGGTTGAGCG TCCGGCGCTG 420
GCAAACACTCAG CGCGTGGAGT TGGTCGGTAG GTCACGGCGC CAGGATGCTG GTGGTCCGCA 480
GCTGGCAGCT CGCGTGGGCC TCGTGCTGCC CGGCTAGCTG GTGGCCGCTG CGGATGTTGA 540
GCCGGTCCTC GCAGTACGCC ACAAGCGTGG CTGACGCCGG ACAAGCGTGG ACGGCTGGGC 600
CAGGGTGCTGC GGCGTGGGCC AAACCTGTCG AGGCTGCGGAT TCGCCGACCG CGGACGAGGA 660
GACTGGTACAC GGCGTACTAC AGGAGTGCAG ATAGACCGG CCGCGTGGTC ATTACTTCGA 720
GCTGGCCGACG GTGCTGGCAG GATGGCGCAT TCCCGTGCGG CAGCAGTGGT GCTGCGCCGGG 780
CCAATGCGAG TCGCGTCGAG GGTGGCGCAT GGCTGCTGCA GATGCACGG CAGGCGGCGG 840
CTTGAGTGGTT TTGGCTGGCGC GCTGACACCG AGACGCGATG CGGAGCGCTG CAGGGCGCGG 900
CTCGAGTACG ACGCGTGGGC TGGCTGGCGC AGCGTGGGC AGCGTGGGC TGGCCGAGCC 960
GATGCAATGCT GGCGTGGACC GCAGCGGCTC TGGCTGGCCG CATCGTCAGG GCTGCGAGAC 1020
CTGCGTTGGAAC TGGCGGCCAC GATCGCGGCC AGCGTGGGC AGCGTGGGC TGGCCGAGGC 1080
GGGCGGCCCC CGCTGGTGAT TGGCGGGCAG CTACGCTCAG CAGCGTGGGC CAGCGAGAGC 1140
CAACGCGTGGT TGGCGGTGGA GCGCGGCGG CTACGCTCAG CAGCGTGGGC CAGCGAGAGC 1200
GCAATGAGACC AGCGAAGACA CACGCTTGAG TACACCGCTC TGAGGGCGAC TCTGCGAGGA 1260
GGCGGTTGCG TCGAGCTGAC GAGCTGCCAG GTAAGGCGGC TGACGCGCCG CGCTGCGGCC 1320
GAATCGCACTT GGTGCTAACG CGAGCGGCGC TGGCGGCTTG AGCGTGGGCC GCTGCGGCCG 1380
GACAGCGGCG ACCTGCGGCG GTAAGGCGGC TGGCGGCTTG AGCGTGGGCC CTTGCAGGCC 1440
GTGGTGGCGCA GTGAGTGGCA ATGCTCGAG CTCTGCGGCC CTCTGCGGCC CTCTGCGGCC 1500
GTCTGACACA ATACCGGTGC ACACCGGCTA GCTTCGAGAA GCACACAGTT ACCCTCGGCC 1560
GTGCATGACGG TCGACGACCG TGCCCGGCGC GCTTCGAGAA GCACACAGTT ACCCTCGGCC 1620
AATCAGTGAA CCGATCGCG CCGCTGTGGT CAGAGGGCGA CAGAGGGCGA CAGAGGGCGA 1680
TACGCTGGAT ACTAGCTTGT GAGCGAAGCC GAGCGAAGCC GAGCGAAGCC GAGCGAAGCC 1740
TTCGCTCGCA CGGAAGACTC GGAATGCGTC TCTGCGCTTG TGGGCGCTGC TGACGGCCAG 1800
GCACGAGATT GCCTGCACTG TGGGCGCTGC TGGGCGCTGC TGACGGCCAG GCCTGCACTG 1860
GCACGAGATT GCCTGCACTG TGGGCGCTGC TGGGCGCTGC TGACGGCCAG GCCTGCACTG 1872

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1482 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CTTGCAGCGAA ACCTGATGCGC GAGGAACAGG GTGTTCTCAGT GAGCGCGGCGG CGCTGCGGCC 60
CGCCGCTCTC CGCGAGAGTC AGGCAAGCTG TGATGCGCAG AAGAGGCTTG ACCAGGCGCC 120
AGCTAGCGGT CGGAAACACCG GGGAAAGCTG ACAGCTGTCG TTCTCTACAG ACCCTCGGCC 180
TCGAGCTGCG GCCAATCAGT GGGCGGCGG AGCGTGGGCC CAGCGTGGGC CAGCGAGAGC 240
GTGTCGCTAGT TCAGGGTAAA AGGCGAAGCG TAGATCTGGC AAGGAGGCTG AAGGAGGCTG 300
TTTGGACGAC CAAATACCG GGGACGCTG CGCGGCGG AGTGGCAATG GTGTCGCTAGT 360
GCCGCAAGAG TGATCGACTG AGGCGACGTC CTCAGGCGGC CGAGCAGCTG GTCATGCTGGC 420
AACCGCCTAA GCTGAGAGCT GCGGAGGTGG TGGCTGGCGC TGTCGCTAGT AGGCGACGTC 480
GAGGCGTACG ACGCACTGCC AGGCGGCGG ACGAGCTGAG GTGTCGCTAGT AGGCGACGTC 540
AACCGCCTAA GCTGAGAGCT GCGGAGGTGG TGGCTGGCGC TGTCGCTAGT AGGCGACGTC 600
GAGGCGTACG ACGCACTGCC AGGCGGCGG ACGAGCTGAG GTGTCGCTAGT AGGCGACGTC 660
GAGGCGTACG ACGCACTGCC AGGCGGCGG ACGAGCTGAG GTGTCGCTAGT AGGCGACGTC 720
GAGGCGTACG ACGCACTGCC AGGCGGCGG ACGAGCTGAG GTGTCGCTAGT AGGCGACGTC 780
GAGGCGTACG ACGCACTGCC AGGCGGCGG ACGAGCTGAG GTGTCGCTAGT AGGCGACGTC 840
CGGTGCTGATCA TGCCGCGGAC CGGACCCGAGC AAGCAGCTCG TCCGCGCCGA AACACTCGCC CCACGGCAGG 900
AAGGGCTTAT TGAGGCTGAC GATCAAGCTG GCCGCTCTATG ACCGGGAGGA CACCAAGCTT 960
AAGAAAGAGGT TGCCGCGCTCT GGCTCTCATAC GAAATGTAAC GACCTGCTG AACCACAGG 1020
AGGGATAGCC GGGCGCAAAC GGCTGCGTGG GGCTGCTGAG GGGCGCGGCT GGGCAAGCGT 1080
GGGCGCTATGC GAGCGCGGAC GCACAGTTGA GTCTCCCGGG TGCCAGCAGG GGCGCAAAAA 1140
CAGACAGTTC TGCCGCGCAG TGATGAAATT CAGGGTGCCC AGATGCGTCA TGCGTCGCG 1200
CTTGGCAAGCG CGACGCGATCT CAAATGTCGA CTTCTCCAAA GACTCTCGAA CCAAGGAGCG 1260
GCGGCCGCGG CTCCGCGGGCTTT CACACCAAGT GGCACCGGCG GTGCAGCGCG CACGGCGAGC 1320
GCAGCGCGCC AGGAATCTCT GTTGCGTGGGC CGGCGCGCAT CGGCGCGCGC 1440
GGACAGTGAC TCCCAGAAAGC CAATGCGAGG GTCTTGGCTC CGCCATGGT GGCTTGGGTC 1482

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 876 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:
GAATTCCGCA CGGCACGCGCG ATAGCTCTATG GGGCGCGCGAG GACCCAGATGG CTCCAGGGTT 60
CTGAGCCTGG GCACCGCGCG GGGCGCAGCA CTCGACCGGT GAGGGCGCTG AACAGCGAGA 120
CGTCGACTCG TGCTCGCTGAG ACSCGACCAA CCGGCGCTGGT GTTGCTTACG ACSCGCGCTT 180
GCCGCAGCAA ATCCTGCTA TCGNCGAGAG CGACTGCGCC AGATGCGTCC GGGAGAAGCC 240
GGGAGAATTC TGGTCCTAGA TACCTGGATG CAGAACAGCC TGACTCGACG CCAGCCGGCC 300
GCCGACAGGC ATACGCCGAGC AGATCTGGCC TCCGCGGGTA GAGATGGCCC CGGGCGCTGC 360
GCCGACAGCG ATGCTGCGGC CGAGTGCCGA TGCCGGCGGC GATGTCCTGG GTGCTGCGGC 420
TTGGCGCGAG CTAAACCGCG ACSCGCGGCT ATGGGACAGG CAGAGAGCTG GCCACCCCGC 480
TCGCCGCGCG CGCGTGGCCT GATGCGAGG AATGTGGTGG CGTGCGGTAG CCGCGCGCAG 540
CGGCCTGTG CGTCTGCGGC CGCGTGGCCC CGAGCAGATC CGCGCTGGG TTCCGGCGGC 600
ATACCGACCG TGCTCGGCGG ACGGCTCGGG TTTTGCGGCG ATCTCGCGCC CGCTGCGCTG 660
TTACTGCGAC CCCGAAGGAG ATGATCGGTG TGCTGCGGCT TTGGCGGGAG GCGGGCGGGG 720
TCGACGCGGAG AATATCAGAC CATTCTCGTC GTGCGCTGGG CCCGCGCGCC AGTTACCCGG 780
AATTCGCGGA TGGGGGCGGT TGCCGGCGGG CAAAAACGCT GGCGCGCGGC 840

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1021 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:
ATCCGGCGCG GCCGCGCGCG CGGAGCGGAGG TCCGACTCGA GAGAGAACAAAT TCCACGCAGGG 60
CAGATTCCATA ACAGAATTAC CGGCGCGGAGC CAAATGTCGA CGATCGCGCTG ATATGCGAC 120
AGGGAGAACG TGCCGCGCGT GGGCGACAGA TTTTACAGCC CAGGGCGCAG CAAAGGACAGA 180
CATCGAGTGA TGCTCGGCGG ACGGCTCGGG TTTTGCGGCG ATCTCGCGCC CGCTGCGCTG 240
GTAGTGGACG TGCCGCGCGG TGTCAGCGAG CCGGGCGGGG CGAGTGCAGT CGCGGCGGCG 300
CAGCGAAGAG CAGTCAGGGA CAAGGTGCGT CGGCGCAGAC CGGTGCGGCG CCAGGCGGGG 360
GATGGCTCGG GCGAGGAGTT CAGTCAGTGG TTCTCGACAG AAGACAGTGA AGAGGCGGCC 420
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGTATGGCAGA</td>
<td>CACCTCAGGCT</td>
<td>82</td>
</tr>
<tr>
<td>AACCTGAGCAG</td>
<td>CATGTAATTTT</td>
<td>480</td>
</tr>
<tr>
<td>GGGGGGCGCC</td>
<td>TCTAATAGGT</td>
<td>540</td>
</tr>
<tr>
<td>TCCGGCACG</td>
<td>CTCTCGTGTG</td>
<td>600</td>
</tr>
<tr>
<td>CGNGNATCG</td>
<td>CTCTGGTCTG</td>
<td>660</td>
</tr>
<tr>
<td>GAGAAGGACG</td>
<td>GTCTGCGGCT</td>
<td>720</td>
</tr>
<tr>
<td>ACCGAGGGGT</td>
<td>TGGTTGTCG</td>
<td>780</td>
</tr>
<tr>
<td>TCCGACGCGG</td>
<td>CTACAGCTTC</td>
<td>840</td>
</tr>
<tr>
<td>CCCACGTCTG</td>
<td>GCCACGCGA</td>
<td>900</td>
</tr>
<tr>
<td>GGGGAGGGGA</td>
<td>TATGACCTTC</td>
<td>960</td>
</tr>
<tr>
<td>GGCGGCCAGC</td>
<td>AGTTTGTGTTT</td>
<td>1020</td>
</tr>
</tbody>
</table>

**Informations for Seq ID No: 21**

- **Sequence Characteristics:**
  - Length: 321 base pairs
  - Type: nucleic acid
  - Strandedness: single
  - Topology: linear

**Sequence Description:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>CGTGCGAGTC</td>
<td>AAGCGAACAA</td>
<td>60</td>
</tr>
<tr>
<td>ACCGAGGGGT</td>
<td>CTCTGGTCTG</td>
<td>120</td>
</tr>
<tr>
<td>CCCACGTCTG</td>
<td>GCCACGCGA</td>
<td>180</td>
</tr>
<tr>
<td>GGGGAGGGGA</td>
<td>TATGACCTTC</td>
<td>240</td>
</tr>
<tr>
<td>GGCGGCCAGC</td>
<td>AGTTTGTGTTT</td>
<td>300</td>
</tr>
</tbody>
</table>

**Informations for Seq ID No: 22**

- **Sequence Characteristics:**
  - Length: 373 base pairs
  - Type: nucleic acid
  - Strandedness: single
  - Topology: linear

**Sequence Description:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
<th>ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTCTATGGCT</td>
<td>TGGTTGGGCC</td>
<td>60</td>
</tr>
<tr>
<td>CGATCGACGG</td>
<td>GCGCGGAGAC</td>
<td>120</td>
</tr>
<tr>
<td>CCTCGGTGTG</td>
<td>GACCGCCCA</td>
<td>180</td>
</tr>
<tr>
<td>ACCGGATGAC</td>
<td>CCCATGCGG</td>
<td>240</td>
</tr>
<tr>
<td>GCAAAGACC</td>
<td>CGCCGTCGTC</td>
<td>300</td>
</tr>
<tr>
<td>GGCTGATCC</td>
<td>TGGAGGCTT</td>
<td>360</td>
</tr>
</tbody>
</table>

**Informations for Seq ID No: 23**

- **Sequence Characteristics:**
  - Length: 352 base pairs
  - Type: nucleic acid
  - Strandedness: single
  - Topology: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GGTACGCGGCG GATGGGATTC CTGGGCGGGG CCGTGGGCTG TGGCGGTGTT TGATCGACAG 60
TGGTATCAGG GGGTGGGCAA GGGTGGGCGT TGGTACCGAG GGGGCTGGTG CGGGGCTGGT 120
TCCTGACGCG CGTGGTACGCG TGCGGCTGATT CAGCGGCGGC CAGTGGTGGC CAGTGGTGGC 180
TGATCGACAG CAGTGGGCGG GGGTGGGCGT CAGCTGGGCT GCAGTGGGCT GGGTGGGCGT 240
GCCCTGAGGT TGTGCGCTAC GCCCGCGGCT GNAAGGCTGG CACCTGCGGC GCCATCGGNT 300
TTGACGACGA GCTTATCGGG NGGATCGGGG NGGATCGGGG NACATGGGAA TGGCAGGGGA 352

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 726 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GAAATCCCGG TTCCATTCCG TGACCACTCG GCTTGCCGTA ATCGACGAAAG TGAACAAAGC 60
GCGGTCGCGG GCCGTCATCGG TGCACAGCGA GTAATCAGCA AGTACCTGGT TATACGACAC 120
CTAGCGCTCA GTTGCTGGCC AGATCGCTTT CGTACGCTCA GTGCTAGTAC CGTGCTGGCG 180
GCCGCGAGCT CAGCTGCGCG CAGCGCATCC TGCGCCAGGG TGGCGCGGTT CGTGCGGTGC 240
GCCGCGAGCT CAGCTGCGCG CAGCGCATCC TGCGCCAGGG TGGCGCGGTT CGTGCGGTGC 300
CCTGCGAGCC GCTGCGCCCG CCTCGCTCGG ATCCCGATCG ATCCCGCTGC ATCCCGCTGC 360
GCCGAGCAGA GCCGCGCCCA AAGCGCGCA ACTACCGCGG ACCACATCTC GAAAGTGGCG 420
TGCTTGACGA TCCGGGTGCT CCGCGCGCG CCGCGGCGTC CGGCGCGGCG GCGCGCGGCG 480
CGTGAGCGGG GCGCGATCGG GAAATCGCGA TATAAACCGG GCGGCGCCCG CGGCAAGCTA 540
CGCGCGCGCG GCGCGGAGAT TTTGCGCCTG GCGCGGCGGT CCGATCGACG ATCGCGGCGG 600
AAAATACCGG AGCGGTTGGG CACCCGCTTG GAGGAGCGGT GAAAGGAACC TATATCGAAC 660
GCCGACAGCG CCTCCACCAT CGACATCGAC AAGGGTGTTA CCGCGACACC CGGGCGCGGG 720
ATCGTG

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 580 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGCGAGCGAG AGCAAGCTCG GCGCCACCAC CGCCTATCGG TTGATCGAGG CGAGCGGGAT 60
GCTGCGCGGAC CATAATCCAGG CATACTGGTG GCCACCTGAG CGACCTTTTG ACTCGCCGAG 120
CTGCGCGAGT GCCTCGCGGG GAAATCATTT GCCGGGCTTG CTGTCAGCTG ATGAAGCCACG 180
ATTGGAGACC ATAGGGGCTG GATTTGCAGG TCTAATGGCG GTATGCGCTG GAAACACCAT 240
GCCGCGCGAT GCCTCGCGCC GACCAGCGTT CGCCAGGGCG CCCAGGCCCA ATCCGGAGGG 300
AGCCCTCTAG GGGCGCGATG AAGCCTCGGCA CGCGCGACGG TCCTTTGGAA GCAAATACAG 360
AGGCGCGCCG CATTTGATGAA CAGTACGCCA TTTGAGGGTT CGTGCTGGCG ATCGCGGCGG 420
TGCAGCGCCA GGAAGCGCAG GCACTGGGTT AGCAGCTCAG AGGCGTTTCT AGCTAGACC 480
AGCCTAACGG CGAAATCGCGG GGTGACCGCG CAGACCTCCC GCGTATGATC CAGCTGCTGC 540
TGCGCGATGT ATGCCCGAGGA GAACTTGTTG ATCCAGGCGG 580
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 160 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AACGAGGCGG CCGGGGTTTT TGCGGGGGCC GGGCGGTGCG GCGGCAACGG CCGGGCCCGG 120
GGTACGGCGG GTTGTGCGGG TGGCGGCGGG GCGGCGTGGG CCGGAGGGCA CCGGATCGCC 160

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 172 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GACACCAGATA CGATGCGGTG GTAAGCGCAAC GGTGCTGACA CGCTGAGGGG GTTGAGCATC 60
CAGCGCCACG CGAGCGGGGT GACATCGG GATCGCGGGC CGTTCGCCGA GCGGCGTGC 120
AAGCCGATGG GAATCGAGAA GTGCGGTTGA ATTCATACCG GAATCGGCCC GCGGCTGCT 180
GACGCAGAAC AGTGGGAGCA CGGCAACAAC AGCTTGGCGT TGGCGCCCGG TGTGCTGCTC 240
GCTTACGGAG GCACAGTACA GACAACGGCC CCGT 272

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 317 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCAGGCAGTG TTGCTGAGAC TGATGCGGCA CGGCTGACCGA GCAGGACGTG CCGGACGTGA 60
AGCAGACGAT GAAGCGGGCA CGCCCATGCC CGGTTGATGCG CTACCCGGGG GTACCCGGC 120
CGGCGAGCG CTACCGGCGG GAGCGAGCGG GGTGATCGGC GTGCGAGCGG GGGAGATCC 180
GTTGCTGACT GCGGCGTGTC GACAGCTGTTA ATCCGTTCA TAAGGTCCCC GCCGTTGCGC 240
GAGAGTGCGG AAGAAGCGGT CAAAGATCCG CGTCGCGCAG ATGGCGCTG 300
CGGCCTGGTT GGGCGGG 317

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 182 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:
GATCGTGGAAG CTGTCGATGA ACAGCGTGGC GCAGGCAGCG CCGGGCACAG CAGTCGCTGA

GCAGGCGCGG ACCACCTGGC CGTGGGCGGC AAGTGTGATG ACCAGCTGGG CACGGCGCCAC
CGCTTCCGGG GCCTAGCAGA ACACCGCGAC ACCGTGCCGG GCAGCGCGCG AGCGCGCCGG

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 308 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GATCAGCGAAG TTTG TGTCGTAAG GGGTGTGCTGA CCGGAAAGTC TGGGGCCTGG CGAAGCGGGG
CGCGTCGGCA GAGGGCAAGA CACCGCCTGTC CGAGCTGCTG CGGCTGCTCT ACGGCGCGCA
GAGGTTGGAG TGGGCGCGCG CCGCGAGGGG TGGCGAAAGG TTGGGCGGCT GACACTCTCAT
GAGACTGGCG GGTAGGGCAT TGACCGATTG TTGACGCGGG TGGCGACAGA TTTGGAACGT
CGGTGTCAG AGCGAGTTCT CGAAGCTTGT CACCCGGTCAA GCCTGACCTC ATCGACACCC
ACGTGTTGG

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 267 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCGACGACGA GCAA CTCAAGG TGGGTGATGG TGGCGCAGCG CATTGAGGAG GGAAGAAATC
CGCGCGGAGC TGCGCGCGCG CAATGTTGCTA TGGTGACGCG CGTGAGGGG CTCGGCCGAT
GCGACGACG TATT CGGGG CCGGAAGAGG TCGGAAAGGA ATGGAGGGG
ACGACGAGA CATTGACACC TACCCAGTGA TCGAGATCGC CGGGAGCTCG CCCGACGGCG
TCGACGCGCG AATCCAGGGG GGTCTGG

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1539 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CTCGTGCCGA AAGA ATGTGA GGGGACAGGA TGAGCAATCA CACCTACCGA GTGATCGAGA
TCGTCGGGAC CTGGCGCGAC GGCTGGCGAC CGCAAATCCA GGGGGCTCTG GCCGGAACTG
CGACGACCAT GGGGGGCGCTG GACTGGTCTG AAATGACGTC ATACCGAGGC CACCTGCTCG
ACGGACGCTG CGCGACCTCC CGTTGACTGA TAAAGTCGCG CTCCGCGCTG AGGATTCCTG
AACTCTACAG CGGGCGCGAT AAATGAGGGT CATCATTAGG CGACTTTCCC AGGACATCTT
GACGGGTCTG GAAACGGGTG GACCGCGGCG TTGGCTGGCG GAGGCGCTGC TTCCAAAATC
CCTGGCAGCAG TGGTGCGCCGG GCGCTACAGA GGAATGCGGT GTGAGTTATGCT
GGTCGACCTG TGTGGCTGCG AGCGGCAAGC AGCGTGCTTC GACGTCGGCT CGCGCGCGGG
GCGGAAGGCG TTGGCCGCTCA CGGCTATCTC GAACCGCGAG GACGCTTACG CGGGCTTCGA
TGACGGCACG GCTTGAGAGC GAGGCACTCG AATCGCCGATCG ACCCCAACCTC AACCCACCATCG
ACTAGACCTT GCACTTTTCAT ATCCGGATGC GTGCTTCCGAT GTGTCGTTCTT TTACTCCGCTG
GGTGCCCGAC AAGAAAACCTG GGCGACCTGGA TCTCCTTGCGT TATGAGACGTT CCGTCGCTGC
GACGCAAGAG AGCGCGGGGA TGGGGGCGGAT AAGGGGCGG ACCTGGGCGG ATCCGGTTCTA
ACCCGCGGTC TGGCGGCGGA TGGCGGGCGG GACGCGGCGG GGGGCGGCGG TGCGGGCGGG
AGCCTGGGTAT ATCGTGATCG ACAACGAGCG AGCGCCCGCG GTTGCAGCGC AGCGCCCGGG
CAACCCCGGT TCTCGCGGCT CGCTCGCCGG TCCGCGGCGG TCCGCGGCGG CGACCGGCGG
GCCAGCAATCG CGCCCAAGTC CGTGGACGATG GATAGCGAGAA ATAGGAAAGTAA CAAAGCAAAA
ATCGATTTGCG ACCCGCGACTG CTCGGCGCGT GGTGGGCTGG ATGAGATGTCG GGCGCCCGGG
CCGACAGGAA CGCGGACTGC TCGTGGCCGG TCCTGGCGGG

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 851 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

CTCGAGGGGTG GCGTGGATGA GCGTACCCGC GGGCAGGGC GACGCTAGCG CGCGGCGGAG TG
GGCGGTGGC AGCGCCTCGC CGGGACGACCG TGATGGGCTG AGCTGGGCCC CGCGGCGGTG
CGGGCGAGAC ATCGGGACGC GCGGCGGACC CGCGGCGGCG CGCGGCGGCG CTGCGCGGGG
CTCGAGACCC CGCGCGAGGC AGCGTGGATG GACGAGATGC CGCGCCGGGC TCGACGGATG
GGCGGCCGCC GCGCGGCCG CCGCGCCGTC TCGCAAGCTG TGCGGGCTGT GAGAGCGCTG
CTGCGCGCAT GCGCGCGCAG TCGACACGAT TGGTGTGCTG GCGGAGAAAC CACATGGCTG
GACCGGACCG TCAGCTGCAAT CGCGGAGCC AATGCGCTGG GCGGAGAAAC CGAAGCGGATG
GGCGGGCGCG TGGGGCGGAG GGGCGGAGCG GCGGAGCGG GAGGCGGGCG CGCGGCGGGT
GGCGGGCGCG TGGGGCGGAG GCGGAGCGG GCGGAGCGG CGCGGCGGGT

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 254 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

GAACGAGTTCG CGGCGCCGAC CGGCGGCCGC CGGCGGGGCC CGGCGGGGCC CGGCGGGGCC CGGCGGGGCC
GATCGATCGG GCGGAAATTT GGACCAGATT CGCCTCCGGC GATAACCCCA TCAATCGACAC 60
CTAGATTATG TCCGCTCAGG GCGCCGAGTA ATGATCGAGA GGAGGACAGC CTTACTGCTG 120
CGGGCACTCG TCGTACGGC TCGATGCACG GGAGGAGTGG GACATTTTGC ACGGACACCC 180
GCAATCCGAAC GTCCGAGGCC CACTCCAGCT TGTGAGGCGG GCGACGCAGT GCGGAGCTGC 240
GCTTTGAACAA GATC 254

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1227 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

GATCTCTACCC GAAGCCGCGCG CGGCAAGGCC GAAGTGGCCTG TTGGACCAGG TAGGAGGGCGA 60
CGATGCGGCG CTGGGCGATCG CGTTGACCC GGCGGGGCTG GCTGGGTGGG CTTATAACTT 120
TTCCTCTCCG GACCGGGAGC TGGAGTGGTA CCAAACGGCC GAGTCTGGTG TCTGCCAGGT 180
GATCCTGCGGC CGGAGTCGCC GCCGCTAATG GAAAGGCGGG TGATGCGATT TCTCGACAC 240
TATTGGAAGC CAAGGTCTAC CAGTGCAGAT CCAAGCGCA CGGCGGCTCG GCGGCGGTGC 300
GATCGCTTCA ACTGATAAAA CGCTAGTACG ACCCGCGGTG GGCAACACAG TACGAGACA 360
CCAAGAGCCT AGCGCCTCTG AAAAGC CAGCTGGGAT GACCGATGGC TCTTCACTGG 420
GGCGCAGGCG GGGAGTGGCC ACCTGCTATG TGAAGCACAG CTGGGGCTGA TATGGGCAAC 480
AGTACACAGT TTGGTCGTAC GAGGTGTACCT CGACCTTGGGA GAACGTGTGG CGGAAACGCT 540
CGCTCTCGAG CTGTCGAGAC GCTGTACAGC AGCCTGTGTC GGCCACGCGC CGTGACAGAT 600
CGGACGACCC ATGGCGAAGC ATGGTGTCGA CATCGCCTGT TCTCGAGCTC TTGGAGTGAC 660
CCGCTATGCC GGGGGTCGCC ATGATGGAGC AGAATGCTGG TGGCGGTGGT GTCGCCAGTG 720
TGCGAGCCCG GTATAGTACG GCGCCGCTCA TAGCGCAAGC CAGCGCGCTG GCTCACAGCA 780
TGGCGATGAG CGCGCGTGTG TGCGGTGCTC TCGGTGAGGA CACCTGGCCG GCACAGCCCG 840
GATATCGGCC GGGCGCGGAC GCGGCCTTGG TCGCCGCTTC CGGCGGCAAG GCGGTTGACG 900
GGCCGCGGCG GCTGTTGGGG TACGCGAGGC CTGCGGTTGG GTGACGCAGC GGCGCGCGGT 960
ACCGGGCGCC TGCCTGGGTC CGSACACCGG GTTGGCAGCG GTGGGGACCG GCAATTTGTG 1020
TTCTCCCTAG GTGGGCGGTG GAAGCGAGTG CTAGGCGAGT AACCCGGCCT CGGTTCAGAC 1080
GGCACAGTGG GAATACTACT GAGCTCCTGA CGGCGCTAGT CGGACACGCT GGGTGACGCT 1140
CTCAAGCCCG CCGGGCGGGG CGCGCAGGCG ATATACTTTG AAGACTAGGC AACCTTAGGA 1200
ACGAAGGAGC TAGGTTTGTG GACGATC 1227

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 181 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

GGGCTTTGGCC CCGATCGCC GGGTGGTGGTC ATCCGGACAGC CCGGGGCAGG GGGGGCGCGC 60
GGGCGGGCC TAACTGGTGG GCACCACGCC AGCACTTTCGG TTGGGGCCGG GCGGGCTCAG 120
GGCGGGCGCG CACCATGGTG GNGTCCGGCC GTCTCCGGCCG ATTTGCTCTAC GCGACGCGCG 180
G 181

(2) INFORMATION FOR SEQ ID NO:37:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 290 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GGGCTGTGGG CGGATTCCGCG GGCTTGCTGAG ACGGCCACCGG CGGTGTCGGC GCACCGGGGGG 60
GGGACCGCGCT GTGGCTGCGG GGGCCTGCGCT CAGGGCGCCG GCCGAGCAGA GGCGAGACTG 120
GGGCGGTGTC CGACGGGCGG CCCGGCGGGG CGAGCGCCGGG GGCGGCGGGG GCAGGCGGGG 180
CCCGGCTGGG CGACCTCCGGG GCAACGGCGG GTAAGGGTTGC CCAGGCGCGG ATCCGGCAGG 240
GACCTCAGAG CGGCCACCGG CTCGGGGGTG ACGCCGGTGAG AGCCCGGGGAC 290

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GATCCAGTTG CATGGNGGCT GTCACTGGAGA GCAT 34

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 155 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GATCGCTGGCT GTGCCCCCCT TTGCGCGGCA CGCCCCCCTG CCGACCCTTA CGAACAAAGC 60
TGCCCCTGTG TGCCAGCAGG CCCGCCACCC CGACCGCGGA GTGGGACCAAT GGCACCGGCG 120
TATCCCCACC ATTCGCCGCGA GNCACCCCGGCACCG 155

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 53 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

ATGGCGTTCA CGGCGCGGCG GGGACCCGGG AGCCCGGGG TGG 53

(2) INFORMATION FOR SEQ ID NO:41:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 132 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

GATCCACCGCC GGTTGCGAGAC CCGCCGCGCG GGCCCACCCCC GCACCGGCGCCC GGCAAGGCAGCAGG 60
GAACCGGCGGG CCAAGGCAGGCC AGGACACCCG AGGCGCGGGGG GGCAGCGCCGCA CCGCCGGGCA 120
AGGGCGGCAA CG 132

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 132 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

GATCGGGCCGGC CGCAGACCGNC GGGGACCGGCG GCAAGGCGCGGC NAACCGGGGCG GCCGNACCGCA 60
CCNAGCAAGA ATTTTCGCGNG TCCCCAATG TGCGGAATGG CCGACAGGGC GGCAACGGGC 120
GCACGGCGCCCA 132

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 702 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CGCAGGCGCAAGG ATCGGTACCC CCGCGACCTCG GCGAGCTGGCG ATTCCCACCGGG GGGGCGGGGTA 60
CGAGAAGGCCG CGTGATACCC TGCGAGCTGCC GAATTTAGGCC GATCGCTTCG CGATGCGCGGCC 120
ATGAGCGGGCC GCATCAATTC TAGTGAGCCAG ACCTTTTCAAT TTACCAGCGA TAATGGCTAT 180
AGACTAAGGG CAAATGACCA CTAATAGCAGC AGTCTGACAGC GCTGATGGGTG ATGAGGCGACA 240
ATGAATTTGAA CAGGGCGGACG GAGTGAGGAGG CCCCCAGTTGCC GAGCCCGCAGC ACTGATTGCT 300
CCCTACCAACC GTCGAGACTTC ACGNAGGNTA AAAAAACGCC CCACAGGNTG ATATTTGGCGA 360
CCGCGGAAATCAT CTGGCGGCAGGC GTGCCCAAGCG GGGCGCGGCG CTGGGACCTCTGGCTGGCGGGAC 420
CGCGCGCGCAAG GNTATCGGCG AGTCTGATGCA GAGGAGCTGC CCGGCGGCTGGC 480
ACGAGGCAACCAT GGAATGACCC GCGCGCGGGAG CACCGGCGGGA CCGGACAGCTCTGGGGGGGGG 540
CGCGCAACTTACCACCGG CGCGGAGGCGC CGCGGCGGCC CGCAACCGGGT CGAACCGCTTC CCGGCGGCTGGC 600
TCAAGGAAAG GCAGAAGGAG CCTGAGAAGCG GCGGCAAGCG CGATCGCTGTC CGCGGCGGCTGGC 660
GGGATGCGGTG AACAACCTCINC ACCCTGAGCCGC TGCAAGCCGCA CG 702

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 298 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

GAAGGCCGGAG CGCTGCTGGG CGACCTGCGG GTCAAAGGGG CATCGCTCGG TGCCGGTGGA 60
GCCGCGGGGG TGCCGCTGCC GCCTGGTGAAG TGGGCGAGTG GGGGCGCGGA ATCGTGTCGG 120
CCCGCTGCCG CGGTGTCAGAT TGGCCGCTTTA GCCAGGGGAA GCCGGCGCGG CGCCGGCGCC 180
CTGGGCGGGG GTGGCATGGG AATGCCGATG GGTGCCGGCC ATCACGGCAAG AGGGGCGCC 240
AAGTCCAGAG GTTCTCGAGA GGAAGACGAG GCCGTCTACA CCAGGATCC TCGTGGCG 298

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1058 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

CGGCACAGGG ATCGAATTCC GTCGCCGGGA GCACAGACGC GCACAGGCCG AGTGGAGGAG 60
CCATGACCTA CTCGCGCGGT ACGCGCGGAT ACGCGCGCGA GCGAGCGCGA GCGCTCTAGC 120
GAGCCGCTAC ACCCTGATGC GCCACGCGGG ATGGGCAGCC GAGAAAGTAC CGTGATCAC 180
TCGAACATCG GCCAGGTTGG CTGGCTCTGG ATCCTGCATG ATCCTGCAGC ACTGGGGTGC 240
CCGCTCGGAT GGTCTCGTAT GCTGCGCTGC TTGCGGCGGT GTTCTCGTG CCTAGGGCCA 300
AGAGCCATGT GACGAGTAAG GCGTGTCTGC GGTACGCGG GTTATCTCGG ATGGTCTCGG 360
CGACGTGTTAA CAAGCGCGAC GCCTATTCGA CGGTGTTGGG AATGGGGTTT GTGTTGCGTT 420
TCACTGGTAT CCAGGCAGGG GCGGCGGTCC TGGCCTCTGT GTGGGAGACC GCGCGTATCA 480
CCGGCCGGGC GCCGCGCGGC AGTGGAGGG CAGCGCGCGA GTACGCGCGG ATCAGGGCG 540
ACGCCGATGA CGGTTCCGAG CGGTGTTGGC ACTACGTTCA GCACGGGTGT CAGACGCGGC 600
CGGAGCTTGC ATCAGCGCGC CGGAGCCTAG CTAGGGTTGA GTTACCAGGG CAGCGCGGC 660
ACCGAGCTGA GGCTCTGTCA CAGGGCGTAT CGGGCGCGGG GGGGCGCGGG CAGCAGCGG 720
CCGCCGCTAC TCGGCGCGGT CGCGCGCGCG CCAGCGCCGC GCACGCGCGC TCGGCGCGGG 780
CGGCGGGAT CGGGTCAGGC GCAGGGGTCC ATACGCGGAC GGGGGTATCC AGCAGCGCA 840
CGGCGGGGAT GACGAGCTGAC CGGTACGGGC CGGAGCGGAG TCGGGCGGCG CAGCGCGCGC 900
GTCGCGCCGC GCACGAGCTG CGCGCGCGCG CGCGCGCGCG CAGCGCGCGG CGCGCGCGCG 960
GCCGCGCGGT CTAACGGCGG GTTCCGCGGT CGCGCGCGCG GGTGGCGGCGA AGAGTGGAAC 1020
GGGTGTCAAG CGGCGCGCGA GATCGCGTG GCGCAAATT 1058

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 327 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

CGCCGAGAGA GACCGAGGCC GCTACCTCGG GCACAGGAGG AGGTAATTTT CAGGGGTCTT 60
CCGGCGACTT GAATCAGCGG ATCAGCGCGG TGGAGTCCAC GCGAGGTTCC TGGGAGGGGC 120
AGTGGCGCGC GCGCGCGCGG AGCGCGCGCG AGCGCGCGCG GCTGGCGCTTC TAAAGAAGGC 180
CAATAAGCA GAAGCGGAGA TCGAGCAGAG TTTGCGGCAA TATCGCTGCG GCAGGGGCTCC 240
AATACCTCAG GCGCGCGAGG GAGCAGCGAG AGCGCAGGCT CTCGCAAATG GCAGTTCTGC 300
CCGGTAAATG GAATAGAAAGA CGGGCGAAG 327
(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 170 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CGGTGCGGAT GATGGCGTTG TCGAACGTGA CCGATTCTGT ACGCGCTGCG TTGAGATCAA 60
CCAACAACGT GTGGCCGTGC GCAAAATGTCG CGNACCCGTTG ATCTCGGTGCT 120
TCTTCTACTCA GAACTGCGCA CCGGCACCC TGCCCTCCGGN TAGCTTTGCGG 170

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 127 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GATCCGGGCGC CAGCGGGGCTT GCGCGGGGCA GCACCGCTTG CGCGCTGGGCG ACGCGGGGGG 60
CGCGGGGGTGG CCGCGGAAAC CGTGCGGTGC TCTTCTGGCAA CGCGCGTGCC GCGGGCGACG 120
GCGCGGTG 127

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

CGCGCGCAAG GCGCGCCACCG CCGCGCGCG CCGCGCGGCC ACGCGCGCGA 60
CGCGCGCTCC GGCTTCAACGG G 81

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 149 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

GATCAGGGCT GGCGCGCTCC GGCCCGAAGG GGCTAAGGG AGGAGCTGCC GGAATGTTTG 60
GCAACGGCGG GCGCGGCGGT CCGCGCGCGT CCAACCAAGCC CGCTAACGCC GGGNCGGCCG 120
GAAACGCGTG TGCCCGGTGG CGATCTTG 149
(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

CGGCACGAGA TCACACCTAC CGAGTGATCG AGATGCTCGG GACCTCGCCC GACGGGCTCG 60
ACGGCGAAT CTGGCGGGGT CTCGCCCCAG CTGGCGCAGAC CATCGCGGCG CATGGAAGT 120
TGCAAGTACA GTCAATTTGA GGCACCCTGG TGCAAGGGAC GTGCGCCACGC TTCCAGGGT 180
CTATAGAAGT CCGGTCCGCC CTGGAGGATT CCTGAACTCTT CAAGCCGCGG CGATAACTGA 240
GGTGCACTAT TAAAGGACCTT TTCCAGAACA TCTCTGACCG CTGGAAACGC GGGTGAGCGG 300
ACGGTGCCCT CGCGCGGCGG CTGGCCTCCA AATCCCTCGG ACAATTCGTC GCCGCG 355

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 999 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

ATGCATCACC ATCACCATCA CATGCATCAG GTGGGCCCCCA ACTTGACACG TCGCAAGGGG 60
CGATGCCGCG CACTGCTAT CGCGGGCGATG GCCAGCGCCA GCTGTGGTAC GCTGCGGTGTG 120
CCGGCGGCGG CCGAGGGCAG CCCGGGCGCA GCCGCCCCGG GGAGAGACGC TACCCCAAC 180
CCGGCGGCGG CCGGGCGCCAG CCGGGCGCCG CCAGAGACGC TCGGCGGCGC CCGGGGCTAT 240
GCCGCGGCGG ACAGGCAGGA TGGGGGCGGG GCCAGCCAAA AGCCAGACGC CTGGGGCGCC 300
GGAGACAGCG GGGCGGCTGG CAAATGACCC ATCTGCTCGG CTGGGCTGATG GAGCCAAAGG 360
CTTACGCGCA GCCAGCGAAG CCAGCGCCTC AAGCCAGCGC CGCGTGGTGG CTCGGACATG 420
GGTGAGTCCT ATATGCCCTTA CCAGGGCGCC CGAGATCCGC AGGAAACGGT CTCGGCTCGC 480
GCAAGCCGGG TGCTCTGCAG GGGCTCTGAT TACGAGTCAA ATCTAGCGGA TCGGAGGCA 540
CGGACAGGGG AGATCTGGAC GGGGCTTATGC GGGCTGGCCC GCGGAGACGC AGCGGACCACG 600
GGGGGGGCTCC AGGCGTGGTT TTGGGGTGAG CTGGGAGGCG CGAAGCCAGC GCCAGGACTG 660
GGGCGGGCGG CGAGGCTTGG CCGCGGCGCG CGCGGCGGGC CCGGGGACGC CCGGGCCGCG 720
GGGGGGGGGG CACAGTGGGC CGGAGGCGAG CCGGCGGCGC CGGGGAGACG CCGGGGGGGC 780
GGGGGGGCGG CGAGGCTTGG CCGGAGGCGAG CCGGCGGCGC CGGGGAGACG CCGGGGGGGC 840
GGGGGGGCGG CGAGGCTTGG CCGGAGGCGAG CCGGCGGCGC CGGGGAGACG CCGGGGGGGC 900
CGGACGACAC CGACGACGCA GCGAGCTTCA CCGGGCCTCA 960

(2) INFORMATION FOR SEQ ID NO:53:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 332 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:
Met His His His His His Gln Val Asp Pro Asn Leu Thr
1  5     10     15
Arg Arg Lys Gly Arg Leu Ala Ala Leu Ala Ala Ile Ala Ala Met Ala Ser
20  25    30
 Ala Ser Leu Val Thr Val Ala Val Pro Ala Thr Ala Asn Ala Asp Pro
35  40    45
 Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro Ser Thr
50  55    60
 Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro Val Ala Pro Pro Pro Pro
65  70    75    80
 Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro Gly Asp Pro Asn Ala Ala
85  90    95
 Pro Pro Pro Ala Asp Pro Asn Ala Pro Pro Pro Pro Val Ile Ala Pro
100 105   110
 Asn Ala Pro Gln Pro Val Arg Ile Asp Asn Pro Val Gly Gly Phe Ser
115 120   125
 Phe Ala Leu Pro Ala Gly Trp Val Glu Ser Asp Ala Ala His Phe Asp
130 135   140
 Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr Gly Asp Pro Pro Phe Pro
145 150   155   160
 Gly Gln Pro Pro Pro Val Ala Asn Asp Thr Arg Ile Val Leu Gly Arg
165 170   175
 Leu Asp Gln Lys Leu Tyr Ala Ser Ala Gln Ala Thr Asp Ser Lys Ala
180 185   190
 Ala Ala Arg Leu Gly Ser Asp Met Gly Glu Phe Tyr Met Pro Tyr Pro
195 200   205
 Gly Thr Arg Ile Asn Gln Glu Thr Val Ser Leu Asp Ala Asn Gly Val
210 215   220
 Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys
225 230   235   240
 Pro Asn Gly Gln Ile Trp Thr Gly Val Ile Gly Ser Pro Ala Ala Asn
245 250   255
 Ala Pro Asp Ala Gly Pro Pro Glu Arg Thr Phe Val Val Trp Leu Gly
260 265   270
 Thr Ala Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Gln
275 280   285
 Ser Ile Arg Pro Leu Val Ala Pro Pro Pro Ala Pro Ala Ala Pro
290 295   300
 Ala Glu Pro Ala Pro Ala Pro Ala Gln Gln Val Ala Pro Thr
305 310   315
 Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala
325 330

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:
Asp Pro Val Asp Ala Val Ile Asn Thr Thr Xaa Asn Tyr Gly Gln Val
<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Val Ala Ala Leu</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Ala Val Glu Ser Gly Met Leu Ala Leu Gly Thr Pro Ala Pro Ser

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 19 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala Ala Lys Glu Gly Arg

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp Pro Ala Trp Gly Pro

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>15</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:58:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 14 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:58:
Asp Ile Gly Ser Glu Ser Thr Glu Asp Gln Gln Xaa Ala Val
1 5 10

(2) INFORMATION FOR SEQ ID NO:59:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:59:
Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro
1 5 10

(2) INFORMATION FOR SEQ ID NO:60:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:60:
Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Ala Ala Ala Ala Pro Pro
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:61:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:61:
Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:62:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:62:
Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Gln Thr Ser
1 5 10 15
(2) INFORMATION FOR SEQ ID NO:63:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 187 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:63:

<table>
<thead>
<tr>
<th>Leu</th>
<th>Leu</th>
<th>Asn</th>
<th>Asn</th>
<th>Leu</th>
<th>Ala</th>
<th>Asp</th>
<th>Pro</th>
<th>Asp</th>
<th>Val</th>
<th>Ser</th>
<th>Phe</th>
<th>Ala</th>
<th>Asp</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

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

(2) INFORMATION FOR SEQ ID NO:64:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 148 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:64:

<table>
<thead>
<tr>
<th>Asp</th>
<th>Glu</th>
<th>Val</th>
<th>Thr</th>
<th>Val</th>
<th>Glu</th>
<th>Thr</th>
<th>Ser</th>
<th>Val</th>
<th>Phe</th>
<th>Arg</th>
<th>Ala</th>
<th>Asp</th>
<th>Phe</th>
<th>Leu</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>10</td>
<td></td>
<td></td>
<td>15</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Ser Glu Leu Asp Ala Pro Ala Gln Ala Gly Thr Glu Ser Ala Val Ser
20  25   30  
Gly Val Glu Gly Leu Pro Pro Gly Ser Ala Leu Leu Val Val Lys Arg
35  40   45  
Gly Pro Asn Ala Gly Ser Arg Phe Leu Leu Asp Gln Ala Ile Thr Ser
50 55 60
Ala Gly Arg His Pro Asp Ser Asp Ile Phe Leu Asp Asp Val Thr Val
65 70 75 80
Ser Arg Arg His Ala Glu Phe Arg Leu Glu Glu Asn Glu Phe Asn Val
85 90 95
Val Asp Val Gly Ser Leu Asn Gly Thr Tyr Val Asn Arg Glu Pro Val
100 105 110
Asp Ser Ala Val Leu Ala Asn Gly Asp Glu Val Gln Ile Gly Lys Leu
115 120 125
Arg Leu Val Phe Leu Thr Gly Pro Lys Gln Gly Glu Asp Asp Gly Ser
130 135 140
Thr Gly Gly Pro
145

(2) INFORMATION FOR SEQ ID NO:65:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 230 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:65:

Thr Ser Asn Arg Pro Ala Arg Arg Gly Arg Arg Ala Pro Arg Asp Thr
1 5 10 15
Gly Pro Asp Arg Ser Ala Ser Leu Ser Leu Val Arg His Arg Arg Glu
20 25 30
Gln Arg Asp Ala Leu Cys Leu Ser Ser Thr Gln Ile Ser Arg Glu Ser
35 40 45
Asn Leu Pro Pro Ala Ala Gly Gly Ala Ala Asn Tyr Ser Arg Arg Asn
50 55 60
Phe Asp Val Arg Ile Lys Ile Phe Met Leu Val Thr Ala Val Val Leu
65 70 75 80
Leu Cys Cys Ser Gly Val Ala Thr Ala Pro Lys Thr Tyr Cys Glu
85 90 95
Glu Leu Lys Gly Thr Asp Thr Gly Gln Ala Cys Gln Ile Glu Gln Met Ser
100 105 110
Asp Pro Ala Tyr Asn Ile Asn Ile Ser Leu Pro Ser Tyr Tyr Pro Asp
115 120 125
Gln Lys Ser Leu Glu Asn Tyr Ile Ala Gln Thr Arg Asp Lys Phe Leu
130 135 140
Ser Ala Ala Thr Ser Ser Thr Pro Arg Glu Ala Pro Tyr Glu Leu Asn
145 150 155 160
Ile Thr Ser Ala Thr Tyr Gln Ser Ala Ile Pro Arg Gly Thr Gln
165 170 175
Ala Val Val Leu Xaa Val Tyr His Asn Ala Gly Gly Thr His Pro Thr
180 185 190
Thr Thr Tyr Lys Ala Phe Asp Trp Asp Gln Ala Tyr Arg Lys Pro Ile
195 200 205
Thr Tyr Asp Thr Leu Trp Gln Ala Asp Thr Asp Pro Leu Pro Val
210 215 220
Phe Pro Ile Val Ala Arg
225 230
(2) INFORMATION FOR SEQ ID NO:66:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 132 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:66:

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

(2) INFORMATION FOR SEQ ID NO:67:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:67:

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

(2) INFORMATION FOR SEQ ID NO:68:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 163 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:68:

Met Thr Asp Asp Ile Leu Leu Ile Asp Thr Asp Glu Arg Val Arg Thr
Leu Thr Leu Asn Arg Pro Gln Ser Arg Asn Ala Leu Ser Ala Ala Leu
Arg Asp Arg Phe Phe Ala Xaa Leu Xaa Asp Ala Glu Xaa Asp Asp Arg
Ile Asp Val Val Ile Leu Thr Gly Ala Asp Pro Val Phe Cys Ala Gly
Leu Asp Leu Lys Val Ala Gly Arg Ala Asp Arg Ala Ala Gly His Leu
Thr Ala Val Gly Gly His Asp Gln Ala Gly Asp Arg Arg Asp Gln Arg
Arg Arg Gly His Arg Arg Ala Arg Thr Gly Ala Val Leu Arg His Pro
Asp Arg Leu Arg Ala Arg Pro Leu Arg His Pro Arg Pro Gly Gly
 Ala Ala Ala His Leu Gly Thr Gln Cys Val Leu Ala Ala Lys Gly Arg
His Arg Xaa Gly Pro Val Asp Glu Pro Arg Leu Pro Asp Arg
Asp Arg Arg

(2) INFORMATION FOR SEQ ID NO:69:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 344 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:69:

Met Lys Phe Val Asn His Ile Glu Pro Val Ala Pro Arg Arg Ala Gly
Gly Ala Val Ala Glu Val Tyr Ala Glu Ala Arg Glu Phe Gly Arg
Leu Pro Glu Pro Leu Ala Met Leu Ser Pro Asp Glu Gly Leu Leu Thr
 Ala Gly Trp Ala Thr Leu Arg Glu Thr Leu Leu Val Gly Gln Val Pro
Arg Gly Arg Lys Glu Ala Val Ala Ala Val Ala Ala Ser Leu Arg
Cys Pro Trp Cys Val Asp Ala His Thr Thr Met Leu Tyr Ala Ala Gly
Gln Thr Asp Thr Ala Ala Ile Leu Ala Gly Thr Ala Pro Ala Ala

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

(2) INFORMATION FOR SEQ ID NO:70:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 485 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:70:

Asp Asp Pro Asp Met Pro Gly Thr Val Ala Lys Ala Val Ala Asp Ala
1  5  10  15
Leu Gly Arg Gly Ile Ala Pro Val Glu Asp Ile Gln Asp Cys Val Glu
20  25  30
Ala Arg Leu Gly Glu Ala Gly Leu Asp Val Ala Arg Val Tyr Ile
35  40  45
Ile Tyr Arg Gln Arg Arg Ala Glu Leu Arg Thr Ala Lys Ala Leu Leu
50  55  60
Gly Val Arg Asp Glu Leu Lys Ser Leu Ala Ala Val Thr Val Leu
65  70  75  80
Arg Gly Arg Tyr Leu Leu His Asp Glu Gln Gly Arg Pro Ala Glu Ser
85  90  95
Thr Gly Glu Leu Met Asp Arg Ser Ala Arg Cys Val Ala Ala Ala Glu
100  105  110
<table>
<thead>
<tr>
<th>Asp</th>
<th>Gln</th>
<th>Tyr</th>
<th>Glu</th>
<th>Pro</th>
<th>Gly</th>
<th>Ser</th>
<th>Ser</th>
<th>Arg</th>
<th>Arg</th>
<th>Trp</th>
<th>Ala</th>
<th>Glu</th>
<th>Arg</th>
<th>Phe</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Leu</td>
<td>Leu</td>
<td>Arg</td>
<td>Asn</td>
<td>Leu</td>
<td>Glu</td>
<td>Phe</td>
<td>Leu</td>
<td>Pro</td>
<td>Asn</td>
<td>Ser</td>
<td>Pro</td>
<td>Thr</td>
<td>Leu</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Ser</td>
<td>Gly</td>
<td>Thr</td>
<td>Asp</td>
<td>Leu</td>
<td>Gly</td>
<td>Leu</td>
<td>Ala</td>
<td>Gly</td>
<td>Cys</td>
<td>Phe</td>
<td>Val</td>
<td>Leu</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Glu</td>
<td>Asp</td>
<td>Ser</td>
<td>Leu</td>
<td>Gln</td>
<td>Ser</td>
<td>Ile</td>
<td>Phe</td>
<td>Ala</td>
<td>Thr</td>
<td>Leu</td>
<td>Gly</td>
<td>Gln</td>
<td>Ala</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Leu</td>
<td>Gln</td>
<td>Arg</td>
<td>Ala</td>
<td>Gly</td>
<td>Gly</td>
<td>Gly</td>
<td>Thr</td>
<td>Gly</td>
<td>Tyr</td>
<td>Ala</td>
<td>Phe</td>
<td>Ser</td>
<td>His</td>
<td>Leu</td>
</tr>
<tr>
<td>180</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Pro</td>
<td>Ala</td>
<td>Gly</td>
<td>Asp</td>
<td>Arg</td>
<td>Val</td>
<td>Ala</td>
<td>Ser</td>
<td>Thr</td>
<td>Gly</td>
<td>Thr</td>
<td>Gly</td>
<td>Thr</td>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td>195</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Val</td>
<td>Ser</td>
<td>Phe</td>
<td>Leu</td>
<td>Arg</td>
<td>Leu</td>
<td>Tyr</td>
<td>Asp</td>
<td>Ser</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Val</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Gly</td>
<td>Gly</td>
<td>Arg</td>
<td>Arg</td>
<td>Gly</td>
<td>Ala</td>
<td>Cys</td>
<td>Met</td>
<td>Ala</td>
<td>Val</td>
<td>Leu</td>
<td>Asp</td>
<td>Val</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>His</td>
<td>Pro</td>
<td>Asp</td>
<td>Ile</td>
<td>Cys</td>
<td>Phe</td>
<td>Val</td>
<td>Thr</td>
<td>Ala</td>
<td>Lys</td>
<td>Ala</td>
<td>Glu</td>
<td>Ser</td>
<td>Pro</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>245</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Leu</td>
<td>Pro</td>
<td>His</td>
<td>Phe</td>
<td>Asn</td>
<td>Leu</td>
<td>Ser</td>
<td>Val</td>
<td>Gly</td>
<td>Thr</td>
<td>Asp</td>
<td>Ala</td>
<td>Phe</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Ala</td>
<td>Val</td>
<td>Glu</td>
<td>Arg</td>
<td>Asn</td>
<td>Leu</td>
<td>His</td>
<td>Arg</td>
<td>Leu</td>
<td>Val</td>
<td>Asn</td>
<td>Pro</td>
<td>Arg</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>275</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Lys</td>
<td>Ile</td>
<td>Val</td>
<td>Ala</td>
<td>Arg</td>
<td>Met</td>
<td>Pro</td>
<td>Ala</td>
<td>Ala</td>
<td>Glu</td>
<td>Leu</td>
<td>Phe</td>
<td>Asp</td>
<td>Ala</td>
<td>Ile</td>
</tr>
<tr>
<td>290</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Lys</td>
<td>Ala</td>
<td>Ala</td>
<td>His</td>
<td>Ala</td>
<td>Gly</td>
<td>Asp</td>
<td>Gly</td>
<td>Leu</td>
<td>Val</td>
<td>Phe</td>
<td>Leu</td>
<td>Asp</td>
<td></td>
<td></td>
</tr>
<tr>
<td>305</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Ile</td>
<td>Asn</td>
<td>Arg</td>
<td>Ala</td>
<td>Gly</td>
<td>Pro</td>
<td>Val</td>
<td>Pro</td>
<td>Gly</td>
<td>Arg</td>
<td>Gly</td>
<td>Arg</td>
<td>Ile</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td>325</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asn</td>
<td>Pro</td>
<td>Cys</td>
<td>Gly</td>
<td>Val</td>
<td>Pro</td>
<td>Leu</td>
<td>Leu</td>
<td>Pro</td>
<td>Tyr</td>
<td>Glu</td>
<td>Ser</td>
<td>Cys</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>340</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Gly</td>
<td>Ser</td>
<td>Ile</td>
<td>Asn</td>
<td>Leu</td>
<td>Ala</td>
<td>Arg</td>
<td>Met</td>
<td>Leu</td>
<td>Ala</td>
<td>Asp</td>
<td>Gly</td>
<td>Arg</td>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td>355</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Asp</td>
<td>Arg</td>
<td>Leu</td>
<td>Glu</td>
<td>Val</td>
<td>Ala</td>
<td>Gly</td>
<td>Val</td>
<td>Ala</td>
<td>Val</td>
<td>Arg</td>
<td>Phe</td>
<td>Leu</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>370</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Val</td>
<td>Ile</td>
<td>Asp</td>
<td>Val</td>
<td>Ser</td>
<td>Arg</td>
<td>Tyr</td>
<td>Pro</td>
<td>Phe</td>
<td>Pro</td>
<td>Glu</td>
<td>Leu</td>
<td>Gly</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td>385</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Arg</td>
<td>Ala</td>
<td>Thr</td>
<td>Arg</td>
<td>Lys</td>
<td>Ile</td>
<td>Gly</td>
<td>Leu</td>
<td>Gln</td>
<td>Gly</td>
<td>Val</td>
<td>Met</td>
<td>Gly</td>
<td>Leu</td>
<td>Ala</td>
</tr>
<tr>
<td>405</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Leu</td>
<td>Gly</td>
<td>Ile</td>
<td>Pro</td>
<td>Tyr</td>
<td>Asp</td>
<td>Ser</td>
<td>Glu</td>
<td>Glu</td>
<td>Ala</td>
<td>Val</td>
<td>Arg</td>
</tr>
<tr>
<td>420</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Ala</td>
<td>Thr</td>
<td>Arg</td>
<td>Leu</td>
<td>Met</td>
<td>Arg</td>
<td>Ile</td>
<td>Gln</td>
<td>Gln</td>
<td>Ala</td>
<td>Ala</td>
<td>His</td>
<td>Thr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>435</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Arg</td>
<td>Leu</td>
<td>Ala</td>
<td>Glu</td>
<td>Arg</td>
<td>Gly</td>
<td>Ala</td>
<td>Phe</td>
<td>Pro</td>
<td>Ala</td>
<td>Phe</td>
<td>Thr</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>450</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Arg</td>
<td>Phe</td>
<td>Ala</td>
<td>Arg</td>
<td>Ser</td>
<td>Gly</td>
<td>Pro</td>
<td>Arg</td>
<td>Asn</td>
<td>Ala</td>
<td>Gln</td>
<td>Val</td>
<td>Thr</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>465</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td>Pro</td>
<td>Thr</td>
<td>Gly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 267 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:72:

Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly Val Gln Val
1  5  10  15
Thr Asn Asp Lys Thr Pro Gly Ala Lys Ile Val Gly Val Val Ala
20  25  30
Val Gly Gly Ala Ala Asn Ala Gly Val Pro Lys Gly Val Val Val Thr
35  40  45
<table>
<thead>
<tr>
<th>Lys Val Asp Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu Val Ala Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
</tr>
<tr>
<td>55</td>
</tr>
<tr>
<td>60</td>
</tr>
<tr>
<td>Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr Phe Gln Asp</td>
</tr>
<tr>
<td>65</td>
</tr>
<tr>
<td>70</td>
</tr>
<tr>
<td>75</td>
</tr>
<tr>
<td>80</td>
</tr>
<tr>
<td>Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly Lys Ala Glu</td>
</tr>
<tr>
<td>85</td>
</tr>
<tr>
<td>90</td>
</tr>
<tr>
<td>95</td>
</tr>
<tr>
<td>Gln</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 364 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 73:

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

(2) INFORMATION FOR SEQ ID NO:74:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 309 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:74:

Gln Ala Ala Gly Arg Ala Val Arg Arg Thr Gly His Ala Glu Asp
  1  5  10  15
Gln Thr His Gln Asp Arg Leu His His Gly Cys Arg Arg Ala Ala Val
  20  25  30
Val Val Arg Glu Asp Ala Ser Val Ser Ala Thr Ser Ala Arg Pro
  35  40  45
Pro Arg Arg His Pro Ala Gln Gly His Arg Arg Arg Val Ala Pro Ser
  50  55  60
Gly Gly Arg Arg Pro His Pro His His Val Gln Pro Asp Arg
  65  70  75  80
Arg Asp Arg Pro Ala Leu Leu Asp Arg Thr Gln Pro Ala Glu His Pro
  85  90  95
Asp Pro His Arg Gly Pro Ala Asp Pro Gly Arg Val Arg Gly Arg
 100 105 110
Gly Arg Leu Arg Arg Val Asp Gly Arg Leu Gln Pro Asp Arg Asp
 115 120 125
Ala Asp His Gly Ala Pro Val Arg Gly Arg Gly Pro His Arg Gly Val
 130 135 140
Gln His Arg Gly Pro Val Phe Val Arg Arg Val Pro Gly Val Arg
 145 150 155 160
Cys Ala His Arg Gly His Arg Arg Val Ala Ala Pro Gly Gln Gly
 165 170 175
Asp Val Leu Arg Ala Gly Leu Arg Val Glu Arg Leu Arg Pro Val Ala
 180 185 190
Ala Val Glu Asn Leu His Arg Gly Ser Gln Arg Ala Asp Gly Arg Val
 195 200 205
Phe Arg Pro Ile Arg Arg Gly Ala Arg Leu Pro Ala Arg Arg Ser Arg
 210 215 220
Ala Gly Pro Gln Gly Arg Leu His Leu Asp Gln Ala Gly Pro Ser Pro
 225 230 235 240
Leu Pro Ala Arg Ala Gly Glu Gln Pro Ser Ser Ala Gly Gly Arg
 245 250 255
Arg Ala Gly Ala Glu Arg Ala Asp Pro Gly Gln Arg Gly Arg His
 260 265 270
His Glu Gly His Asp Pro Gly Arg Gln Gly Ala Gln Arg Gly Thr
275 Ala Gly Val Ala His Ala Ala Ala Gly Pro Arg Arg Ala Ala Val Arg
290 Asn Arg Pro Arg Arg
305
(2) INFORMATION FOR SEQ ID NO:75:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 580 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:75:

Ser Ala Val Trp Cys Leu Asn Gly Phe Thr Gly Arg His Arg His Gly
1  5 10 15
Arg Cys Arg Val Arg Ala Ser Gly Trp Arg Ser Ser Asn Arg Trp Cys
20 25 30
Ser Thr Thr Ala Asp Cys Cys Ala Ser Lys Thr Pro Thr Glu Ala Ala
35 40 45 50
Ser Pro Leu Glu Arg Arg Phe Thr Cys Cys Ser Pro Ala Val Gly Cys
55 60
Arg Phe Arg Ser Phe Pro Val Arg Arg Leu Ala Gly Arg Thr
65 70 75 80
Ser Arg Thr Leu Gly Val Arg Arg Thr Leu Ser Gln Trp Asn Leu Ser
85 90 95
Pro Arg Ala Gin Pro Ser Cys Ala Val Thr Val Glu Ser His Thr His
100 105 110
 Ala Ser Pro Arg Met Ala Lys Leu Ala Arg Val Val Gly Leu Val Gin
115 120 125
Glu Glu Gin Pro Ser Asp Met Thr Asn His Pro Arg Tyr Ser Pro Pro
130 135 140
Pro Gin Gln Pro Gly Thr Pro Gly Tyr Ala Gin Gly Gln Gln Thr Gln
145 150 155 160
Tyr Ser Gin Gln Pro Phe Asp Trp Arg Tyr Pro Ser Pro Pro Pro Gin
165 170 175
Pro Thr Gin Tyr Arg Glu Pro Tyr Glu Ala Leu Gly Gly Thr Arg Pro
180 185 190
Gly Leu Ile Pro Gly Val Ile Pro Thr Met Thr Pro Pro Pro Gly Met
195 200 205
Val Arg Gln Arg Pro Arg Ala Gly Met Leu Ala Ile Gly Ala Val Thr
210 215 220
Ile Ala Val Val Ser Ala Gly Ile Gly Gly Ala Ala Ala Ser Leu Val
225 230 235 240
Gly Phe Asn Arg Ala Pro Ala Gly Pro Ser Gly Gly Pro Val Ala Ala
245 250 255
Ser Ala Ala Pro Ser Ile Pro Ala Ala Asn Met Pro Pro Gly Ser Val
260 265 270
Glu Gin Val Ala Ala Lys Val Val Pro Ser Val Val Met Leu Glu Thr
275 280 285
Asp Leu Gly Arg Gin Ser Glu Glu Gly Ser Gly Ile Ile Leu Ser Ala
290 295 300
Glu Gly Leu Ile Leu Thr Asn Asn His Val Ile Ala Ala Ala Ala Lys
305  Pro Pro Leu Gly Ser Pro Pro Pro Pro Lys Thr Thr Val Thr Phe Ser Asp
     325  330  335
310  Gly Arg Thr Ala Pro Phe Thr Val Val Gly Ala Asp Pro Thr Ser Asp
     340  345  350
315  Ile Ala Val Val Arg Val Gln Gly Val Ser Gly Leu Thr Pro Ile Ser
     355  360  365
320  Leu Gly Ser Ser Ser Asp Leu Arg Val Gly Gln Pro Val Leu Ala Ile
     370  375  380
325  Gly Ser Pro Leu Gly Leu Glu Gly Thr Val Thr Thr Gly Ile Val Ser
     385  390  395  400
330  Ala Leu Asn Arg Pro Val Ser Thr Thr Gly Glu Ala Gly Asn Gln Asn
     405  410  415
335  Thr Val Leu Asp Ala Ile Gln Thr Asp Ala Ala Ile Asn Pro Gly Asn
     420  425  430  435
340  Ser Gly Gly Ala Leu Val Asn Met Asn Ala Gln Leu Val Gly Val Asn
     440  445
345  Ser Ala Ile Ala Thr Leu Gly Ala Asp Ser Ala Asp Ala Gln Ser Gly
     450  455  460
350  Ser Ile Gly Leu Gly Phe Ala Ile Pro Val Asp Gln Ala Lys Arg Ile
     465  470  475  480
355  Ala Asp Glu Leu Ile Ser Thr Gly Lys Ala Ser His Ala Ser Leu Gly
     485  490  495
360  Val Gln Val Thr Asn Asp Lys Asp Thr Pro Gly Ala Lys Ile Val Glu
     500  505  510
365  Val Val Ala Gly Gly Ala Ala Asn Ala Gly Val Pro Lys Gly Val
     515  520  525
370  Val Val Thr Lys Val Asp Arg Pro Ile Asn Ser Ala Asp Ala Leu
     530  535  540
375  Val Ala Ala Val Arg Ser Lys Ala Pro Gly Ala Thr Val Ala Leu Thr
     545  550  555  560
380  Phe Gln Asp Pro Ser Gly Gly Ser Arg Thr Val Gln Val Thr Leu Gly
     565  570  575
385  Lys Ala Glu Gln
     580

(2) INFORMATION FOR SEQ ID NO:76:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 233 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:76:

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

(2) INFORMATION FOR SEQ ID NO:77:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 66 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:77:

Val Ile Asp Ile Ile Gly Thr Ser Pro Thr Ser Thr Trp Glu Gln Ala Ala
1  5 10 15
Ala Glu Ala Val Gln Arg Ala Arg Asp Ser Val Asp Asp Ile Arg Val
20 25 30
Ala Arg Val Ile Glu Gln Asp Met Ala Val Asp Ser Ala Gly Lys Ile
35 40
Thr Tyr Arg Ile Lys Leu Glu Val Ser Phe Lys Met Arg Pro Ala Gln
50 55 60
Pro Arg
65

(2) INFORMATION FOR SEQ ID NO:78:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 69 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:78:

Val Pro Pro Ala Pro Pro Leu Pro Pro Leu Pro Pro Ser Pro Ile Ser
1  5 10 15
Cys Ala Ser Pro Pro Ser Pro Pro Leu Pro Pro Ala Pro Pro Val Ala
20 25 30
Pro Gly Pro Pro Met Pro Pro Leu Asp Pro Trp Pro Pro Ala Pro Pro
35 40 45
Leu Pro Tyr Ser Thr Pro Pro Gly Ala Pro Leu Pro Pro Ser Pro Pro
50 55 60
Ser Pro Pro Leu Pro
65

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 355 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO: 79:

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

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 205 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 80:

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

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 286 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:81:

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

(2) INFORMATION FOR SEQ ID NO:82:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 173 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:82:

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

(2) INFORMATION FOR SEQ ID NO:83:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:83:

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

(2) INFORMATION FOR SEQ ID NO:84:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 125 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:84:

Val Leu Ser Val Pro Val Gly Asp Gly Phe Trp Xaa Arg Val Val Val Asn
1 5 10 15
Pro Leu Gly Gln Pro Ile Asp Gly Arg Gly Asp Val Asp Ser Asp Thr
(2) INFORMATION FOR SEQ ID NO:85:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 117 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:85:

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

(2) INFORMATION FOR SEQ ID NO:86:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 103 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:86:

Met Tyr Arg Phe Ala Cys Arg Thr Leu Met Leu Ala Ala Cys Ile Leu
1  5  10  15
Ala Thr Gly Val Ala Gly Leu Gly Val Gly Ala Gln Ser Ala Ala Gln
20 25 30
Thr Ala Pro Val Pro Asp Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp
35
Pro Ala Trp Gly Pro Asn Trp Asp Pro Tyr Thr Cys His Asp Asp Phe
40
His Arg Asp Ser Asp Gly Pro Asp His Ser Arg Asp Tyr Pro Gly Pro
55
Ile Leu Glu Gly Pro Val Leu Asp Asp Pro Gly Ala Ala Pro Pro
70
90
Pro Ala Ala Gly Gly Gly Ala
100

(2) INFORMATION FOR SEQ ID NO:87:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 88 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:87:

Val Gln Cys Arg Val Trp Leu Glu Ile Gln Trp Arg Gly Met Leu Gly
1      5 10 15
Ala Asp Gln Ala Arg Ala Gly Gly Pro Ala Arg Ile Trp Arg Gly His
20     25 30
Ser Met Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala
35     40 45
Thr Lys Glu Gly Arg Gly Ile Val Met Arg Val Pro Leu Glu Gly Gly
50     55 60
Gly Arg Leu Val Val Glu Leu Thr Pro Asp Glu Ala Ala Ala Leu Gly
65     70 75 80
Asp Glu Leu Lys Gly Val Thr Ser
85

(2) INFORMATION FOR SEQ ID NO:88:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 95 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:88:

Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile
1      5 10 15
Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly
20     25 30
Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala
35     40 45
 Ala Val Val Arg Phe Gin Glu Ala Ala Asn Lys Gin Lys Glu Leu
50     55 60
Asp Glu Ile Ser Thr Asn Ile Arg Gin Ala Gly Val Gln Tyr Ser Arg
65     70 75 80
Ala Asp Glu Glu Gin Gin Gin Ala Leu Ser Ser Gin Met Gly Phe
(2) INFORMATION FOR SEQ ID NO:89:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 166 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:89:

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

(2) INFORMATION FOR SEQ ID NO:90:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 5 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:90:

Arg Ala Glu Arg Met 1 5

(2) INFORMATION FOR SEQ ID NO:91:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 263 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:91:

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

(2) INFORMATION FOR SEQ ID NO:92:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 303 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:92:

Met Thr Tyr Ser Pro Gly Asn Pro Gly Tyr Pro Gln Ala Gln Pro Ala
      1      5     10     15
Gly Ser Tyr Gly Gly Val Thr Pro Ser Phe Ala His Ala Asp Glu Gly
     20     25     30
Ala Ser Lys Leu Pro Met Tyr Leu Asn Ile Ala Val Ala Val Leu Gly
     35     40     45
Leu Ala Ala Tyr Phe Ala Ser Phe Gly Pro Met Phe Thr Leu Ser Thr
(2) INFORMATION FOR SEQ ID NO:93:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 28 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:93:

Gly Cys Gly Glu Thr Asp Ala Ala Thr Leu Ala Gln Gln Ala Gly Asn
  1  5  10
Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile
  20  25

(2) INFORMATION FOR SEQ ID NO:94:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 16 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:94:

Asp Glu Val Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:95:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:95:

Gly Cys Gly Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala
1 5 10 15
Ala Gly Thr Ala Ala Gln Ala Ala Val Val Arg
20 25

(2) INFORMATION FOR SEQ ID NO:96:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:96:

Gly Cys Gly Gly Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu
1 5 10 15
Ala Ala Asn Lys Gln Lys Gln Glu Leu Asp Glu
20 25

(2) INFORMATION FOR SEQ ID NO:97:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 27 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:97:

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

(2) INFORMATION FOR SEQ ID NO:98:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 28 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:98:

Gly Cys Gly Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg Ala Asp Glu
1       5       10    15
Glu Gln Gln Gln Ala Leu Ser Ser Glu Met Gly Phe
20       25

(2) INFORMATION FOR SEQ ID NO:99:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 507 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:99:

ATGAAGATGG TGAATCTGAT CGCCGACGTT CTGACCGCCG CCGCTGCAA TGCGGCGGCT
60
GCCGCCGCTT TGACTCTGAT CATGGCTGCC GGGCGGGTGC TACACAGAT GCAGCGGCCT
120
GTCTCCGGCC GCAGCATGCC GTTGGACGCG GACACCGGCC CTGACCTCGGC GCACCGCGCC
180
CACATTGACA CAGCTCTGGC GATTCACAGC TGTCGTTTGC GACAAAGGCC
240
AGTCTGCTGC AGGGGGGCGT GGGGGCCGCC GAGGGCGGCC TCCCGACCGA CAAGCTGAGG
300
AAAGCCGCAG GCAGCGGGGA TCTGGCGCTG TGATCCAGGC TAGCAAGACAT GCCAGCGCGG
360
GCCGCCGCTT CGCCACCGCC CGAGGTCCCT GTTCGCTGGGC CGAACGTCCT GTCGCCGGTC
420
AGCCAGAACG TCAGGCTCGT GAAATCAAGGC GGTCTGCCGC TGTCACGCGC ATGCGCGATG
480
GAGTTCGCTGC AGGCCGAGCC GAACCTGA
507

(2) INFORMATION FOR SEQ ID NO:100:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 168 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:100:

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

(2) INFORMATION FOR SEQ ID NO:101:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 500 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:101:

CGTGGCCATG TCGTTGACCG TCGGGGCCGG GTGCGCCCTCC GCAGATCCCG TGGACGCCTT 60
CATTAACACC ACCTGCAATT ACCGGCCAGTT AGTAGCTGGC CTCAAGCGGA CGGATCCGGG 120
GGTGGCGCGA CGATTCAACG CTCACCGCTT GGGCGAGTCC TATTGCGCA ATTTCTCGCC 180
GGCGACGCGA CCTCAAGCGCG CTGCCATGTC CGGCGAATTG CAGCTCTGTC CGGGGCGGCG 240
AACAGCATGC GCCCTGTGGC AGTCGGTGTC CGGCTCTGCC AACAATTA AAGCCCCATGC 300
GGGGCCCCRTCC CGCGACGCCG GCATGCCTGC GGGGTCTAGG CGAGATTGCC CGGCCTCTCA 360
ACGGGGCAGA TCCCGCGCAG CGCAGCTGTC GGGGGCTGTA GGGAGATTG CCCGCTCTCT 420
CAACCGGCGG CATCCTGTGC CCAATCTGCT CAGCCCCTGG GATCCACTAG TTCTAGACGG 480
GCCGCCACCG CGTTGAGCT 500

(2) INFORMATION FOR SEQ ID NO:102:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 96 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:102:

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

(2) INFORMATION FOR SEQ ID NO:103:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 154 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:103:

ATGACAGAGC AGCAGTGAAA TTTCCGGGCT ATCGAGCCCG CGGCAAAGCC AATCCAGGGA
AATGTCAGCT CCACTCATT CCTCCTTGAC GAGGGGAAGC AGTCCCTGAC CAAGCTCGCA
GGGCCCTGGG GCGTACCGGG TTTCTGAAAGG TACC

(2) INFORMATION FOR SEQ ID NO:104:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 51 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:104:

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

(2) INFORMATION FOR SEQ ID NO:105:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 282 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:105:

CGGTCGCGCA CTTCAGGGT ACTATGAAAG TCGGCTTCCG NCTGGAGGAT TCCTGAACCT
TCAAGGGGG CGGATACTG AGCTGCATCA TTAAAGCATC TTTCCAGAAC ATCTGGAGGC
GCTGAAACGG CGGCACGCC GACGGTGAGT CGGGCGAGGC GCTGCTTCCA AAATCCTGGA
GACAATTGCG CGGGCGCGCC TACAAGGAGG TCGTGCTGTA ATTCGNGGNG TATCTGGTGG
ACCTGTGTGCG TCTCGNACCG GACGAAAGCC TGCTCGACGT CG

(2) INFORMATION FOR SEQ ID NO:106:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3058 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
<table>
<thead>
<tr>
<th>Sequence Description: Seq ID No: 106</th>
</tr>
</thead>
<tbody>
<tr>
<td>GATCGTACCC GTGGGAGTGCC TGCGGCGGTT TGAGGATGGA GTGCACGCTGT CTTCGCTGTGAT 60</td>
</tr>
<tr>
<td>GGCATAACCA GAGATGGTGG CGCGCGGCAG TGACACCCCTG CAGAACATCG GTGCTTACCAC 120</td>
</tr>
<tr>
<td>TGTTGGCTAGC AATGCCCGCTG CGCCGCGCCC GACGACTGGG GTGCTGCCCC CGCGGTGCAG 180</td>
</tr>
<tr>
<td>TGAGGTGTGCG CGCGCTGACTG CGCGCGACCTT CGCCGCACAT GCGCGCATGT ATCAGTCCGT 240</td>
</tr>
<tr>
<td>GACGGCTGGG GCGTGGCTCGA TTCAATGACCA GCTCGTGCGG ACCCTTGGCCA GCAGCGCCAG 300</td>
</tr>
<tr>
<td>CTCGATGGCG GCCACTGAAG TGCCCAATTG CGCGCGCGGC AGCTAAAGCCA GGAACGGTGC 360</td>
</tr>
<tr>
<td>GCAGGAGAAA CCAAGAGAAA TAGGGACAGA TAATGGTGGA TTTCGCGGCCITT ACCACCAGG 420</td>
</tr>
<tr>
<td>AGATCAACTC CGCGAGGATG TACCGCGGCC CGCGCTGCTG CTCGCTGCTG GCCGCGGCCG 480</td>
</tr>
<tr>
<td>AGATGTGGGA CAGCGTGCGG CAGTGACCTGT TTTCCGGCGCC GTGCGGCTTT CAGTGGGTGG 540</td>
</tr>
<tr>
<td>TCTGGGTTCTG GAGCGTGGGG TGGTGATAG GTTCGCTGGC GCACGTGATG GTCCGGCCGGG 600</td>
</tr>
<tr>
<td>CCTGCGCTTA TGTTGCCTTG ATGAGCGCTCA CGCGCGGCGA GGCGACGCTG ACCCGCGCCC 660</td>
</tr>
<tr>
<td>AGGTCGCGGT TGGCTCGCGG GCTTACGAGA CGCGGTATGG GCTGACGGTG CCCCGCGCCG 720</td>
</tr>
<tr>
<td>TGATCGCGCA GAAACGCTCTA GAACTGATAG TTCTGATAGA GACCAACCTC TTGGCGACA 780</td>
</tr>
<tr>
<td>ACACCAGGCC GATCGCGGTC AAGAGGGCGG AATACGCGCGA GATGTGGGCC CAAGACGCGG 840</td>
</tr>
<tr>
<td>CGCGATGTTT TGCGTACGCC CGCGCGACCG CGACGCGGAC GGCGACATTG GTGCCTGTCG 900</td>
</tr>
<tr>
<td>AGGAGCGGCC GGAGATGACG AGCCGGGCTG GCTCTCCGCA GCAGCGCGCC GCAGTCCGAG 960</td>
</tr>
<tr>
<td>AGGCCTCCGA CACCAGCGCC CGCAACGACT TGATGAAACAA TGTCGCCCAG CGCGTGCAAC 1020</td>
</tr>
<tr>
<td>AGCTGGGCCCA GCCACAGCGA GCACAGCCGC TTCCTCCCAA GCTGGGTCGG CTGTTGAAAGA 1080</td>
</tr>
<tr>
<td>CGCGTTCCGC GCACTCGCTCG GCAGTACGCA ACATGGTGTC GATGGGCAAAC AACACATGT 1140</td>
</tr>
<tr>
<td>CGATGACCAA CTGGGCTGTC TGATGACCAC ACACCTTGAG TCTGATGTTG AAGGGCTTTG 1200</td>
</tr>
<tr>
<td>CTCGCGGCAG GCACGCGCAG CGCGCTGCAA CGCGCGGCCA AAACGCGTGC GCGCGCATGA 1260</td>
</tr>
<tr>
<td>GTCGCGCTGG CAGCGTCCTG GTGCTTCGAA GTCTGGGCGCG TGCGGTTGCG GCGCAACTTGG 1320</td>
</tr>
<tr>
<td>GTCGCGCAGC CTCGCGCTTT GCGCGGTCTG CGCGCGGCGA CGCGCGCGGGG CGCGCAACGG 1380</td>
</tr>
<tr>
<td>TCCGCAGCAG GCGCGCCCA TCCGGCAAAA CGCGCGCCAC AAACGCGTGC GCGCGCATGA 1440</td>
</tr>
<tr>
<td>GGGGCCGGCA GATGCTGGGC GGGCTGGCGG TGCCGCAAGA GACCCATGTATG AAGCGCGCAT 1500</td>
</tr>
<tr>
<td>GGGCGGCGG CAGCGTGGGG TTTCCGGCGGC GACCCATGTG AAGCGCGCAT TCCCGCGCGG 1560</td>
</tr>
<tr>
<td>CCGCGTGGGA GAGCGGCGCC AGACTGTCGCT TTATTTGACCA GTGATCGCGG GTCTCGGTGT 1620</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:107:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 391 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:107:

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

(2) INFORMATION FOR SEQ ID NO:108:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1725 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:108:

GACGTCAAGA CCGGCCGCTGC AGGGCTGGAG GTGTGCCTGC TTTGATCTGC GTTCAGGTTG 60
ACGTCCCCTG GCCTGTCCGC GCGGTGGATG CAGACTCGAT GGGCCTCTTT AGTGGCACTA 120
ATTTGTGTCG AGTGCCTGCG AGGTATAGGA CTTCACTGAT GGTATTATGT A GCTTACCC 180
CGTGGGGGGG TCAGTTGGGC CGGACCAGTC GTCAACCAAG CTTGCGGTGC GCGCCAGGCG 240
GGCGATCAGA TCCTGGACTC ACCAATCAA ACTTGAGCTCC CGGGCGGATG CTGGGGCTAA 300
ATGAGGGAGA GCACGGCTGT CTTTACCTGC GCAACCGGAG ATGTGGCGGG CCGCGGGCTGG 360
CGAAGTCTGT TCCCTGGGGG CAACGCTGAG GTGCTACCAAT GCCGCGCCAG CGTCGGCGAC 420
GACGTGGGTTG GTGCCGCGCG CTCGGAGACA ACCGTCGCTG GTGTGTCGCA CACAAATTCCG 480
TAGGCTAGAG CCGAGTATG AGACGCGGAG CCGCAAGGCC GCCTGATGCC ATGACCGAGT 540
TTGGGCAAC CGTCGCCACAT GCTGTAGTTC ATATCGGGAC ACCAGGCGAC CCGACGTCG 600
GGTACCCGGC TAGCTGACTT GACGCTATTC GAGCGGAAGG ATATACTGAG TGTGGGATT 660
<table>
<thead>
<tr>
<th>Val</th>
<th>Val</th>
<th>Asp</th>
<th>Phe</th>
<th>Gly</th>
<th>Ala</th>
<th>Leu</th>
<th>Pro</th>
<th>Pro</th>
<th>Gly</th>
<th>Ile</th>
<th>Asn</th>
<th>Ser</th>
<th>Ala</th>
<th>Arg</th>
<th>Met</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. INFORMATION FOR SEQ ID NO:109:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 359 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:109:

Val Val Asp Phe Gly Ala Leu Pro Pro Gly Ile Asn Ser Ala Arg Met
1 5 10 15

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Ala Lys Met Trp
20 25 30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ser Ala Phe Glu Ser
35 40 45
Val Val Trp Gly Leu Thr Val Gly Ser Trp Ile Gly Ser Ser Ala Gly

Leu Met Ala Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr

Ala Gly Gln Ala Gln Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala

Ala Tyr Glu Thr Ala Tyr Arg Leu Thr Val Pro Pro Pro Val Ile Ala

Glu Asn Arg Thr Glu Leu Met Thr Leu Thr Ala Thr Asn Leu Leu Gly

Gln Asn Thr Pro Ala Ile Glu Ala Asn Gln Ala Ala Tyr Ser Gln Met

Trp Gly Gln Asp Ala Glu Ala Met Tyr Gly Tyr Ala Ala Thr Ala Ala

Thr Ala Thr Glu Ala Leu Leu Pro Phe Glu Asp Ala Pro Leu Ile Thr

Asn Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Glu Ala Ile

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu

Gln Gln Leu Ala Gln Pro Ala Gln Gly Val Val Pro Ser Ser Lys Leu

Gly Gly Leu Trp Thr Ala Val Ser Pro His Leu Ser Pro Leu Ser Asn

Val Ser Ser Ile Ala Asn His Met Ser Met Met Gly Thr Gly Val

Ser Met Thr Asn Thr Leu His Ser Met Leu Lys Gly Leu Ala Pro Ala

Ala Ala Gln Ala Val Glu Thr Ala Ala Glu Asn Gly Val Trp Ala Met

Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Ser Gly Leu

Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser

Leu Ser Val Pro Pro Ala Thr Ala Ala Asn Gln Ala Val Thr Pro
Ala Ala Arg Ala Leu Pro Leu Thr Ser Leu Thr Ser Ala Ala Gln Thr
340 345 350
Ala Pro Gly His Met Leu Gly
355

(2) INFORMATION FOR SEQ ID NO:110:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3027 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:110:

AGTTCAGTCG AGAATGATAAC TGACGGGCTG TATGCCAGAT GGCCTGAGACA ACCGAACCCAC
60
CGTCCGGACGC GGGGACATCTG CAAAGCCACGCG GATGTGCTT TTGGCGCGAAAG GCCGAACCCG
120
CCGAAGCCGA AGCCCGTGGCC GCCGCCCGGGG GCCGCCGGGGG CGCGTGCGGGG CGGTTGACCG
180
GTGAGGCCGCT GTGGCCGAGGC CGCGCGGGC GGCGGCCAGCG GGGAGCGACCC AGGAGGCGGG
240
GGAAGACGCC GAAAACGTATGC GCAATGCTAGA CCGCAGGCAG GGGCCAGACCC AAGGGCGGGC
300
ACGGTGGGCA TCCGGCGGCC GCGGGTTCG CGCGCGGGTTTA TCAAGACTTG CACGAGATCC
360
CATGCGGCCG CGAGTCGCTGA TTATCCGGGC TTCGCCAGAGTACATTGGT
420
GCAACCCAGT GAGGCGCCCCG AACCCCAACCC CGCGCCCGCC CGGCTCCGCG GGCGAGCACA
480
GCAAGTGTGC ATCAACATGATC CCGCGTCCTG AATCCGACAAG GCCAAAGAAC AGCTGC CGCGG
540
TTGCTAGCAG CGTGCCAGCC GCCAATCTCG GCCGACGTCT AAGTGACGT TCCGACGCAGG GACGCCGATT
600
CACCAAGGGT GCGAAGCAGT GCCAAAGTGGT CACCGAAGACG AGCTGAACCG GACAGCCGGT
660
CGAATCAGTG CAGCAGTCGT CCTGGCCTCG GCACGTCCAG GGGTCACGAAA
720
TTCGCCTGGG GCCAAAGAGC ACCAGTGTCG GGGCGGGAGTC AAGTGACCG TGGCGAGACA
780
GGGGGACAG TACAAAGATGG CAGAAGTTGTT TTGGTGCCAC AGCGGATGTA GCTAGCCGAC
840
GTCAACACCG AAACACCTGCG CGCCACCGCA GTCGCTGGAAC TGACTGACAG CGACGGCCGA
900
GCCGGGATGTT CGCGGCCAGCA GCGATTGACG CACAGACCTGT CAAGCCCAAAC TACGCCGAG
960
AAGGTCAGGC GCCACCCGCG CCGCTGGGAT GCAGTGTTAAC CTTGGAACCG GCTTCCGGTG
1020
ATTCTCATTC TGCTCATCGT TGGCTGTTGG GGGGGCGGCC GATGGCTATA CCGGAAGC
1080
TACGACCCGG TCAGCGACGC GACTCGCCGG CGCCCGCTGGT TGCCCCGCGCG GCGCGCTCGT
1140
ACGGGACCATCG CGGTCCTGGT TTATCCGACC AGCGAGTGCTC ACGGAAGAATC CCGTCCGCGC
1200
AGGTCCGACC TCCCGGCCGA TTTCCCTGTC TATACGACCA GTTCCAGCCAG CAGATCGTGG
CTCCGCGGCC CAAACAGAAG TCACGAAAAC CCAGCCGCAA GGTGGTGGCG GCAGCCGCTT
CGAGACTACA TCCGAGTCCG GCGGTAGTCC TGGTTTTTGT CGACCAGAGC ACTACAGATA
AGGAGAGCCC CAATCCGTTG ATGGCCGAGCA GCAGCGTTGAT GGATGGACCTA GCCAAGCTGG
AGGGCAATTG GCTGATCACC AAGTTCACC CCGGTTAGGT TGCCGGAGTG GGTCGCAAAG
TCTGACGGGG GGCGCGTTGG CGTCTCGGTC GAGATACCCG CGTTCTCCGG GACATCCAG
GCCGACCTTC AACAGATCTT CGGCAGCTGT CTAATCGGCC GGGTTATTTA AGATTAGTTG
CCACCTGTAT CACCTGATTG TGCAAGTTTT GACCTGATTG TAGTCTCGCC GCAGGGCGGC
TGGTGACCTT TGGCATCTGG GGGTGACTA CTTGAGAGAA TTTGACCTGT TGCCGACGTT
GTTGGCTGTC CATCATTTGT GCTAGTTATG GCCGAGCGGA AGGATTATCG AAGTTGTTGA
CTGCGGGGGG TTACACCAGG AGATCAACTT CGCGAGGATG TACGGCCGCC CGGGTTCCGC
CTCGCTGTGTG CGCGCCGAGA AGATGTGGGA CAGCGTGCGG ATGGACCTGT TTTCCGGCCG
GTCGCCGTTTT CAGTGCGGTG TCTGGGTGCT CACGAGCGGA TCGGTGATAG GTCCTCGGCG
GGTCTCTATGG GTGGCGGGGG CGTCCCGCTCA TGGGCGTGGG ATGGAGCTCA CGCGGCGGC
GCCGAGCGTC ACCGGCGCGG AGGTCCGGGT TGCGGGCGCG GCCTACGAGA CGGCGTATGG
GCTGACGGTTG CCCCCGCGGG TGATCGCGGA GAACCTGGCT GAACGGTCTG TAATTGTGGC
GACCAACCTC TTGGGGCAAA ACACCCCGGC GATCGCGGTC AACGAGGGCG AATACGGGGA
GATGGGGCC CAAGACCGCG CGCGAGTTGT TGCTAGTGCC GCCACGCGCG CGACGGCGAC
CGAGGCCTTG CGGTGCTTGG AGGACGGCCC ACGTGACCC CACACCGCCG GCTCCTGGGA
GCAGGCCGGC GCGGTGCGAG AGGCCCCAGA CACCGCCCGG GCGAACAGTG TGATGAAACA
TGTCGCACCA GCCGTCGACA AACTGCGCCA GCGACGGAAAA AGCATCTGGG GTCGCGCCA
ACTGAGTGAA CTCTGGAAGG CACATCGCCG GCATCTGTCG CGCTCAGCA ACATCGGTGC
GATGCTCAAC AACACGGTG CGATGACCAA CTGGCTGGTG TGATCGCGCA GCACCTTGGA
CTCAGGTTTG AAGGGCTTTG CTCCGCGCGG GGGCTCGGGG GGTGAACCGG CCGCGAAAAA
CGGGGGCTAG CGATGAGGTG CGTGGGCGCC CGAGCTGGTG TGCGTGCTGG GTCTCTGGGG
TCTGGCGCGG CGGGTGCGGG CCAACTTTGG TGCGGCAGCC TGGGGCTGGT GCTGGTGCCTT
GCCGAGGCC TGGCGCGCGG CCAACCGAGG GGTCACTCGGC GCGCGGCGGG CCGCTCGGCT

128
GACCAGCCTG ACCACGGCG CCAAAACGGC CCCCAGACAC ATGCTGGGCG GGCTACCGCT

GGGGCAACTG ACCAATAGCG GCGGCCGGTT CGGCGGGGTT AGCAATGCGT TGCGGATGGC

GCCCGGGCGC TAGTGAATGC CCCGTGTCGCC CGCGGCGGG TAAGCCCGAT CCCGACGGCA

TGGGGGCGCT CTAATCGGGC AGCCGATC

(2) INFORMATION FOR SEQ ID NO:111:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 396 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(Xi) SEQUENCE DESCRIPTION: SEQ ID NO:111:

Val Val Asp Phe Gly Ala Leu Pro Pro Glu Ile Asn Ser Ala Arg Met
   1  5 10 15

Tyr Ala Gly Pro Gly Ser Ala Ser Leu Val Ala Ala Lys Met Trp
   20 25 30

Asp Ser Val Ala Ser Asp Leu Phe Ser Ala Ser Ala Phe Gln Ser
   35 40 45

Val Val Trp Gly Leu Thr Thr Gly Ser Trp Ile Gly Ser Ser Ala Gly
   50 55 60

Leu Met Val Ala Ala Ala Ser Pro Tyr Val Ala Trp Met Ser Val Thr
   65 70 75 80

Ala Gly Gln Ala Glu Leu Thr Ala Ala Gln Val Arg Val Ala Ala Ala
   85 90

Ala Tyr Glu Thr Ala Tyr Gly Leu Thr Val Pro Pro Pro Val Ile Ala
   100 105 110

Glu Asn Arg Ala Glu Leu Met Ile Leu Ile Ala Thr Asn Leu Leu Gly
   115 120 125

Gln Asn Thr Pro Ala Ile Ala Val Asn Glu Ala Glu Tyr Gly Glu Met
   130 135 140

Trp Ala Gln Asp Ala Ala Met Phe Gly Tyr Ala Ala Thr Ala Ala
   145 150 155 160

Thr Ala Thr Gln Ala Ala Leu Phe Pro Phe Gln Ala Pro Leu Ile Thr
   165 170 175

Asn Pro Gly Gly Leu Leu Glu Gln Ala Val Ala Val Glu Gln Ala Ile
   180 185 190

Asp Thr Ala Ala Ala Asn Gln Leu Met Asn Asn Val Pro Gln Ala Leu
Gln Gln Leu Ala Gln Pro Thr Lys Ser Ile Trp Pro Phe Asp Gln Leu
210  
215  
220
Ser Glu Leu Trp Lys Ala Ile Ser Pro His Leu Ser Pro Leu Ser Asn
225  
230  
235  
240
Ile Val Ser Met Leu Asn Asn His Val Ser Met Thr Asn Ser Gly Val
245  
250  
255
Ser Met Ala Ser Thr Leu His Ser Met Leu Lys Gly Phe Ala Pro Ala
260  
265  
270
Ala Ala Gln Ala Val Glu Thr Ala Ala Gln Asn Gly Val Gln Ala Met
275  
280  
285
Ser Ser Leu Gly Ser Gln Leu Gly Ser Ser Leu Gly Ser Gly Ser Gly Leu
290  
295  
300
Gly Ala Gly Val Ala Ala Asn Leu Gly Arg Ala Ala Ser Val Gly Ser
305  
310  
315  
320
Leu Ser Val Pro Gln Ala Trp Ala Ala Ala Asn Gln Ala Val Thr Pro
325  
330  
335
Ala Ala Arg Ala Leu Pro Leu Thr Ser Thr Ser Thr Thr Ser Ala Gln Thr
340  
345  
350
Ala Pro Gly His Met Leu Gly Gly Leu Pro Leu Gln Leu Thr Asn
355  
360  
365
Ser Gly Gly Phe Gly Gly Val Ser Asn Ala Leu Arg Met Pro Pro
370  
375  
380
Arg Ala Tyr Val Met Pro Arg Val Pro Ala Ala Gly
385  
390  
395

(2) INFORMATION FOR SEQ ID NO:112:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1616 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:112:

CATCGGAGGG AGTGATCACC ATGCTGTGSC ACCGCAATGCC ACCGGAGTAA ATACCGCGACG 60
GCTGATGSGCC GGC CGCGGGCT CGGCTC CAAT GCTTGCGCGG GCCGCGGGAT GGCAGACGCT 120
TTCGGCGGCT CTGAGCGGCTC AGGC CGCGGTCA GTGACGGCGG CGCGCTA ACTC TCTGGAGA 180
AGCTGGACT GAGGGTTGGGCA GCGACAGGGC GCTTGCGGCT GCACCGCGGA TGCTGGTCTG 240
(2) INFORMATION FOR SEQ ID NO:113:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 432 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:113:

CTAGTGGATG GGACCATGCG CATTTTCTGC AGTCTCAGTG CTTTCTGTTG TGACATTTTG

60

GCACGCCC GC GAAACGAAG CACTGGGTC GAAGAACGCC TGCCGCTGCA TATGCTCCGG

120

AGCTTTCCATA CTTTGGCGGC GCCGGAAGAG CTTGTCGTAG TCGGCGCGCA TGACAACCTC

180

TCAGAGTCCG CTCAAACGTA TAAACACGAG AAAGGGCGAG ACCGACGGAA GTTCGAAACTC

240

GCCGATCCC GTGTTCGCT ATTCTACCGC AACTCCGGCT TGCCCTATGC GAACATCCCA

300

GTGACGTTCG CTTCGGTGGA AGCCATGCC TGACCGGCTT CGTGTACGT CGCGCCAGG

360

TTCTGCAAGC CGTTGGTCAG CTGCGTGACC GTGGCGTCCC ATTTTTGCTG GACACCCCTGG

420

TACGCCCTCGG AA 432

(2) INFORMATION FOR SEQ ID NO:114:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 368 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:114:

Met Leu Trp His Ala Met Pro Pro Glu Xaa Asn Thr Ala Arg Leu Met
1  5  10  15

Ala Gly Ala Gly Pro Ala Pro Met Leu Ala Ala Ala Gly Trp Gln
20  25  30

Thr Leu Ser Ala Ala Leu Asp Ala Gln Ala Val Glu Leu Thr Ala Arg
35  40  45

Leu Asn Ser Leu Gly Glu Ala Trp Thr Gly Gly Gly Ser Asp Lys Ala
50  55  60

Leu Ala Ala Ala Thr Pro Met Val Val Trp Leu Gln Thr Ala Ser Thr
65  70  75  80

Gln Ala Lys Thr Arg Ala Met Gln Ala Thr Ala Gln Ala Ala Tyr
85  90

Thr Gln Ala Met Ala Thr Thr Pro Ser Leu Pro Glu Ile Ala Ala Asn
100 105 110

His Ile Thr Gln Ala Val Leu Thr Ala Thr Asn Phe Phe Gly Ile Asn
115 120 125

Thr Ile Pro Ile Ala Leu Thr Glu Met Asp Tyr Phe Ile Arg Met Trp
130 135 140
Asn Gln Ala Ala Leu Ala Met Glu Val Tyr Gln Ala Glu Thr Ala Val
145    150    155    160
Asn Thr Leu Phe Glu Lys Leu Glu Pro Met Ala Ser Ile Leu Asp Pro
165    170
Gly Ala Ser Gln Ser Thr Thr Asn Pro Ile Phe Gly Met Pro Ser Pro
180    185    190
Gly Ser Thr Pro Val Gly Gln Leu Pro Pro Ala Ala Thr Gln Thr
195    200    205
Leu Gly Gln Leu Gly Glu Met Ser Gly Pro Met Gln Gln Leu Thr Gln
210    215    220
Pro Leu Gln Gln Val Thr Ser Leu Phe Ser Gln Val Gly Gly Thr Gly
225    230    235    240
Gly Gly Asn Pro Ala Asp Glu Glu Ala Ala Gln Met Gly Leu Leu Gly
245    250    255
Thr Ser Pro Leu Ser Asn His Pro Leu Ala Gly Gly Ser Gly Pro Ser
260    265    270
Ala Gly Ala Gly Leu Leu Arg Ala Glu Ser Leu Pro Gly Ala Gly Gly
275    280    285
Ser Leu Thr Arg Thr Pro Leu Met Ser Gln Leu Ile Glu Lys Pro Val
290    295    300
Ala Pro Ser Val Met Pro Ala Ala Ala Gly Ser Ser Ala Thr Gly
305    310    315    320
Gly Ala Ala Pro Val Gly Ala Gly Ala Met Gly Gln Gln Gly Ala Gln Ser
325    330    335
Gly Gly Ser Thr Arg Pro Gly Leu Val Ala Pro Ala Pro Leu Ala Gln
340    345    350
Glu Arg Glu Glu Asp Asp Glu Asp Trp Asp Glu Glu Asp Asp Trp
355    360    365

(2) INFORMATION FOR SEQ ID NO:115:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 100 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:115:
Met Ala Glu Met Lys Thr Asp Ala Ala Thr Leu Ala Gln Glu Ala Gly
1    5    10    15
Asn Phe Glu Arg Ile Ser Gly Asp Leu Lys Thr Gln Ile Asp Gln Val
20 25 30
Glu Ser Thr Ala Gly Ser Leu Gln Gly Gln Trp Arg Gly Ala Ala Gly
35 40 45
Thr Ala Ala Gln Ala Ala Val Val Arg Phe Gln Glu Ala Ala Asn Lys
50 55 60
Gln Lys Glu Leu Asp Glu Ile Ser Thr Asn Ile Arg Gln Ala Gly
65 70 75 80
Val Gln Tyr Ser Arg Ala Asp Glu Glu Gln Gln Gln Ala Leu Ser Ser
85 90 95
Gln Met Gly Phe
100

(2) INFORMATION FOR SEQ ID NO:116:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 396 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:116:

GATCTCCGCG GACCTGAAAA CCCAGATCGA CCAGGTGGAG TCGACGCGAG GTTGGTTGCA
GGGCCAGTG CCGGCGCCGG CGGGGACGGC CGCCCCAGCC GCGGTTGGCG GTTTCCAAGA
AGCAGCCCAAT AGGAACTCGA CGAGATCTCG ACGGATATTC GTCAAGCCGG
CGTCCACAT TGAGGCGCG AGAGAGAAGCA CAGAGCTCGG CTGTCCTCGC AAATGGGCTT
CTGACCCGCT ATTAGGAAAA GAAACGCGGG AAAACATGA CAGACCGCAGA GTGGAAATTC
GGGGTATCG AGGCGCGCGG AAGCGCAATC CAGGGAATG TCAGTCCAT TCATCCCTC
CTTGCAGAGG GGAAGCAGTC CCTGACCAAG GTGCGA

(2) INFORMATION FOR SEQ ID NO:117:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: Single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:117:

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

(2) INFORMATION FOR SEQ ID NO:118:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 387 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:118:

GTGGATCCCG ATCCCGTTT TCAGTATTCT AGCGGAACCT GGGTTGGCC TATGCGAACA  60
TCCCAAGTGAC GTTGCTTTGC GTGAAGGCCA TTGCCTGACC GGCTTCGCCT ATCGTCCCGG
CCAGGGTTCG CAGCGGGTTG TTCAAGCTCGG TAGCCGGTTC GTCCCCATTT GCTGGACAC
CCTGGTAGCC CTCCGAAACG CTACCAGCCC AGGCCCGCTGC GAGCTTGGTC AGGGACGTCT
TCCCTCTGGC AAGGAGGGAA TGAATGGACG TGACATTTC CTGGAATGGCG CTGGGCAGGG
CCTGATACC CGCGAAATTC CACTGCTGCT CTGTCATGTT TTGCTCAGGT TTCTTTCTGT
ATTAGCGGCT CAGAAGCCCA TTTGCGA  387

(2) INFORMATION FOR SEQ ID NO:119:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 272 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:119:

CGGCACGGAG ATCTCGGGTG GCCCAACGGG GCTGGGCGAGG GCTCCCTTCC GGGGCGAGGC  60
TGCGGCGCGG ATGGTTCCTC TGCCAAGCCG CGCGCCTGGA TGGATGGGAC AGTTGCTACC
TTCCCGACCT TTCGTTCCGGT GCTCTTCGCCA TAGGCGTCAG CCGGCGCAGG ACCTGGGAG
TGTTGGGGGG CAGGGGGGCT CGGTGTTCCG GCCGGGGACG CAGACGGGCT TAGACGGGACG  240
GGCGGGGGTT CGCCGATTTG GATCTTTGCCC CA

(2) INFORMATION FOR SEQ ID NO:120:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 20 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:120:

Asp Pro Val Asp Ala Val Ile Asn Thr Thr Cys Asn Tyr Gly Gln Val
1    5     10
Val Ala Ala Leu
    15

(2) INFORMATION FOR SEQ ID NO:121:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:121:

Ala Val Glu Ser Gly Met Leu Ala Leu Gly Thr Pro Ala Pro Ser
1    5     10
    15

(2) INFORMATION FOR SEQ ID NO:122:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 19 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:122:

Ala Ala Met Lys Pro Arg Thr Gly Asp Gly Pro Leu Glu Ala Ala Lys
1    5     10
    15

Glu Gly Arg

(2) INFORMATION FOR SEQ ID NO:123:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 15 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:123:

Tyr Tyr Trp Cys Pro Gly Gln Pro Phe Asp Pro Ala Trp Gly Pro
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:124:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:124:

Asp Ile Gly Ser Glu Ser Thr Glu Asp Gln Gln Xaa Ala Val
1  5  10

(2) INFORMATION FOR SEQ ID NO:125:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 13 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:125:

Ala Glu Glu Ser Ile Ser Thr Xaa Glu Xaa Ile Val Pro
1  5  10

(2) INFORMATION FOR SEQ ID NO:126:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:126:

Asp Pro Glu Pro Ala Pro Pro Val Pro Thr Thr Ala Ala Ser Pro Pro
1  5  10  15

Ser

(2) INFORMATION FOR SEQ ID NO:127:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:127:

Ala Pro Lys Thr Tyr Xaa Glu Glu Leu Lys Gly Thr Asp Thr Gly
1   5   10   15

(2) INFORMATION FOR SEQ ID NO:128:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:128:

Asp Pro Ala Ser Ala Pro Asp Val Pro Thr Ala Ala Gln Leu Thr Ser
1   5   10   15
Leu Leu Asn Ser Leu Ala Asp Pro Asn Val Ser Phe Ala Asn
20   25   30

(2) INFORMATION FOR SEQ ID NO:129:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:129:

Asp Pro Pro Asp Pro His Gln Xaa Asp Met Thr Lys Gly Tyr Tyr Pro
1   5   10   15
Gly Gly Arg Arg Xaa Phe
20

(2) INFORMATION FOR SEQ ID NO:130:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:130:

Asp Pro Gly Tyr Thr Pro Gly
1   5

(2) INFORMATION FOR SEQ ID NO:131:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: /note= "The Second Residue Can Be Either a Pro or Thr"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:131:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Xaa</td>
<td>Xaa</td>
<td>Phe</td>
<td>Thr</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:132:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ix) FEATURE:
(D) OTHER INFORMATION: /note= "The Third Residue Can Be Either a Gln or Leu"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:132:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Xaa</td>
<td>Pro</td>
<td>Xaa</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:133:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:133:

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Xaa</td>
<td>Xaa</td>
<td>Glu</td>
<td>Lys</td>
<td>Pro</td>
<td>Phe</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:134:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:134:

<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Xaa</td>
<td>Asp</td>
<td>Ser</td>
<td>Glu</td>
<td>Lys</td>
<td>Ser</td>
<td>Ala</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:135:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:135:

Ala Gly Asp Thr Xaa Ile Tyr Ile Val Gly Asn Leu Thr Ala Asp
 1     5          10        15

(2) INFORMATION FOR SEQ ID NO:136:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:136:

Ala Pro Glu Ser Gly Ala Gly Leu Gly Gly Thr Val Gln Ala Gly
 1     5          10        15

(2) INFORMATION FOR SEQ ID NO:137:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 21 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:137:

Xaa Tyr Ile Ala Tyr Xaa Thr Thr Ala Gly Ile Val Pro Gly Lys Ile
 1 5          10        15
Asn Val His Leu Val
 20

(2) INFORMATION FOR SEQ ID NO:138:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 882 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:138:
GCAACGCTGT CGTGGGCTTT CGGTGATCG GTTTCGCCCCT GCTGGCGGGTG GCGGTGGCGG
60
TGACCATCCG ACCGACCGCG GCCCTCAAAAC CGTGTAGAGGG ACACCAAACAC GCCCCGCCAG
120
GGAAGTTCAT GCCGTTGTGG CGAACGCAAC AGCAGGGC CGTCCCGGCG CCTCGGGCCTG
180
ATGATCCAC CGTGGATTCC CAGGGCCGCA CCATTCCGGC TGTACAGAAC GTGTTGGCCGC
240
GGCGGGTTAC CTCACCCGGG GTGGGTGGGA CGCCGGGTTTC GCCGGCGCCC GAAAGGGCCCG
300
CGGTGCCGGG TGTTGTGGCT GCCCGGTGGC CAATCCCCGT CCGATCATCA ATCCCCCGGT
360
TCCCGGTTTG GCAGCCTGGA ATGCGCGACCA TCCCAACGGG ACCGGCGACG AGCGCGCGTA
420
CCACGTGCCG GACGACCGCG CGGACCCCGC CGGGCGCGAC CGGGTGGACC AGCGCGCCAA
480
CGACCGCCGC GACCAGCGCG GTGACCGCACG CGGCAACGAC GCGCGGACCG AGCGCGCGTA
540
CCACGCCCACC AACGACCGTCG GCCCGGAAGC CGTCCGCCCC GAGACGCTTG GCTCCGGCA
600
CCGTGCCGCC GACCAGGGTC GCTCGAGCGA CGGCCAGCGC GAGGACGCTG GCTCGGGCAC
660
CGACCGAGCA GCCACCGAAG CAACCAACC AACAGATGCC ACCCCCGCAG CAGACCGTTG
720
CCCCCGCAGC GGTGGCCCGC GTCGGCGAGC CGCGTGCCCG TGCGCCGACA GCACCGCGGC
780
GGGCCGACTT ATCCGGCCGG TTCTGATCAAC GTGCGCGGGCT CGCTAAGCGGT CGGAGGACAT
840
GGCGGGTGAT CGGGTGACCG TGTTGCTGGC CTGTCTCAAC GA
882

(2) INFORMATION FOR SEQ ID NO:139:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 815 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:139:

CCATCAACCA ACCGCTCGCG CGCCCGCCGC CGCGGATCC CGCGTCGGCG CGACGCCCGC 60
CGGTGCCTCC GGTGCCCGCG TTGGCCCGCT CGCCGCGGCT CGCCGGACG GGCTGGGTGC
120
CTAGGGCGCT GTTACCGGCG TTGGTGCCCG GGACGCCCGC GCCACCGCG GTACCGCGGA
180
TGCGCCGCTT GCCGGCGGGC GCACGTTGCC CACGGTTGCC ACCGTTCGCC CGTTGCGCA
240
CGAGCCACCC CGCGGGGACA CCAGCCACC CGCGGGCGCC CGCACCGCGC GGGTGCGCCTG
300
TGGTGCGCGT ACCGCGGCGC CGCGCTTGCC CGCGTCACC GCCGACGGAA CTACCGGCGG
360
ACGCGGCGTG CGCGGGCGGC CGCCCGCCAC GCCCATGGSC ACCGGCGTCA CGCGCGGGTG
420
GGAGTGCGCC GATTAGGGCA CTGACCGGCG CAACCAGCGC AAGTACTTCTC GGTACCAGG

CACCAGAGA CAGACACACA GCAAGGTTT GTCCGGCGGA TGGGTGAAT GCCAGCGGAT

AGCGGCTAGC TGTCGCTGTC GGTCAACCTC GATCATGATG TCGAGTTGAC CCGTAGCGCG

ccccccgag aaggcgctga actccgcttg gaggccagtc cggctcggtt gggcagggtc

ccagcccaat acggggatac cgggtgtcna agccgcccgag agccgacgctt cggtgcgcgg

acngttgtcg gggtggcctg ttacgctgggt gtntctgaac acaagtagca ggtctgctcc

ccaggggca tccaccagcc gttgccgctag ctcgt

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 1152 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 140:

acgagccgcc gctgtagggtc tcagatcaga gatctccgg actcaccggg ggggctcagc

cttctccag aacaaactgct gaagatcttc gccccgcaga cagccgctga tttgaccttc

TATGACCCTG TGAACGACGA GATCATCGCC GAGATTGATA TGGCACCCTG GGGCTAACAG

GTGGCCAAGA TGTTGCAGC GTATGCTCC GTACCCCGTG CGGGATCGT CTTTGGCGAC

GGCCGGGTGA TGCGTGGGAG CGAGAAGCCGACTACCGGG CGAGAAGAGC

GAGCGCATCA CGCTGTGGG GCGGCCAGAG ATGACACACG CTTTCATCCTG TGAGATGCTC

AACCTGACGC GCGACATCCA GCTCTCCACG ACCAGCAGGC ACTACAGGGG CGGATCTCA

ACACCCGACG TGTCATAGCC GCCGCGCTCG CGTCGACAGG TTACCGCAG GCAGACAGCT

GCGTCTGGCC TGTCGTGAA CAGCGAGATC GTTCGAGGGA AGATCCTGAA TACGCCAGGC

TTGATTCCGG CACACACGTC GGGCAAGAC GTGCTGAGA GCATCCGGAC GATGAAGCAC

TCGGTCGGCT GGCGCTGAGC ATCCGGACT CGTCGAGGTT TACGCCAGGAT

GGCGAAGG CATACCTCAG CGCGCTGGGG CATAGTCTGC CGCGAGTGG GCATTCAGAG

GGCCTCGCGA CTGGCCCGGC GTTGAGCGCC TTCAACGTA CTTGCAGCCCT CGGCTATTGC

CTGCTGTACA AGAACATCAT AGGGCGATC GAGCGTCACA GCCTGAACGC GTATATCGGT
TTCCCTACACC AGGATTCAG AGGGCAGCAG ACGTCTCGTG CCGAATTCGG CACGAGCTCC
GCTGAAACGG CTGGCCGCGT GCTCAGTGGC CGTACGTTAAT CGGCTGCGCC CAGGCCGGCC
CGCCGGCGGA ATACACAGAC ATCGGCGAAC GAATGCGCGC CCAGCCCGTTT GGAGCCGCTGC
ATACCGCCGG CACACTCACC GGCAGCGAAC AGGCTTGGCA CGGTGCGGCG GCGGCGTCC
GGTCTTACCTT CGACACCGCC CATCAACCTAG TGACAGCTCG GCCCCGACTTC CATGCGCTGC
GTCCGCGACG AG

(2) INFORMATION FOR SEQ ID NO:141:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 655 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:141:
CTGGTGGCGA TTGGCGAGGG TGACTATTGC GGTGGTGTAAN GCCGCGATGAG TGCCGAGGC
60
CAGCAATGGC GCAAACGACC GGAATCCCGGT CAACGGCAGCC ACGCGGTTCCA CGTGCGCGAT
120
CGGCTGCAAT CCGCTCTTGG CCGGCTTCCC TGGGGCGAGG GTCATCGGAC GTGTCTTCGC
180
CGTGTGTTGC CGCATTATAG CCGGCCGGCG GCGCGCGCGG GCCGGTATGG CGAAANGTCG
240
ATCAGCACCG CCGAGATAAG GGTCTTGGCA AGCTTTTGTG GGGTGGCGCG GGGCGCGTTC
300
GGCGGCAATT CTACTAGCGA GAACTTCGGC CGCATACGGA TCTGACCAGA GTGCCTCGGG
360
TGCGAGCCAC CCTCATTGGC GATGCGCGCC AGCGATGGCGG CTUGCACGAT GCTGCGCC
420
TTCCGGGACG CGACGCGTGA GGTGGAACATG GCCGGTCTAC GCTGGCGGTC CTGGCCGACG
480
TCCCCGACTT GTGCACGGTTG CCACCCGCACG ACGCGCGGGT CGGTGCCCAT CAGGAATGCC
540
TCACCGCGCC GCACCTCGAC GCACGTGGC GCACCGATGT CAGCCATCGG GACATCATGC
600
TGCGGTTCCA ACTCTCTGAC AGCTCGGCGG AACGCTCGA TTCCCCGACC GCCCA
655

(2) INFORMATION FOR SEQ ID NO:142:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 267 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:142:

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

(2) INFORMATION FOR SEQ ID NO:143:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 174 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:143:

Ile Asn Gln Pro Leu Ala Pro Pro Ala Pro Pro Asp Pro Pro Ser Pro  
1      5    10  15

Pro Arg Pro Pro Val Pro Pro Val Pro Pro Leu Pro Pro Ser Pro Pro  
20     25    30

Ser Pro Pro Thr Gly Trp Val Pro Arg Ala Leu Leu Pro Pro Trp Leu  
35     40    45

Ala Gly Thr Pro Pro Ala Pro Pro Val Pro Pro Met Ala Pro Leu Pro  
50     55    60

Pro Ala Ala Pro Leu Pro Pro Leu Pro Pro Leu Pro Pro Leu Pro Thr  
65     70    75  80

Ser His Pro Pro Arg Pro Pro Ala Pro Pro Ala Pro Pro Ala Pro Pro  
85     90    95

Ala Cys Pro Phe Val Pro Val Pro Ala Pro Pro Leu Pro Pro Ser  
100    105    110

Pro Pro Thr Glu Leu Pro Ala Asp Ala Ala Cys Pro Pro Ala Pro Pro  
115    120    125

Ala Pro Pro Leu Ala Pro Pro Ser Pro Pro Ala Gly Ser Ala Ala Ile  
130    135    140

Arg Ala Leu Thr Gly Ala Thr Ser Ala Ser Thr Leu Gly His Arg Ala  
145    150    155  160

Leu Pro Asp Asp Thr Thr Ala Arg Gly Cys Arg Arg Thr Gly  
165    170

(ii) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 35 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:144:
Gln Pro Pro Ala Glu Val Ser Asp Gln Arg Val Ser Gly Leu Thr Gly
1  5  10  15

Ala Val Gln Pro Ser Pro Arg Thr Thr Ala Glu Asp Pro Arg Pro Arg
20  25  30

Asn Arg Arg
35

(2) INFORMATION FOR SEQ ID NO:145:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 104 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:145:

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

(2) INFORMATION FOR SEQ ID NO:146:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 53 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
   (A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:146:

GGATCCATAT GGGCCATCAT CATCATC ATC ACGTGATCGA CATCATCGGG ACC

(2) INFORMATION FOR SEQ ID NO:147:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 42 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:147:

CCTGAATTCA GGCCTCGTT GCGCGGCT CATCTTGAC GA

(2) INFORMATION FOR SEQ ID NO:148:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR Primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:148:

GGATCCTGCA GGCTCGAAAC CACCAGCGG T

(2) INFORMATION FOR SEQ ID NO:149:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 31 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:149:

CTCTGAATTCCAGGCTGGAAATCGTCGCCAGT

(2) INFORMATION FOR SEQ ID NO:150:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:150:

GGATCCAGGCGCTGAGATGAAAGCCGATGCCGCT

(2) INFORMATION FOR SEQ ID NO:151:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PCR primer"

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:151:

GAGAAGATTCATCCAGGGGATTTGCGAGGACA

(2) INFORMATION FOR SEQ ID NO:152:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1993 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Mycobacterium tuberculosis

(ix) FEATURE:
TGTTCTTCGA CCGCAGGCTG GTGGAAGGAA  GGGCCACCAGA  ACAGCTGTTC TTCTGCCCGA  
AGCATGCCGA AACGCCCGAA  TACGTCGCGG  GACTGTCGGA  GGAGCTCAGG  GAGGCCAACG  
GCCGAAAATTG AAGAGCACAG AAAGGTATGG C GTG AAA ATT GGT TTG CAT ACG  
Val Lys Ile Arg Leu His Thr 1 5  
CTG TTG GCC GTG TTG ACC GCT GCC CCG CTG CTA GCA GCG GCG GCC  
Leu Leu Ala Val Leu Thr Ala Ala Pro Leu Leu Ala Ala Gly 10 15 20  
TGT GCG TCG AAA CCA CCG AGC GGT TCG CCT GAA ACG GCC GGC GCC GCC  
Cys Gly Ser Lys Pro Pro Ser Gly Ser Pro Glu Thr Gly Ala Gly Ala 25 30 35  
GGT ACT GTC GCC ACT ACC CCC GCC TCG TCG CGG GTG ACG TTG GCG GAG  
Gly Thr Val Ala Thr Pro Ala Ser Ser Pro Val Thr Leu Ala Glu 40 45 50 55  
ACC GGT AGC ACG CTG CTC TAC CCG CTG TCG TTC AAC CTG TGG GGT CCG GCC  
Thr Gly Ser Thr Leu Tyr Pro Leu Phe Asn Leu Trp Gly Pro Ala 60 65 70  
TTT CAC GAG AGG TAT CCG AAC GTC ACG ATC ACC GCT CAG GGC ACC GGT  
Phe His Glu Arg Tyr Pro Val Ser Val Thr Ala Gln Gly Thr Gly 75 80 85  
TCT GGT GCC GGG ATC GCG CAG GCC GCC GCC GGG ACG GTC AAC ATT GGG  
Ser Gly Ala Gly Ile Ala Glu Ala Ala Gly Thr Val Asn Ile Gly 90 95 100  
GCC TCC GAC GCC TAT CGT TCG GAA GGT GAT ATG GCC GCG CAC AAG GGG  
Ala Ser Asp Ala Tyr Leu Ser Glu Asp Met Ala Ala His Lys Gly 105 110 115  
CTG ATG AAC ATC GCG CTA GCC ATC TCC GCT CAG CAG GTC TAC AAC  
Leu Met Asn Ile Ala Ala Ile Ser Ala Gln Gln Val Asn Tyr Asn 120 125 130 135  
CTG CCC GGA GTG AGC GAG CAC TCT AAG CTG AAC GGA AAA GTC CTG GCG  
Leu Pro Gly Val Ser Glu His Leu Lys Leu Asn Gly Lys Val Leu Ala 140 145 150  
GCC ATG TAC CAG GCC ACC ATC AAA ACC TGG GAC GCC CAG ATC GCT  
Ala Met Tyr Glu Thr Ile Lys Thr Trp Asp Asp Pro Glu Ile Ala 155 160 165  
GCC CTC AAG CCC GCC GTG AAC CTG CCC GCC GCC GCC GCC ACG GTA GTT CCG CTG  
Ala Leu Asn Pro Gly Val Asn Leu Pro Gly Thr Ala Val Val Pro Leu 700
CAC CGC TCC GAC GGG TCC GGT GAC ACC TTC TTG TTC ACC CAG TAC CTG
His Arg Ser Asp Gly Ser Gly Asp Thr Phe Leu Phe Thr Gln Tyr Leu
185 190 195

TCC AAG CAA GAT CCC GAG GGC TGG GCC AAG TCG CCC GCC TTC GGC ACC
Ser Lys Gln Asp Pro Glu Gly Trp Gly Lys Ser Pro Gly Phe Gly Thr
200 205 210 215

ACC GTC GAC TTC CCG GCG GTG CCG GGT GCG CTG GGT GAG AAC GCC AAG
Thr Val Asp Phe Pro Ala Val Pro Gly Ala Leu Gly Glu Asn Gly Asn
220 225 230

GCC GCC ATG GTG ACC GGT TGC GCC GAG ACA CCG GCC TGC GTG GCC TAT
Gly Gly Met Val Thr Gly Cys Ala Glu Thr Pro Gly Cys Val Ala Tyr
235 240 245

ATC GGC ATC AGC TTC CTC GAC CAG GCC AGT CAA CCG GGA CTC GCC GAG
Ile Gly Ile Ser Phe Leu Asp Glu Ala Ser Gln Arg Gly Leu Gly Glu
250 255 260

GCC CAA CTA GCC AAT AGC TCT GCC AAT TTC TTG CCC GCC GCG CAA
Ala Glu Leu Gly Ser Ser Ser Gly Asn Phe Leu Leu Pro Asp Ala Gln
265 270 275

AGC ATT CAG GCC GCG GGT GCC TTC GCA TCG AAA ACC CCG GCC AAG
Ser Ile Glu Ala Ala Ala Gly Phe Ala Ser Lys Thr Pro Ala Asn
280 285 290 295

CAG GCG ATT TCG ATG ATC GAC GGG CCC GCC CCG GAC GCC TAC CCC ATC
Gln Ala Ile Ser Met Ile Asp Gly Pro Ala Pro Asp Gly Tyr Pro Ile
300 305 310

ATC AAC TAC GAG TAC GCC ATC TTC ACG AAC CCG CAA AAG GAC GCC GCC
Ile Asn Tyr Glu Tyr Ala Ile Val Asn Asn Arg Glu Lys Asp Ala Ala
315 320 325

ACC GCG CAG ACC TTG CAG GCA TTT CTG CAC TGG GCG ATC ACC AAC GCC
Thr Ala Glu Thr Leu Glu Ala Phe Leu His Trp Ala Ile Thr Asp Gly
330 335 340

AAC AAG GCC TCG TTC CTC GAC CAG GTT CAT TTC CAG CCG CTG CCC
Asn Lys Ala Ser Phe Leu Asp Glu Val His Phe Glu Pro Leu Pro Pro
345 350 355

GCG GTG GTG AAG TTG TCT GAC GCG TTG ATC GCG ACG ATT TCC AGC
Ala Val Val Lys Ser Asp Ala Leu Ile Ala Thr Ile Ser Ser
360 365 370

TAGCCTCGTT GACCACCAAG CGACAGCAAC CTCCGTCGAG CCATCGGGT GCTTTGCGGA
1333

GCATGCTGCG CCCTGCGCGT GAAGTGGCG GCCGCTGCGGC GGCCCAGCGG TGTTGGGGTG
1393

GGAATGATCGC GGGATCCCG GCTGTTGGTG GCTGCTCATCG
1453
AGGCAGATGGG TGGCATCAGG CTCAACGGGT TGCTATTCTT CACGGCCACC GAATGGAATC
CAGCAACAC CTACGGCGAA ACCGTTGCTA CGGACCGGTC GCCCCATCGG TGGGCGGCAA
CTAGGGGGCG TTGCCGCTGA TGTCGGGGAC GCTGCGGACC TGGGCAATCG CCGTGATCAT
CGCGTGCCCG GTCTCTGCTG GAGGCGCGCT GTTGATCTGG GAACGCGTGC CGAAGCGGTT
GGCCGAGGCT GTGGGAATAG TCCTGGGAAT TCTCGGCCGA ATCCCCAGCG TGGTGCTCGG
TTTGTGGGGG GCAATGACGG TCGGCGCGTT CATCGCTCAT CACATCGCTC CGGGTGATGC
TCACAAAAGCT CCCGATGTGG CGGTGCTGAA CTACTTGGCG GCGCCACGGG GCAACGGGSA
GGGCATGTG TGTTCCGCGTC TGTTCTGCGC GGGATGTGGTC GTTCCCATTA TGGCCACAC
CACTCATGAC CTGTCCCCGG AGGTGCCGGT GTTGCCCCCG GGGGGCGGGA TCGGGAAATT

(2) INFORMATION FOR SEQ ID NO:153:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 374 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:153:

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

(2) INFORMATION FOR SEQ ID NO:154:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 1993 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:154:

TGTTCTTCGA CGCGCAGCTG TTGGAGGAAG GGCCCCACCGA ACAGCTGTTTCC TCTCTGCAGCA 60
AGCATGCGGA AACGGGCGGA TACGCAGCGG GACACGTGGGG GAGCCTGCAAG GACGCCAACG 120
GCGGAAATTG AAAAGCTAC AGAACGATTG CGTGAATAAT CGTTGCCATA CGGTGTGGGC 180
CGTGGTGGGC CGTGGGCGGCG TCCTGCTAGC AGCGGCGGCG TGGCTGTCGA ACCACCGAG 240
CGGGTTGCTT GAAGCGGGCG CGCCCGCCGG TACTGTCCGG ACTACCCCGG CCTCGTGCCGC 300
GGTGAGCTTG CGCCAGACCG CGCACGGCTT GGTACACCCG CTGTATCCACG TGGGCTGTCG 360
GGCCCTTTCAC CAGAGGTATC CGGAAGTACG GATCTCCCGT CAGGACGGAC GGGCTGGTGC 420
CGGCTGCTGG CAGGCCGCGG CGCGGAGGCT CAACTTTGGG GCTCCGACCG CCTATGTGGG 480
GGAAGCCTGT GATGCAGCGC ACAAGGCAGCT GATGAACATC GCGCTAGCGA TCTCGCTCA 540
GCAGGTCAAC TACACCTGCG CCGAGTGAGC CGAGCACCTC AAGCTGAAAC GAAAGATCAT 600
GCGGCCTGTT TACCAGGCAA CCATCAAAAA ATGGGAGACG CGCGAGTACG CTTCCGCTCA 660
CCCGGCGTGG AACCTGGGCG GCACCGCGGT AGTTCGGCTG CAGGCGCTTCG ACGGCTGGG 720
TGACACCTTC TTGTTCCACC AGTACTCGTC CAAAGCAGAT CCCGAGGCTG GGGGAAGTGC 780
GCCGCCTTTC GCCAGCACC CGCTGGCTCC CGGGCTCGCG GGTGCGCTGG TGAGAAACGG 840
CAACGGCGGC ATGGTGACCG GTTGCGCCCG GACACCGGCC TGGCTGACCTG ATATCGGCAAT 900
CACCTTCTTC GACCATGGCA GTCAACGGGG ACTCCGGCGG GCCCAACTAG GCCATAGCTC 960
TGGCAATTTC TTGTGGGGCG ACCGCGCAAAG CTATCAAGCC GCGGCCGCTG GCTTGGCATC 1020
GAAAACCCCG GCCGAACCGG CGATTTCGAT GATCGACGGG CCCGCCGCCGG AGGGCTACC 1080
GATCATCAAC TACGGTAGCA CCACTCGGCA CACCCGGCAA AAGGACGCCC CCACCCGCCA 1140
GACCTTGGAG GCCATTCTGC ACTGGGGGAT CACCGAGGCC AACAAGGCCT GTTCTCCGTA 1200
CCAGGTTCAT TCCGACGGCG TGGCCGCGCG GGTGCTGAAG TTGTGGCTAG CCGGTGATCG 1260
GACGATTTCG AGCTAGCTTC GTGGACACCC AGCGACAGCC AACCTCCTCG GCCGCGCCAC 1320
GCTGCTTTGC GGAGCATGCG GGGGCTGGCC GGTGAAGGCG GCCGCCGCTGG CCCGGCCCAT 1380
CGTTGTCGCG GTGGTGGTCG TGGGCTGATC CGCGCTGCTT GGTGCTGGGT CTTGTCCTTT 1440
GTTGTCGTCG TGGGAGGGAT GGTGCGATCG AGGCTCAACC GGGTCTATTG CTTCCGCGCC 1500
ACCGAATGGCA ATCCGGCAGA CCACCTCGGC GAAAGGCTTG ATCGCCACGC TCGGCCCATC 1560
CGTGCCGGCC CTAACCTGCG GCGTTGGCGC TGATCGGCTG GACGGTGGCG ACCTCGCCGAA 1620
<table>
<thead>
<tr>
<th>Met</th>
<th>Lys</th>
<th>Ile</th>
<th>Arg</th>
<th>Leu</th>
<th>His</th>
<th>Thr</th>
<th>Leu</th>
<th>Leu</th>
<th>Ala</th>
<th>Val</th>
<th>Leu</th>
<th>Thr</th>
<th>Ala</th>
<th>Ala</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Gly</td>
<td>Cys</td>
<td>Gly</td>
<td>Ser</td>
<td>Lys</td>
<td>Pro</td>
<td>Pro</td>
<td>Ser</td>
<td>Gly</td>
<td>Ser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>30</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Glu</td>
<td>Thr</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Thr</td>
<td>Val</td>
<td>Ala</td>
<td>Thr</td>
<td>Thr</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>45</td>
</tr>
<tr>
<td>Ser</td>
<td>Pro</td>
<td>Val</td>
<td>Thr</td>
<td>Leu</td>
<td>Ala</td>
<td>Glu</td>
<td>Thr</td>
<td>Gly</td>
<td>Ser</td>
<td>Thr</td>
<td>Leu</td>
<td>Tyr</td>
<td>Pro</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>60</td>
</tr>
<tr>
<td>Phe</td>
<td>Asn</td>
<td>Leu</td>
<td>Trp</td>
<td>Gly</td>
<td>Pro</td>
<td>Ala</td>
<td>Phe</td>
<td>His</td>
<td>Glu</td>
<td>Arg</td>
<td>Tyr</td>
<td>Pro</td>
<td>Asn</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>70</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Ala</td>
<td>Gln</td>
<td>Gly</td>
<td>Thr</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Ile</td>
<td>Ala</td>
<td>Gln</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>90</td>
</tr>
<tr>
<td>Ala</td>
<td>Gly</td>
<td>Thr</td>
<td>Val</td>
<td>Asn</td>
<td>Ile</td>
<td>Gly</td>
<td>Ala</td>
<td>Ser</td>
<td>Asp</td>
<td>Ala</td>
<td>Tyr</td>
<td>Leu</td>
<td>Ser</td>
<td>Glu</td>
<td>Gly</td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>105</td>
</tr>
<tr>
<td>Asp</td>
<td>Met</td>
<td>Ala</td>
<td>Ala</td>
<td>His</td>
<td>Lys</td>
<td>Gly</td>
<td>Leu</td>
<td>Met</td>
<td>Asn</td>
<td>Ile</td>
<td>Ala</td>
<td>Leu</td>
<td>Ala</td>
<td>Ile</td>
<td>Ser</td>
</tr>
<tr>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>120</td>
</tr>
<tr>
<td>Ala</td>
<td>Gln</td>
<td>Gln</td>
<td>Val</td>
<td>Asn</td>
<td>Tyr</td>
<td>Asn</td>
<td>Leu</td>
<td>Pro</td>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
<td>Glu</td>
<td>His</td>
<td>Leu</td>
<td>Lys</td>
</tr>
<tr>
<td>130</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>135</td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Gly</td>
<td>Lys</td>
<td>Val</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Met</td>
<td>Tyr</td>
<td>Gln</td>
<td>Gly</td>
<td>Thr</td>
<td>Ile</td>
<td>Lys</td>
<td>Thr</td>
</tr>
<tr>
<td>145</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>150</td>
</tr>
</tbody>
</table>

**Información para el SEQ ID No: 155:**

(i) **CARACTERÍSTICAS DE SECUENCIA:**
(A) **LONGITUD:** 374 aminoácidos
(B) **TIPO:** aminoácido
(C) **DIRECCIÓN:** lineal
(D) **TOPOLOGÍA:** lineal

(xi) **DESCRIPCIÓN DE SECUENCIA:** SEQ ID NO: 155:

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

(2) INFORMATION FOR SEQ ID NO:156:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1777 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:156:

GGTCTTGACC ACCACCTGGG TGTGAAAGTC GGTGCCCGGA TTGAAGTCCA GGTACTCGTG 60
GGTGGCGCGG GCCAAAACAT AGCCACAAGC ATGCGGACG CCGCGGTAGC CGTTGACGCGT 120
GTAGCGAAAC GGCAACGC GG CCGCGTGGG CACCTTGTTTC AGCGCTGATT TGCAACAAC 180
CTCTTGGAAG GTGATGCGGG CGAATTTGGG CGCGGAAAC CTGGGGACCA CGCCGATCCG 240
CTGCAACCGC GCAGCGCCGG TGTCGCAACGG GCATCCCGGT CACCGCGACG GCTTGGCCGGG 300
CCCAACGCACT ACCATTATTC GAACACACCTG TCTTACTTTC GTCAACGCTG GCCGCTACCG 360
AGCGCGCAGC AGGATGTGTAT ATGCCCATCTC TGCCCGCACA GACAGAGGGC AGGCCCTTATG 420
ACACGTATTG CGTGCAGGACC CTACGGGCAG CGGAAGTACC TAGAATAGCC CGGGAAGGCC 480
ATGGCSTATA TCGAGCAAGG CAAGGGTGAC GCCATCGTCT TCCAGCACGG CAACCCCAGC 540
TCGCTTTACT TGTCGGCGCAA CATCATCGCCG CACTTGAGG GGTGGGCGG GCTGTTGCCC 600
TGGGATCTGA TCGGGATGGC CGCGTCCGAC AAGCTCAGCC CATCGGAGAC CGACCCGTAT 660
AGCTATGGCG AGCAACAGGA CTTTTTGTTG CGGCTCTGGG ATGCGCTCGA CCTCAGGAC 720
CACGTGCTAC TGGTGCTCAA CGACTGCGGC TCGGCGCTCG TGCTCGACTG CAGCTAAAGC 780
ATCGCGCGCC GAGTGCAGGG GATCGCGTTC ATGGAAGGCA TGTCGACCAC GATGACGTGG 840
GCGGACTGGC CGCCGGCCGT GCGGTTGCTG TCCAGGGTTT TCGCATCCGC TCAAGGGGAG 900
CCAAATGGCGT TGAGCGCAAA CATCTTTGTCG GAACGGGCTCC TGCCCCGCGGC GATCTCGCGA 960
CAGCTCAGCG AGCAGGAATT GAACCACATG CGCGGCCCAT TGCTGAACGG CGGCGAGGAC 1020
CTCGCCTCCE CGATTGGCTGT GCCACGAAAC TTTCAATCG ACGTGGAGCC CGCCGAGGTC 1080
GTGGGGTGG CACCCAGTGG CCCAGCTCAGCTG CTCCGAAAGA CCGACATGCC GAAACTGTTG 1140
ATCAAAGCCG AGCCCGGCGC GATCATCACC GGCGCCATCC GTGACTATGT CAGGAGCTTG 1200
CCAAACCCAGA CGAAATCAC AGTGGCGGGC GTGCAATTTGG TGGAGGAGGA CAGGGATGCC 1260
GTGCTATCGT GGCGGGCGGC TGCGGAGCAT CGGGCAGCTG GGAGCGCTCT CATTTCAAGA 1320
GACCAAGATG GTGATTTCCG CGGAAGCGGG CGCCCGTGTT GTCAACTCTG AAGACTCTCT 1380
GCTCCGGCA GAGATTTCTCA GGAAAAGGG CACCAATCGC AGCCGGCTTC TCGCAAAGCA 1440
GGTGGACAAA TATACGTGGC AGGCAACAGG TCTTCTATT TGCCCGCAGCA ATTAGTGCCT 1500
GCCCTTTATG GGCGTCAAGT CGAGGAACCG GACGGATCTA CGCGTATCCG ATGGGACCTA 1560
TGGAACCCGT ATCATGGAAG CTCCGAAATC TTGGAAACGC GGGGGTCCCT GCCGCCGTCCG 1620
ATCATCCACG AGGGCTGCTC TCAACAGCC CACATCTACT AGCTTGTTAC ATGGCAGCGC 1680
GCCGATCCGG AGGAGGGTCTG GCGCGCTCTGG AGAGCGGAAG GTATAGGCGC GTGCTTTTAT 1740
TCAGTGCCGC TTCAGATTC GCAGGCACGGC CGTCGCT

(2) INFORMATION FOR SEQ ID NO:157:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 324 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:157:

GAGATTGAAT CGTACGCCTGCT CCCTTAGACGG CTCGATCCCG TGAATGCCCA TATCAGCAC 60

GGCCATGTTC TGGCTGCGT ACGCCCGGC ATGCGCCTCA GTGCTGAAC CCAGCGTTTG 120

ATCAGTAATTC CGCCGAGCAT GCGGGCAGGA GCGCGCCAGG ATGCGCCTGA CGCCGCGCCC 180

CGCCGTCGCG CAGCGACCAT CTGGATGGTC AGCCCGCGTG CGCGACGTGA GCGACGGGGT 240

GGCCGCGTGTG CTCCACAGTG GTACTCCCGTG GACGCACGGG CGCGTGCGCT GGGTGAAGAGC 300

CGTGACCGACC GCAGGCGATT CAGA 324

(2) INFORMATION FOR SEQ ID NO:158:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1338 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:158:

GGGTACCCGC CGCGTTTGCC GCCTTAGAGG CCGTAGGACC CTGAACTGAT CTTGCACGCG 60

AAACGTGCAC GAAACCGCTCG TGGCGCCGGT GTGGGTGCTC AAGGTTGCGG GTGATGGTCTG 120

CGATGCACGCG CGCGCGACCC GCCGACTACGG GTGCAAGGAC GTGCGCGGGCC GCGACGGCA 180

GCAGGCATTC CCAGCAGTAGG CGTCGAGCAGGG GTGGCTGCTA TGGCTGGCTG 240

GGAAGTCGCG GTTGGCGCGC AGATTGGCGT CCTCGATGAC CTGCGACCGC AGCAGCGCGG 300

GTGGGGCGCC TGGAGCGGAC GCGACGTGCG CGAAGTTGAG AAGCGCGCGG CGGTGCTGCG 360

GTCTGTATCA AGCTGGCGGC CATCGTGCCG TTCGCTGCGGC ACGGTTGCGG CTGGACGCGG 420

CTGAAATCAGC TAGTACGAGG CAGTTGGCGA CAGATCGCCG TCAGTGGGTC TGATGCGACAC 480

GAGTACGTAC ATACGTATTC GATGGACTGC ATGGACCCGTT TGATGCGCAGA GACCAAGAGAC 540

AAGGTGCCGG GGCAAAACGG CTGCAGAGATG ATCCGCGCTC CATCAGGCGCC CTGGCGCGGGTG 600

GGCGTCATTC GCAGATCGGT CCGGCTCGCG GATGCTGCTG GACGCCAGCG CTGAAAGAGG 660
GGAGCGCGGC GGTGACGGCG CTGCTGGACG GCCGGCACAG GGGTGCCTGTG CTTAAGGGCG 720
GGATCGGCAG GAAGCGGCTTC GAGGCTCGGGG CCCACGAGTTT CTTGGTTGCTT CCGACGGTGG 780
CGCACAGCCA CGCCGAGCGG GTCCGCCCCAG AGCACCGCGA CCTGCTGGGGG CGGCGGCAGCG 840
CCGACAGCAC CGACGAGCTG GTGCTACTGC GGGCGGCCAGC GAAAGTTGGT GCCGCACTGC 900
CGGTTAACCG GCCAGAGGCT CTGGACGCCCA TCGAGGATGCT GCACATCTGG ACGGCGAGCT 960
CGGTCGCGGC CGACCGGCTCG ACGTTTTGCGC CCAAGCACAAC ACTGCCGCTTC TTGGTGCGCT 1020
CGCGATCCCC GCTGGCCGAG CCGTGGCGCG TGGCAGCGTAG GCCCGAGTAC GCCGCGTTGCA 1080
CCAGCTGGGT GCAGCTGGCG GTGACGCGGCA GTTTGGGGGCG GCCGAGTCAG CAGGAGGCG 1140
CGCTGGCCGA GGTCCGGCGCC CGGGCTCCCGG AGGCCCTGGG TTGACTGGGC GGCATCGCTT 1200
GGCTCTGAGC TGTAAGCGCCA GTCGGCGCTG CAGAGTACCT GTGGCTCGTTT CGTCCCTGC 1260
TGGCTCAAT TGACGGCGGG GCCAACAGCA GCATTGGCGG GCCAATCTGC GGCAGCGCCG 1320
GCGCCCAACG CTACAAACC 1338

(2) INFORMATION FOR SEQ ID NO:159:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 321 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:159:

CCGCGGGCAG CGCCGAGCAG GCCGGTACGG GCCGACAGCC GCTGCTGGTG 60
TGGGCTTCCG CGCGAAGCGC GACCGTGCAA CGCTGGCGGG CAAAGCGGGT AAGGGGGG 120
TAGGTGGGCC CGCGGCTGACA GCGCGGCGGG CGCGGCGGAG GGCAAAGGCG 180
GCACCGGGCG TGCCGGCCCG GCCGGCAACG ACGCAGCGCAG CACGGGCAAT CGCGCGGCTA 240
AGGCCGGCGA CGGCAGGATC GCAGGTGGGCC GCAGGGCCGC GCAGGCGGGC 300
ACGGCGCACA TGCCGGCAAC C 321

(2) INFORMATION FOR SEQ ID NO:160:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 492 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:160:

GAAGACCGGG CCCGCCATA TCGATCGGCT CGGCCGACTAC TTTCCGCAGAA CGTGACCGG 60
GCAGCGTGG GCTGATCATC ACCGGTCGCT ACCGCCCAAC CGGCACCCGA TGGCTGCTGC 120
CGTTCGCCCT CGAACTCGTC ACTTCCCGGC AAGCCCGGACG GCACCGCCGGA ATCAACAGGG 180
CGTCCACGGA TTCGCGGTGCTA AAGATCCTGCT TGCAAAATCTC GACGCCGCGA CGCTACGCTCT 240
ACCACCACT TCGGTCGACG GCCTCGCGGA TCAAGGGGCC GATCAACCCCG TTTCGTCCGC 300
GAGCACTATG GGTCGC GGGA GTAAGAAGGGA CCGTCCCAGG TTTCCGCCC CGGCGCCCTG 360
TGCCCGCGGA TGCCGGCTAC GAGCGTGTCG AAATCATGGG CAGCGAAGGG TATCTGCTGCA 420
ATCGTTCCCT GCCGCCGGCC ACCAACAGGC GCACCGGACTCG TGGGGCCGGAG ACACCGGCCA 480
ACCGTCGCGCG GT 492

(2) INFORMATION FOR SEQ ID NO:161:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 536 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:161:

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

Ala Gln Phe Glu Glu Ala Glu Arg Ile Thr Arg Ile Arg Leu Asp Leu 450
455 460

Trp Asn Arg Tyr His Glu Ser Phe Glu Ser Leu Glu Gln Arg Gly Leu 465
470 475 480

Leu Arg Arg Pro Ile Ile Pro Gln Gly Cys Ser His Asn Ala His Met 485
490 495

Tyr Tyr Val Leu Leu Ala Pro Ser Ala Asp Arg Glu Glu Val Leu Ala 500
505 510

Arg Leu Thr Ser Glu Gly Ile Gly Ala Val Phe His Tyr Val Pro Leu 515
520 525

His Asp Ser Pro Ala Gly Arg Arg 530
535

(2) INFORMATION FOR SEQ ID NO:162:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 284 amino acids
(B) TYPE: amino acid
(C) STRAINEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:162:

Asn Glu Ser Ala Pro Arg Ser Pro Met Leu Pro Ser Ala Arg Pro Arg 1
5 10 15

Tyr Asp Ala Ile Ala Val Leu Leu Asn Glu Met His Ala Gly His Cys 20
25 30

Asp Phe Gly Leu Val Gly Pro Ala Pro Asp Ile Val Thr Asp Ala Ala 35
40 45

Gly Asp Arg Ala Gly Leu Gly Val Asp Glu Gln Phe Arg His Val 50
55 60

Gly Phe Leu Glu Pro Ala Pro Val Leu Val Asp Gln Arg Asp Asp Leu 65
70 75 80

Gly Gly Leu Thr Val Asp Trp Lys Val Ser Trp Pro Arg Gln Arg Gly 85
90

Ala Thr Val Leu Ala Ala Val His Glu Trp Pro Pro Ile Val Val His 100
105 110

Phe Leu Val Ala Glu Leu Ser Gln Asp Arg Pro Gly Gln His Pro Phe 115
120 125
Asp Lys Asp Val Val Leu Gln Arg His Trp Leu Ala Leu Arg Arg Ser
130  135  140

Glu Thr Leu Glu His Thr Pro His Gly Arg Arg Pro Val Arg Pro Arg
145  150  155  160

His Arg Gly Asp Arg Phe His Glu Arg Asp Pro Leu His Ser Val
165  170  175

Ala Met Leu Val Ser Pro Val Glu Ala Glu Arg Arg Ala Pro Val Val
180  185  190

Gln His Gln Tyr His Val Val Ala Glu Glu Arg Ile Pro Glu Arg
195  200  205

Glu Gln Lys Val Ser Leu Leu Ala Ile Ala Ile Ala Val Gly Ser Arg
210  215  220

Trp Ala Glu Leu Val Arg Arg Ala His Pro Asp Gln Ile Ala Gly His
225  230  235  240

Gln Pro Ala Gln Pro Phe Gln Val Arg His Asp Val Ala Pro Gln Val
245  250  255

Arg Arg Arg Gly Val Ala Val Leu Lys Asp Gly Val Thr Leu Ala
260  265  270

Phe Val Asp Ile Arg His Ala Leu Pro Gly Asp Phe
275  280

(2) INFORMATION FOR SEQ ID NO:163:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 264 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:163:

ATGAACATGT CGTCGGTGGT GGTCGCGCAAG GCCTTTGCGC GATTCGCAG CTACTCCCTC  60
GCCATGCAAG CGATCGCCGG TTTCTCGAT GCGTGCGCC AAGAGCTGCC GGGTAGCGGA 120
ATCGCGCTCT CGGTGATCCA CCCCGCGCCCTG ACCCAGACAC CGCTGTGCC CAACGTCGAC 180
CCCGGCGACA TGCCGCGCCC GCTTCGCGCC CTCAGCGCCTA TTCCCGTTCC CTCGGGTCGCC 240
GCAGCGGTCG TTGACGCGTG GCGG  264

(2) INFORMATION FOR SEQ ID NO:164:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 1171 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:164:

TAGTCGGCGA CGATGACGTC GCGGTCCAGG CCGACCAGGG CAAGCACCAG GGGGACCAGG 60
AGCCCGGTGC GATCCTTACC GCGGAAGCAG TGGTGAGCA CCGGGGCTCC GGGGCAAGGC 120
AGTGGACGCA CACGATGTAG CGCGGCGTGT GCTCCATTGCG GCGTTGCGAAG TGGCGATAC 180
TCGTCGGTCA TGGAGGGGCT GGGCAGCTGA TTTATCAGACT GCGTGGATTCT GCCGAGCTCG 240
CCGTTGGACC GCTCATTGCT TAGCAGCCTCT TGGAAATCGGG TTTGCTGGCG GGGCTGAGTCG 300
TCGGCGTCACT CATCGGGCAAG GTGGGGGAAG CCGACCGAGG GTAGTCGACT GGGGCTCCCGA 360
ACCCTCGCTG GACCCGCGCG GCGAACCTCC CGGGACGACC GCAGGTCGCG AACGTCGCGT 420
ATCCCCAGCC GGGCGACGGT TGCCCCCTGCT GCGAAATCCG GCACGAGCCG GGGGAGCCAC 480
CGGCCATCAC CAAGAACGCG TTGCCCCAGTA CGGATCCTCA ATTCGGCAGTCT GGGGAGCCAC 540
ATCTCTGTCG CGCCCATCTG CAGATCCCGCC TCGTGCCTTG ACAAAGACCG CCGCAGAGTT 600
GCCACGGGCT ATCGGAGATT GAACCGCGCA GCGATTTCTT CAATCGCTGC GGGCTGCCGC 660
ACTATTGCAA CTTCCGCGCG GTGCGCGGATG TCAGCAAGCGA TGGAGTCTCT GACGAACCTG 720
CCCCAGCTAAC CACCCGCGAT GCATCCGCGC TGGACGCCGAG GATGCCGGCG GTGATTCTTG 780
CGGCCAGCTG CGTACCGGTT GATCCACCGC TGGCGCGGTC GCGGGGGGAG GGGGAGCCAC 840
TTATCGACCT CGGCGTATGC CGACGGCAAG CTTGCGCGGT TCAGGAGGTT CAAGAACCTC 900
ACCATCGCGA CGCGCAGCGCA GGTGCGCGAC TGACCTTACG TCGGCGAGCG GCACATCGCG 960
GAGTACAGCA ACACTCGCGGC CTCCAGCGGT TGGTCAGCTA ACGACGGTAC TGCAAAACGG 1020
CGCACCAGGC TCGTTGCGCA GTCACCGAGC GGCTCGCGACA CCATGTTGCT GGGGAGCTGA 1080
ACCATCGCGG ACCGCCGCTGTA TACCACGGGG GCCACAGTTG TGGCGGAAAG TGTCGGCGGC 1140
GGGGCGCTGG CAGTGTGCCG GGGGCGCGCA C 1171

(2) INFORMATION FOR SEQ ID NO:165:

(i) SEQUENCE CHARACTERISTIC:
(A) LENGTH: 227 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:165:
GCAAAGGCGG CACCGGCGGG GCCGGCAGTA ACAGCCTCGA CCCCCTGCTTA GCCGCGGCAAG 60
ACGGCGGCAA AGGGCGGACC GCGGGCAGCG GCGGCAACGC GGCGGCGGCG GCACCCAGCT 120
TCACCAAGGG CGCCGACGGC AACGCCGCGA ACGGCGGTGA CGCGGGGTCG GCCGGCAAGG 180
GGCGGAAACGG CGGAAACGGC CGAGAACAAC CCACCGCCGC CGGCGG 227

(2) INFORMATION FOR SEQ ID NO:166:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 304 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:166:
CTCAGCACC ATGGGCGGAGG AGGGCGGTAG CGGTGGCGGC GGCTCTACCC CAGCGCGCAA 60
GGCGCCACC GGGTCACCTC CAACGCGCGG CGCGGCGCCGG GCGGACGGCG GCAACGGCGG 120
CAGCCTCCAA GTGTGGCGCG GCAAGCGCCG CGACGGCGGC AATGGCGGCA ACGCGCGCAG 180
CGCGGCGCAG GGGCAGACGC CGCGCGCGCG GGCTTTGGTG GCATAGTGCG 240
CAACGCCACC AACCTGCGGG AAAAGGCGGC AAAGGCTAACC CCCCAGGGCA ACGGGGCGGC 300
GGCG 304

(2) INFORMATION FOR SEQ ID NO:167:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 1439 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:167:
GTGGGACGCT GCCGAGGCTG TATAACAGGG ACAACATCGA CCAGCGCCCG CTGGGTGAGC 60
TGATGGACCT ATTTAACGAT GCCCGCTCCA GCCGGCAGGG CGAGCACCAG GCCGGGAGTC 120
TGATGGGTGA GGTCTACGAA TACTTCTCG GCAATTTCGC TCAGCGCGGA AAGAAGGGGG 180
GTGGGCGAGTT TTCCCTCCCG CCCAGCGCTG TCAAGGTGAT CTGGAGGTTG CTGGAGGCGT 240
CGAGTGCCGC GGTGTATGAC CGTGTGCCTGG GTTCGCGAGG CATGGTGTG CAGACCGAGA 300
AGTTTCATCTA CGGAGCAGCGG GCGATCAGGA AGGATGACTTC GATCTATGCG CAGGAAGACA 360
TTGAGGAGAC CTGGCGGATG GCCAAGATGTA ACTCTGCGCAT CCACGCGATC GACAAACAAGG 420
GGCTCGGCGC CGATGGAGT GATACCTTCG CCCCACGACCA GCACCCGGAC GTGCAAGATG
ACTACGTGAT GGCAATCCG CGTTCAACA TCAAAAGACTG GGCCCCGCAAC GAGGAAGACC
CAGCCTGGGC CTTCGGTTTG CGCAGCGACA ATACGGCGCA CTACGCAATT ATCCGCAACA
TCTGTGACAA CTTGCGCGCG GGAGGGCGGG GCGGCGTGTT GATGCCCCAC CGGCTGATGT
CGTGCAGACTC CAACCGCAAG GGGATATTG CGCGCAAAAT CTTGGAGGCG GATTTGTTTT
CGTCGATGCG CGCAGTACCC ACCGAGCTGT TCCGCAAGAC CGGAATCCCG GTGGTCCCTG
GGTTTTCTGC CAAAACAAG CGCGCAAGTA AGCAAGGTCG TATCAAGCCG TGCGGCGGAG
TGCTGTTCAT CGACGGCTCGG GAACAGGGAC ACCATTGTGGA GCGCGCAGCGA GGGCCGCTGA
CCAACGGAGA GATCGTGCGC ATCCGGGATA CTTTCACGAC GAGCAAGACC ACCGGCAAGC
CCGGCTCGCG TGGTCGCGGC GCTTATGGGG GCACTGCGGT CAAGCGCGCG CGGCTGGCTG
GGGGCGCGG CGCGAAGCGG GGGTGGCGGG CGGCTGCTCT CGGCAACGCT GTGCGGCGCG
ACGGGCGCA CGCGCGCAAC GGGCGAAGCG CGCGACCGGG CACGACGCGG GGGCGCGCGG
GCAAGGGCGG CAACGGCAGC AGGCGGTGGG CCCAGGGCTC AGGCGTGCTC AAGTCAAGGC
CGGCGCAGCG CGGAACCCGG GCAATTGGGG GCAACGGCGG CAACCGGCTCC GCAGGGCGCG
CGGCGCAGGG CGGGTGCGGG GCAGCGCGGG CGCGCGCGGG CGGCGCGCGG CGGCGCGCGG
CGCGCGGGGG CGCGCGCGGG GCGCGGAGG CGCGCGCCGG CGCGGCACGG CGCGCGGGG
TCACCGTCAC CGCGGCGCCG CGCGGCAAGC CGGCGCAATGG CGCGCAACGC GCACGGCGGC

(2) INFORMATION FOR SEQ ID NO:168:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 329 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:168:

GGCCCGGCAG GGCCGGAATT TCTCGTGGCT TGATGGTGCC TGAGGATAAC GGGCGGCTATG
GTGCGTAACGG CGGGATGGGC GGGCCTGGGG GGGCCTGGCG CGGGGGCGGG GGGCGGGCGG
TGATCAGCCT GCTGCGCGCG CAAGGCAGCG GAGGGCGGGG CGGACCGCGG GGGCGGGCGG
GTGTTGCGGG TGAGGGCGGC CGCGGCGCGG CGCGCAACCC GGCTTTCAAC CGAGGGCTGG
GCGGGCGCG CGGGCTGGATC AGCCCGTGCG GGCGGCAAGG CGGCCGGCGG GGGCGGGGGA
CGCGCGCGCG CGCGGCGCAT GCGGCGGCGG

60
120
180
240
300
329
(2) INFORMATION FOR SEQ ID NO:169:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 80 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:169:

GCAACGGTGG CAACGGCGGC ACCAGCACGA CCGTGGGGAT GGCGGGAGGT AACTGTGGTG 60
CCGCCGGCT GTACGCGCAAC 80

(2) INFORMATION FOR SEQ ID NO:170:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 392 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:170:

GGGCTGTGTC GCACTCACC CGCCGCTTC GGCGACGTG TCGCCGCCAAT ATCCAGCTCA 60
AGGCCCTACTA CTGACGCGTG GAGGACGCC GCATCAAGGT GCAGGTCAGC CCCAAAGGAA 120
TCAAGGTCAT CGACCGCGAC GGCGCTGAG GCCGTCGTGC GCCGGCTCGG GCAGGATCCG 180
CCCCGGGCCA CTTGCGCGGC CAACGCGGCT CATCGCTCCG AACCAGCGGCC ATCCCTGTGAG 240
CACAACGTAG GGGCCGCAAC GAGATGTGTC CAATTGTCAA GGGGTGTTCG ACCGGAGGGA 300
CCGGTTATAC GTATGTCAC ATATGTCAT GCACAAGACC GCGATAACGA TCCGGTGATC 360
CGCCGACAGC CCAGAGGTGC AAAGCGTTTA CA 392

(2) INFORMATION FOR SEQ ID NO:171:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 535 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:171:

ACGGCGGCCA CGCGCGGCAC CGGTGCGGCC GTGTTGCGGG CGCGGGCCGG GCGCGAGGGG 60
GGTATCAGCG GTGCCGGCGG CACCAACGCG TCTGGTGCCG CTGGCCGGCCG CGCGGGACAA 120
GGCGGCGCGG GGGCGCTGG GCGGGCGCGG CCGATAACC CCACCGCGCAT CGCGGGCGCC 180
GGCGGCACCG CGGCGACCGG CGGAGCCGCC GGAGGCGCGG CGGCGGCTGG CGCCATCGGT 240
ACCGGCAGCA CGGGCGGCAG GCGGGCGAG CTGGTGAGTG CAGCGCGGTC CTGGTGCCGC 300
GGTACGGTTG TGGTGAGTGG TGGTGAGTGG CGGGCGGCGG CTGGTGAGTGG TGGTGAGTGG 360
GTCACCGGAT GCACGCGAGG CGGCGCGGCGG GGGCGGGGAG AAGGGCGACC GGGCGGCAAC 420
AGCGGTGAGG CGGGCGAGGCG CGCTGCACGC GGCCGGCGCC GACCTGAGGCG CAGGCGCGCC 480
ACCGGAGGTTG CCGCGGCGAGAC AACCACCGAG GTGCTGAGTTT CGCGCGCG 535

(2) INFORMATION FOR SEQ ID NO:172:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 690 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:172:

CCGACGTGCG CGGGCGGATA CGGGCGACAC CGACTACTAC ATCATCGGCA CCGAGAAATCG 60
GCCGCTGCTG CAACCCGCTGC GGCGCGGCGG GGGCGCTGGA GATCGAGCTG CCGCACTCGT 120
CCAGCCGAGG CGGGCGGCTG CGGGCGGCTG CGGGCGGCTG CAGGGCGGCTG CCGCACTCGT 180
GAGCAGCTGAC GGGCGGTGGA GGGCGGGTGG CCGGGCGCTG CAGGGCGGCTG CCGCACTCGT 240
CATGCGGAGG GCCGCGGAGG CGGCGGCGGAG CTGGTGAGTG CGGCGGCGGAG CCGCGGCGGAG 300
GCCAGCGGTG CTGGCGCGAC CGCTCGCGCT CCGCGAGATCG CAGCTGCGCG ACACCGCGCG 360
TCTGGTGCGC CGGGCGGCGG CGGCGGCGGAG GCGGGCGGAG GTGCTGAGTT GCGGGCGGAG 420
GATCATCTCA ACCGAATCGA CGCGCTCGGT CGGCCACCGT GACATCGGC TGCGCTGAGT 480
ACCACCTGAG CGCGTCACA CACGAAGCTG TTGGCTGAGG AACTGCGTGC GGGCAACTCG 540
ATCAACGACA TGCGTATCCG CCTGGCGGCGG ACCGTAGGTT TAGCGACGAT CGATAGCGG 600
CGGCGGAGGAA TTGGTTGGGAG GCCGTTGACGG GCCGTTGACGG ACCGTACCAG ACACGAGG 660
CCTGCTGATC TAACGGGATC CCGAGCGGCT 690

(2) INFORMATION FOR SEQ ID NO:173:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 407 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:173:
ACGGTGACGG CGGTACTGGC GGCGGCCACG GGGCAAGCGG CGGGAAATCCC GGGTGGCTCT  60
TGGGCAACGC CGGGGGTTGGC GGCAACCTTG GGCAGGCGAC AGCCCCGACT GCAGTGGCCG  120
GCTCTGgggg CACCggCggG GACGgggGgGA CCGGGggGGG GGGGGGCTTG TGATGGGGCG  180
CCGGCCGGCC GGGCCgACGG GCACgACTggCC gCCGGGGGGG TGCGGGTTGC gACGgGGGGCg  240
GGGGGGGGGG GGGGGggGGG gGGGGGGGCA gGCGGGGGCG GGGGGTTCAA GGGGGCCCTGC  300
TGGTggGGGG GGGGGGGGGG GGGGGggGGG GGGGGGGGAC gGCGGGATGG gGTGGGGGGG  360
GTgACggGgCT CGAGGGCGGC ATGGGGCCGC TGGGTTGgATAC CGGgTGGC  407

(2) INFORMATION FOR SEQ ID NO:174:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 468 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:174:

GATGGGTCA GCGCATCGCC TGGGCgGgCA gGCGATCCAC CGGgTTCAACG AAGAACATCG  60
TGCCGGGCgG GGGGCgGGgA GCgCCGGCTGC gGGCGGGGGG GGGTGACCGG TCCAGCGGAC  120
ACGAGCGTCC TGGGCggGgCC TGGGACGGgA ACGATgCGTG GTCAACGGTG TgAGATCgCCg  180
GGATGGGGGC CTGGCCgCAAC GCATTCCGGG GGCgCCGGGC GTCTTgGTgA TGCTGGACgA  240
TGACggGGAT GTCTGGGGGC ACCAGGGgGC gGGGGGGGGG gGTGGGGGGG CTGGgGgATgA  300
GGGGGGGGGC GGGGGGGgGC gGGGGGGgGA GGAGGGGGGG GTGGGGGGGG CGGGGGGGGC  360
GGGGGGGGGG GGGGGGGGGG GGGGGGGGGG GGGGGGGGGG ATATCCgTTG GGGGGGGGGG  420
CGATAATGGG GGGCGATGGC CCAGGGgCGT ACAGGGCGGC CACCACgCG  468

(2) INFORMATION FOR SEQ ID NO:175:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:175:

GGTGGAACCG GGGCGACGGG TGCGATCGGC GGCgCCGGCC AGAGAGGGGC CGACGGCGCC  60
GGgcccCAATG CTAAGGGCGC AAGGGCGGAC AAGGGGGGGG GGGTGgGTAA GGCTGGGGGC  120
GGCGGCGCCG GCAGCAATGG CGCCGGCGCC GCAGCAACGC CGCGCGGCAG GTCAGACGAC 180
GGCGGCACGG GCACCGGCCG CGACCGGGCG GCAGCGCCGC 219

(2) INFORMATION FOR SEQ ID NO:176:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 494 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:176:
TAGCTCCGGC GAGGGCGGCA AGGGCGGCGA CGCTGGCAAC GCAGCTGGACG CCCTGCAGCCG 60
CAACAGTCTCC GTCACCGGGCG GGACCGGGCGA CGCTGGCCCGC GCAGCCGCGGC GCGCGCGGGGG 120
CGCAGGGCCGTT TGGCGCGCGCA AGGCGCGGCTT GGCAGCAGCCGAG GCGGTCAGAG GCCGCCCCGAA 180
CGCGGGCGGT ACCGTCAGCA CGCTGGCAGGG TGGCAGCCGGC AACGGCGCTG TGGCGCGCCG 240
GGCGGGCGAC GCGTCTGTGG CGCCGGCGGG CGCCGAGGGG GCCTGGCAGTG GCCAGGGCGG 300
CAATGCGGGCG GCTCCAGCG CAGGCGGCGC CGCCGCGGGG GCGCGGCCGCG GTGGCAGGGG 360
CAAGCGCGCG GCTCAGGGCCG CATTCCGCGCT GACCTGGGAC AGCGGCGGCA AATGCGCTGC 420
CATCGCATCA GGGCGGCTACT GCCCGGAACA CCTGGGATG GCCCGGGCGGC 480
ATTTCCGATG CACC 494

(2) INFORMATION FOR SEQ ID NO:177:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 220 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:177:
GGCCCGGTGG TGGCGCGGGCG CAGCTCTTTAC GCGCGGAGGG CGCGCGCCGGT GCGCTGGGGG 60
TTGCGCGCGCCA CGCGCGCGCA AGGGGGGCTT GCGGTTCCGGC AGGCGCGGCG GCAGCGCCGGG 120
CGCGCCGCAAG AAGTCTACGG GTGGTTACCGG GGGTCGCGGG GGGCGGCGCG GGGCGCGGGGGG 180
GCCAGAGCGG CAAAGGCGGTT GCCGGCGGCCGA TCAGCGGGCCT 220

(2) INFORMATION FOR SEQ ID NO:178:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 388 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:178:
ATGGCGGCAA CGGGGGCCCC GCCGGGTGCTG SCGGGGCCGG CGACTACAAT TTCCACCGGC 60
GGCGAGGGTG GTGCCGCCGG CCAAGGCGGC CAAGCCGGCC TGGGCGGGGC AAGCACCACC 120
TGATCCGGCT AGGCGCACC GCCGAAAGCG ATCCAAACGG CGACGATGCC GCCTTTCTTG 180
CCGGTTTGGG CCAGGCCGCG ATCCACCTAC GTGACCCAGG CCACGACATA ACGGCCGCA 240
AGGCGATGTT GGGGCTGTTG GCTAAGGCGG TAACAGGTCT ACAGCTGGTC CCGGACCTGC 300
GGGACTACAA TCCGGGCTG ACCATGGACA GCGCGGCCAA GTTGCGCTGC ATCGCACAG 360
GCGGTACTG CCCCCAACAC CTGGAAACA 388

(2) INFORMATION FOR SEQ ID NO:179:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 400 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:179:
GCAAAGGGCG CACCGGCGGG GCCGCGATGA AGACGCTCGA CCGGCTGCTA GCCGCCCAAG 60
ACGGCCGCCA AGGCGGCAAC GCCGGCCACG CCGGCAAGCG GCCGCAGGGC GCCACACCCG 120
TCACCCAAGG CGCGCAAGGC AACCGGCGGA CGCGCCGCTA CGCGGGGTCT GCCGCCAACGG 180
GGGAAACCGG CCGAAAGCGG GCAGACAACA CAACACCCCGG GCCGGCGGCC ACCACCGGG 240
GCCGACGCGG GCCCGCCGGG GCCGGCGGAA CGGGCGGAAAC CGGCAGAGCC GCCGGCACC 300
GCACCCGGCG CCAACAGGGG AACCGGGCGA AGGCGCGCCGG CCCGGCGCAA GCCGCCACCG 360
GCGCGACGG TGACACTTCTCA GCCAGCCACGG GTGGTGCCGG 400

(2) INFORMATION FOR SEQ ID NO:180:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 538 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:180:
GGCAACGGCG GCAACGGCGG CATCGCCGGC ATGGGGCGGC AAGCGGCGTTC CGGGACGGGC 60
AGCGGCAACG GCGGCCAACG GCAGCAGGCG GGCAACGCGC GCCATGGGCGG
CAACCCGCGG ACCGGCAGCG GCGACGGCGG TGCCTGGCGGG AACGGCGCGG GCGGCCGAC
GGCGCGCACC GCAGCGCGAC GCAGCGTCAC CGTGACTGGC GCACCGCGCG AGACGGCGTG
CACGCGCGGT GCAGCGCGTA ACAGCGCGCAC CGAGGAGAT AACACCGCAA ACATGACTGC
GCAGCGCGCG GGTGACGGTG GCAAGCGCGG CGACGGTGGA TCGCCGGCGG GGGCGCGGGC
CGCGCGCGGT GGTGACGGTG CGCGCGCCCGT CGCCGCGGCA CGCCCGCGGC GCGCGCGCCG
CGATGCGCGG AACGGCGGCA TGCCGCGGCA CGCCCGACTC ACTGACGACC CGCGCGCGCA
CGGGCGACCC GCAGCGCAACG CGCGCCACCG GCAGCGCGCA TCGGCACG

(2) INFORMATION FOR SEQ ID NO:181:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 239 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:181:

GGCGCGCTGG TGCCGCGGGC CGCTCTTCA GCAGCGCGAGG CGCGCGCGGT GCAGTTGGGG
TTGGCGCAC CGCGCGCCAG GGTGGGCTGG CGTGCGGGGG AGGGCGCGGC GCAGCGCCCG
CGCCAGCAC AGGTCTACCC GGTGGTACCG GTGTCGCGCC CGCGCGCGGC GCAGTGCGGC
GCCAGCGGG CAAGCGCATT GCAGCGCGGCA TCAACGGCCTC CGTGTTGCC GCAGCGCACC

(2) INFORMATION FOR SEQ ID NO:182:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 985 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:182:

AGCAGCGGCTA CCGTGCGGCG CGGGTCGCGC GCAGGAGAAGG CGGAAGCGGGC
GGCAACAGCG GTGTGGGCGG CACCAACCGG TCCGGCCCGC CGGGCGTGGC AGGGCGCAAG
GGCGCGCACCG GAGGTGCGCG CGGTCGCGGC GCAGACAACC CCACCAGGTGC TGTTTTGCGG
GTTGGCGCG CGCGCAACAG TGCGCCGGCC GGGCGCGGCG GGCGCGCGGG GCAGACGGGT
ACGGCGCGCA CGCGCGCGGT TGTCGGCGCC ACCGTTAGTG CAGCCATCGG CGGGCGGCGC
GGCCCGCGGG GTGACGGGCG CGATGGGGGC AGCGGCTCGC GCCTGGCGGCT CTGGCGCTTT
GACGGCGGCC AAGGCCGCCA AGGCAGGCGC GCAGGCAAGG CCGCGCCGGC GCGCATCAAC 420
GGGCGGCGGC GGCCCGCAGG CACCGGCGGC GACGGCGCGC AGGGGCCGAG CGGCGCCGCA 480
GGTCTCGCGGC ACAACCGCGGC GCTGCAGGCG GACGGTGCGG CCGGTCGGCG CGCGCGCAAC 540
GGCGCCAGCC CGGCGCTGGG CCTGACAGCC AAGGCCGGCC AGCGGCGCGC CGCGGGCAAT 600
GGCGGCAACG GGCGCGCGCG CGTGCTGGGC GGCGCGCGCG ACGAATTTT CAACGCGCGC 660
CGGGTGTGGT CCGGGCGCCA AGGGCGCAAA CGGGCTTGGG GCGGGCAAG CACCACCTGA 720
TCGGCTTAGC CGCAACCGGC AAAGGCGATC CAACGAGGCA GCAAGCGGCC TCCCTTGCGC 780
CGTGGAACCA GGGCGCGATC ACCTAACCTG ACCAGGCGCA GGCCATAACG GCCGCGCAAG 840
CGATGTGTGG GCTGGTGCTG AAGCGGTCGA CAGGGTCTACG GCTGGTGCAG GACCTGGCGG 900
AATACCAATCC CGGGTCGACC ATGGACAGCG CGGCCAACCT CGCTGCCATC GCATCGGCCG 960
CGTACTGGCC CGAAGACCTG GAACA 985

(2) INFORMATION FOR SEQ ID NO:183:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2138 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:183:

CGGCAGGAGG ATCAGGTACCC CGGCAGCATCG GCAGCTGCAG ATTCGGCGGG TTTCCCGACC 60
CGAGGAAGCG CGCGTACGGT TGGCGCGTGC GAGTAGGGGC GATCGGTTGC CGATGCGGGC 120
ATGAAGGGGC GGCATCAAAT TAGTGCAAGA ACCTTGCGAT TTAGCGACGA TAAATGCTAT 180
AGCAATAGGG AAGATGATCC GATATCGACG AGTGCAGAGA CCGTGACGTTG GATCAGCAAG 240
AGATTGCGA CAGGGCGCAAC GAGGTGAGGG CCGCGATGCC GGACCCACGG ACTGATGTCC 300
CCATCACAAG GTGCGAACTC AGGCAGGCTA AAACCGCCGC CCAACAGTGC GATTTGCGCTG 360
CCGACAACAT CGGGGAATCC CTGGCGCGCG GTGCCAAAGA GGGCGACGGT CGGGCGACCT 420
CGCTGCGCAA CCAGGCGCAAC GCGTGATTGC AGGTTGATGA GGAGGCTGCG ACCGGCGCTG 480
AACAAGCAGG CGAGGAAACT GTGCGAAGGC AAGTGGCGGC GGGCCTGGGA GGGACAGTT 540
CGCGCGAACT AACAGGAGCG CCGAGGTTGG CCGACGCGGG TGAAACCAAC TTGATGGAATC 600
TCAAAGAGGC GGCAAGGAGG CTCGAAAGCG GGCAACAGGG CGCATCGCTC GCAGCTTTTG 660
(2) INFORMATION FOR SEQ ID NO:184:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 460 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:184:

Met Thr Gln Ser Gln Thr Val Thr Val Val Asp Gln Gln Glu Ile Leu Asn
1 5 10 15

Arg Ala Asn Glu Val Glu Ala Pro Met Ala Asp Pro Pro Thr Asp Val
20 25 30

Pro Ile Thr Pro Cys Glu Leu Thr Ala Ala Lys Asn Ala Ala Ala Glu Gln
35 40 45

Leu Val Leu Ser Ala Asp Asn Met Arg Glu Tyr Leu Ala Ala Gly Ala
50 55 60

Lys Glu Arg Gln Arg Leu Ala Thr Ser Leu Arg Asn Ala Ala Lys Ala
65 70 75 80

Tyr Gly Glu Val Asp Glu Ala Ala Thr Ala Leu Asp Asn Asp Gly
85 90 95

Glu Gly Thr Val Gln Ala Ala Glu Ser Ala Gly Ala Val Gly Gly Asp Ser
100 105 110

Ser Ala Glu Leu Thr Asp Thr Pro Arg Val Ala Thr Ala Gly Glu Pro
115 120 125

Asn Phe Met Asp Leu Lys Glu Ala Ala Arg Lys Leu Glu Thr Gly Asp
130 135 140

Gln Gly Ala Ser Leu Ala His Phe Ala Asp Gly Trp Asn Thr Phe Asn
145 150 155 160

Leu Thr Leu Gin Gly Asp Val Lys Arg Phe Arg Gly Phe Asp Asn Trp
165 170 175

Glu Gly Asp Ala Ala Thr Ala Cys Glu Ala Ala Ser Leu Asp Gln Gln Arg
180 185 190

Gln Trp Ile Leu His Met Ala Lys Leu Ser Ala Ala Met Ala Lys Gin
195 200 205

Ala Gin Tyr Val Ala Gin Leu His Val Trp Ala Arg Arg Glu His Pro
210 215 220

Thr Tyr Glu Asp Ile Val Gly Leu Glu Arg Leu Tyr Ala Glu Asn Pro
225 230 235 240

Ser Ala Arg Asp Gin Ile Leu Pro Val Tyr Ala Glu Tyr Gln Gln Arg
245 250 255

Ser Glu Lys Val Leu Thr Glu Tyr Asn Asn Lys Ala Ala Ala Leu Glu Pro
260 265 270
Val Asn Pro Pro Lys Pro Pro Pro Pro Ala Ile Lys Ile Asp Pro Pro Pro Pro 275 280 285
Pro Pro Gln Glu Gln Gly Leu Ile Pro Gly Phe Leu Met Pro Pro Ser 290 295 300
Asp Gly Ser Gly Val Thr Pro Gly Thr Gly Met Pro Ala Ala Ala Pro Met 305 310 315 320
Val Pro Pro Thr Gly Ser Pro Gly Gly Gly Gly Leu Pro Ala Asp Thr Ala 325 330 335
Ala Gln Leu Thr Ser Ala Gly Arg Glu Ala Ala Ala Leu Ser Gly Asp 340 345 350
Pro Ser Ala Pro Leu Gly Ser Ala Ile Gly Gly Ala Glu Ser Val Arg 370 375 380 385
Pro Ala Gly Ala Gly Asp Ile Ala Gly Leu Gly Gln Gly Arg Ala Gly 390 395 400
Ala His Gln Gly Gly Gly Ala Lys Ser Lys Gly Ser Gln Gln Glu 420 425 430
Asp Glu Ala Leu Tyr Thr Glu Asp Arg Ala Trp Thr Glu Ala Val Ile 435 440 445
Gly Asn Arg Arg Arg Gin Asp Ser Lys Glu Ser Lys 450 455 460

(2) INFORMATION FOR SEQ ID NO:185:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 277 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:185:

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

Thr Thr Arg Arg Ser Gly Arg Arg Asn His Arg Cys Gly Trp His Pro
65
70
75
80

Gly Thr Gly Ser His Thr Gly Ala Val Arg Arg Arg His Gln Glu Ala
85
90
95

Arg Asp Gln Ser Leu Leu Leu Arg Arg Gly Arg Val Asp Leu Asp
100
105
110

Gly Gly Arg Leu Arg Arg Val Tyr Arg Phe Gln Gly Cys Leu Val
115
120
125

Val Val Phe Gly Gln His Leu Leu Arg Pro Leu Leu Ile Leu Arg Val
130
135
140

His Arg Glu Asn Leu Val Ala Gly Arg Arg Val Phe Arg Val Lys Pro
145
150
155
160

Phe Glu Pro Asp Tyr Val Phe Ile Ser Arg Met Phe Pro Pro Ser Pro
165
170
175

His Val Gln Leu Arg Asp Ile Leu Ser Leu Leu Gly His Arg Ser Ala
180
185
190

Gln Phe Gly His Val Glu Tyr Pro Leu Pro Leu Leu Ile Glu Arg Ser
195
200
205

Leu Ala Ser Gly Ser Arg Ile Ala Phe Pro Val Val Lys Pro Pro Glu
210
215
220

Pro Leu Asp Val Ala Leu Gln Arg Glu Val Glu Ser Val Pro Pro Ile
225
230
235
240

Arg Lys Val Arg Glu Arg Cys Ala Leu Val Ala Arg Phe Glu Leu Pro
245
250
255

Cys Arg Phe Phe Glu Ile His Glu Val Gly Phe Thr Gly Arg Gly His
260
265
270

Pro Arg Arg Ile Gly
275

(2) INFORMATION FOR SEQ ID NO:186:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 192 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:186:
<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arg Val Ala Ala Ser Phe Ile Asp Trp Leu Asp Ser Pro Asp Ser Pro</td>
<td>1 5 10 15</td>
</tr>
<tr>
<td>Leu Asp Pro Ser Leu Val Ser Ser Leu Leu Asn Ala Val Ser Cys Gly</td>
<td>20 25 30</td>
</tr>
<tr>
<td>Ala Glu Ser Ser Ala Ser Ser Ala Arg Ser Gly Asn Gly Ser Arg</td>
<td>35 40 45</td>
</tr>
<tr>
<td>Trp Thr Ser Met Pro Ser Gly Thr Arg Pro Gly Pro Arg Arg Ala Thr</td>
<td>50 55 60</td>
</tr>
<tr>
<td>Ser Arg Asp Arg Arg Ser Ala Thr Ser Val Ile Pro Ser Arg Arg</td>
<td>65 70 75 80</td>
</tr>
<tr>
<td>Ser Val Ala Pro Arg Ala Glu Phe Gly Thr Arg Leu Ala Ser His Arg</td>
<td>85 90 95</td>
</tr>
<tr>
<td>Ala Ser Pro Ser Asn Ala Cys Pro Val Arg Ile Val Thr Ser Ala Ser</td>
<td>100 105 110</td>
</tr>
<tr>
<td>Gly Arg Pro Ile Ser Ser Pro Pro Ile Val Arg Ser Arg Ser Cys Val</td>
<td>115 120 125</td>
</tr>
<tr>
<td>Asp Lys Asn Gly Arg Arg Cys Ala Ser Gly Tyr Arg Arg Leu Asn Arg</td>
<td>130 135 140</td>
</tr>
<tr>
<td>Ala Arg Ser Ser Ser Ile Ala Ala Ala Cys Arg Thr Ile Gly Thr Phe</td>
<td>145 150 155</td>
</tr>
<tr>
<td>Arg Arg Ser Arg Tyr Ser Ala Ser Met Arg Val Ser Thr Asn Ser Pro</td>
<td>160 165 170 175</td>
</tr>
<tr>
<td>His Val Thr His Gly Val Ala Pro Gly Val Thr Arg Arg Ile Gly Gly</td>
<td>180 185 190</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:187:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 196 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:187:
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Glu Arg Pro Gln Met Cys'Gln Arg Val Ser Glu Ile Glu Pro Arg</td>
<td>1 5 10 15</td>
</tr>
<tr>
<td>Thr Gln Phe Phe Asn Arg Cys Ala Leu Pro His Tyr Trp His Phe Pro</td>
<td>20 25 30</td>
</tr>
<tr>
<td>Ala Val Ala Val Phe Ser Lys His Ala Ser Leu Asp Glu Leu Ala Pro</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
</tr>
<tr>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>Arg</td>
<td>Asn</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td>Ile</td>
<td>Phe</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>Ala</td>
<td>Gly</td>
</tr>
<tr>
<td>100</td>
<td>105</td>
</tr>
<tr>
<td>Gln</td>
<td>Gly</td>
</tr>
<tr>
<td>115</td>
<td>120</td>
</tr>
<tr>
<td>Gln</td>
<td>Gln</td>
</tr>
<tr>
<td>130</td>
<td>135</td>
</tr>
<tr>
<td>Gln</td>
<td>Thr</td>
</tr>
<tr>
<td>145</td>
<td>150</td>
</tr>
<tr>
<td>His</td>
<td>Val</td>
</tr>
<tr>
<td>165</td>
<td>170</td>
</tr>
<tr>
<td>Arg</td>
<td>His</td>
</tr>
<tr>
<td>180</td>
<td>185</td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
</tr>
<tr>
<td>195</td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:188:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 311 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:188:

Val Arg Cys Gly Thr Leu Val Pro Val Pro Val Pro Met Val Glu Phe Leu Thr  
1  5  10  15
Ser Thr Asn Ala Pro Ser Leu Pro Ser Ala Tyr Ala Glu Val Asp Lys  
20  25  30
Leu Ile Gly Leu Pro Ala Gly Thr Ala Lys Arg Trp Ile Asn Gly Tyr  
35  40  45
Glu Arg Gly Gly Lys Asp His Pro Pro Ile Leu Arg Val Thr Pro Gly  
50  55  60
Ala Thr Pro Trp Val Thr Trp Gly Glu Phe Val Glu Thr Arg Met Leu
Ala Glu Tyr Arg Asp Arg Arg Lys Val Pro Ile Val Arg Gln Arg Ala
   85
Ala Ile Glu Glu Leu Arg Ala Arg Phe Asn Leu Arg Tyr Pro Leu Ala
   105
His Leu Arg Pro Phe Leu Ser Thr His Glu Arg Asp Leu Thr Met Gly
   125
Gly Glu Glu Ile Gly Leu Pro Asp Ala Glu Val Thr Ile Arg Thr Gly
   140
Gln Ala Leu Leu Gly Asp Ala Arg Trp Leu Ala Ser Leu Val Pro Asn
   155
Ser Ala Arg Gly Ala Thr Leu Arg Arg Leu Gly Ile Thr Asp Val Ala
   175
Asp Leu Arg Ser Ser Arg Glu Val Ala Arg Arg Gly Pro Gly Arg Val
   190
Pro Asp Gly Ile Asp Val His Leu Leu Pro Phe Pro Asp Leu Ala Asp
   205
Asp Ala Asp Asp Ser Ala Pro His Glu Thr Ala Phe Lys Arg Leu
   220
Leu Thr Asn Asp Gly Ser Asn Gly Glu Ser Gly Glu Ser Ser Gln Ser
   235
Ile Asn Asp Ala Ala Thr Arg Tyr Met Thr Asp Glu Tyr Arg Gln Phe
   255
Pro Thr Arg Asn Gly Ala Gln Arg Ala Leu His Arg Val Val Thr Leu
   270
Leu Ala Ala Gly Arg Pro Val Leu Thr His Cys Phe Ala Gly Lys Asp
   285
Arg Thr Gly Phe Val Val Ala Leu Val Leu Glu Ala Val Gly Leu Asp
   305
Arg Asp Val Ile Val Ala Asp
   310

(2) INFORMATION FOR SEQ ID NO:189:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 2072 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:189:

CTCGTGCCGA TTTGCGACAGA GCTGAGCAGC CCAAGGGGCC GTTCGGCGAAA GTCATCGAGG 60
CATTCGCCGA CGGGCTGCGCC GGCAAGGTGA AGCAAAATCA AACCAGCTTG AACAGCCTGT 120
CGCAGGCCTT GAACCCTTGG AATGAGGGCC GGCAGGACTT CTTCGGCGTGG GTACCCAGCC 180
TGGCGCTATTT CTTCAACGCGC CTACATCAGG AGCAAAACAAC GTCGCTCGCG TTTGAAACAGA 240
ACCTTGCAGA GTTCACCGAC AGGTTCACCC ACTCCGATGC GGACCTGTGC AACGCCATCC 300
AGCAATTCGA CAGCTTGCTTC GCGTCGCGCC GCGCTTCTCT CGCACAAGAAC CGCGAGGTGC 360
TGACGCATGA TGCTCATAAT TCCGCGACCC TGACCCACAC GTTGGTCCAG CCCGATCCGT 420
TGATGGGTTT GGAAGCCGTC CTGACATCCT TCCGAGCGCT GGCAGGCCAAC ATTTACCAGC 480
TTTACCATGC GACACAGGTT GGCCTTGGTGT CGCTTTTCCGC GTTCAGCAAT TTGGCCAACC 540
CGATGGAATT CATCTGCAAC TGATTCAGGC CGGTAGCGGC GCTCGTTTAT CAAGAGTCGG 600
CGAAGCTCTG TGCCGATGTAT TGCGGCACAG TCCCTCGATGC GATCAAGGTC AACTACTTTT 660
CGTTCGGCCT GAACGGGCCC AGCCACCGCT CGACAACGCC TAAAGAGATGC CGGTACTCCG 720
AGCCCCGCTT GCAGAGCGGC AACGGTTACA AGSACACCAC GGTGCGCGCCG ATTGGGTGC 780
CGGATACGCC GTTGTCACAC CGCACAAGCC AGCCCCGTTTG GGTGGTTGGCA CCGGGGATGC 840
AAGGGGTCTA GGTGGGACCG ATCACGCGAG GTTGGCTGAC GCGCGAGTCC CTGGCGAAAC 900
TCATTGGGTC TCCGCAATTC GCCCTCTCGT GTCCAGGGCT GCAAACCCCG CCCCGAACCC 960
CGAATGCTTA CGACAGGTAC CCGGGTGCTC CGCGATCGGC TTACAGGGCC CCACAGTGTC 1020
CGATACACC GCCGCTCTCT GGGCCCGAGC TAATCCGCGG TCCGCTGGCACA CGGGTTTGG 1080
CGGCCTCATG CTGCCCAAGA GATCGCGCGG CAGCGTTCGA AAACCTCGAC TACATGGGCC 1140
TCCCTGTTGCT GTGCGCGGCC CTGGCGACCT TCCGTCTTGG GGTGTCATCT AGCCCCGCCG 1200
GTGGGAGCATG GGGCGATCGC CAGGCTTGTGA TACCGCGCAT CACGCCCCTG GCGTTGATCG 1260
CGGCATTCTG CGCAGATCGG TGGTACGGCA CAGAAACATCC GCTCATAGAC ATGGGCTTGT 1320
TCCAGAACCAG AGGGTCCGCG CAGGCAACA TAACGATGAC GGTGCTCTCC CGGGCTGCTG 1380
TTGCGCTTCTT CTTGCTGCTC CGAGCTACC-"TCCAGCAAGT GTTGGCACCAA TCAAGGATGC 1440
AATCGGGGGT GCATATCATC CCACAGGGCC TCGGTGCCAT GCTGGCGATG CGCATGCAGG 1500
GAGCCATGAT GCAGCGAGCG GAAGCGGCCA AGATCTGCTG GGTGGGATAC ATGGCTGATCG 1560
CTGGGGGTTT GGCGACTCTC GCCCTGGGTTG TCGGCGCGCA AGCGGACTAC TTACCACATC 1620
TGCCGACCCG GCTGGCAATCT ATGGGCAATGG GCATGGGCTG CTCCATGATG CCACTGTCCG
GGCAGGCAGT GCAGACCCCTG GCCCCCATAC TGATCGCTCG CGGTTGCGACG CGTATCCAGG
TCAACCAGCA GGTGGGCGGT TCGATAGGGA CGGCACTGAT GTCGGTGTCTG CTCAACTTACC
AGTTTCAATCA CAGCGAAATC ATCCGCTACTG CAAGAAAGAT CGCACTGACC CCGAGAGATG
CGGCCGCGCG GGGGGCGGCG GTTGACCTCTT CCTCGCTTACCC GGCACGAAACC AACCTCGCG
CCCAACTGCT GCATGACCTT TGCCACGCGCT AGCGCGGTGGT ATTTGCTGATA GCGACCGGCG
TAGTGGTCTC GACGCTGATC CCCCAGCGCAT TCTCGCGGAA ACAACCGGCT AGTCAATGAA
GAGCACCGTGT GTATCCGCGA TGACGTCTGC TT

(2) INFORMATION FOR SEQ ID NO:190:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1923 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:190:

TCACCCCCGA GAAGTCGTTTC GTGACGACC TGGAACATCGA CTGCCTGTCG ATGGTCGAGA
TCGGCGTGC AACCAGGGAC AAGTCGCGGC TCAAGATCCC CGACGGAGGAC CTGCCGCGTC
TCGGTACCGT CCGGACGTTT GTGCCCTACA TCCGAAGCCT CGAGAAAGAA AACCCCGGAG
CGGCTACGGC GTGCGGCGCG AAGATTTGAGT CGGAGAACCC CGATGCGGGCA CGGAGAATCC
GGTGCGTCTT ACCCATCAGC CAAAGCTCGAG ACCCGCTGCTG TCTCTTTCGCA GCTCTACSCCA
GGTGGGCCTG TGCGCCCGTTT CACGCAATGG TTCCCACCAC AGGAAGGGAC CCTCGCGAAA
GGTGACTGAT CGCGGACCA CATAGTCGAT GCCACCGGTG CTGACAAATG CGGCGGCTG
GAGTGGGCAG GGGCCGGATT GGGCATTGTC GTGCAAGGGC AGCGGATCCC GGGCGCGGCC
CGGCGTTGCT GTGTGGTGGG GGGCGCGGAG GGCACGAGC AGAACGACGA TGGCGCGCAT
GAACAGCGCC ACGGCAATCA CGACCAGCGA ATTTCCGACG CATACCTCCT CGTACGGCTG
CGCGCGGTTC GGTGACTCGG TGGCATACG ATGGCGGCGT TTAACGTAAAC AGCTTTCGGC
GGCCCGGCG TCACAAGCC CGAGTTGCTC GCCCGGAAC CCGCGCGGTC TGGCGCGCGG
TCACCCGACG TCACCTGACT ACCATCCGGG TGGCGGCTGG AGTGCAACTC AAACACACTC
AACGGCAACG GTTTCTCAGG TCACCACTGC AACCCTCGAC CGCAAATGCT GTACCTTTTC
GACCGCGGCC AGGTCGCGAG TCACCGACCTT TGGCGCCGCCCA GCTTTCAGCGG TGAAGCCCAAC 900
CAGGGCATCG TAGGTTGCCC CACCCGGTGAC ATCGTGCTCG GCGAGGTTGT CGGTCAAGCC 960
GGCATATGAG CAGCCATCCA GTGCCAGGTA GTGGTGGGAG GTGATGTGCC CCAAGTAGGC 1020
GTGGACGGCA ACAGGCCGCA TACGATGCGG CCGGGGTCAG CCGGTTCAAGA CGGAATAGGT 1080
TTCACCAGCC GGCTGCGCGGA TCAGAATGAC GCCACGGTGTG AGCCGGCGCCA CGGGCGCCCTC 1140
GTGGGCTTTGC TGGCAGGCTCG CGAATCCGCG AACCAGCACG CTGGTGCTCG GTGGGATCAC 1200
CGGCTGGTGCC GATCGAGGCT TTCCGGAAAC ATATTCGTCGG TCAACGGGGG CAGGGGAGCT 1260
TCTGGGCGTG CGACGAGAAC GAAAGCTTCC CGAAGCAGTT CGAACAACGCT CGGGGCAGGG 1320
TCAATCTCGA TGGGCCGACTC GGCTGCTGGTG ATCTCCACCT GTCTGGTCCC GCGCAAGCCA 1380
AGGGCTGTCG GAATCCGGTT GGGAACTCACC AGAGCTCCTG CGAATCGATG GTGGTTGCC 1440
ATGTTAGGAA ATTTACCATG GCAAGTCTCC TACGGGATGTG CTGGGCCGGGA TGTCGGGACG 1500
ATCCGCTAGC GATGGAAGCG ATGGTTTGCG AAGATGCTGGA GGGAGCGGTC GGGGGGCTTG 1560
ATGGTGGTCTG ATCCCCGGTT GACCGGATGC GGGCTGCTGC TCACTGAGAG TGGCGCTGGT 1620
CGGCGAGCTCA CCGGGCTGGA TGTCGACGGT GTGCGCACAC CGTGGGATGC GGCTTGCCGG 1680
CAGGCCTTGT TGCCCATCAG CGATGCCTGC GACACTGGCC TGGACACCCA TCATCCGGAG 1740
GTGGTGTTCTA TCGAACGGGT GTTCTTCAG CTCAACGTGA CCACGGGTAT GGGCCCGCGG 1800
CAGGCCGGGC GGGTGATCCG CCTGGCGGCC GGGCAGCTGG GTGGGACGCT GTGATCTGCA 1860
ACCACACGCS AGGTCAAGGCG GGGCGTCACT GGCAAAGCGTT CGCCAGACAA GGTCGAGGTC 1920
ACC 1923

(2) INFORMATION FOR SEQ ID NO:191:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1055 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:191:

CTGGCGTGCC AGTGTCACCG CCCGATATGC GTGGGACATT AACCTGCGCG CACCCCGGCGA 60
CCCGTCGCGA CCCAATCGGG ACCACCGGGT CGTCAATTG CCGAAGGCTCG CCAAGTGCGT 120
GCCAATGTTG GTGCTGGGTT TGGTGAAGC CAAGGCTGGCC TACGGGCTGC CCTACCCCGA 180
AACGAACCGA GTGAGGTAAT CGGTTCCCCG GCGGCGGATT CCCAGCGCGA TCTGGCTAGCC 240
GGGGATGGTT CAGACGTAAC GGTGCGCTAG GTGGAACCC CGCACCAGGC CGCTGGAGCG 300
GCTCATTGCA GGAAATTAG AAAACCCGGG ATATTGTCG CGGATTGCTA TACGATGCTG 360
AGTGCTTGGT GGTGCGCTTT TACCGATTTG TGTTGCGAAGC GTTGAGACCC TGGCCTGGAA 420
GGGGACCAAG TGTTTTGCC CTCTGGTCCG CTCCTGGCGG CGGACCGGGT GGTGCCGAAA 480
CGGGCGTGTG CGGGAAATGT CGGGCAGTTTG TGCGCTCCGC TCAGCTGGGG AGTGGCTGTTG 540
CCACCGATAG ATTATGACCA CTGCTGCCCT CGCGCATGATG AGCTGACGATTC TGGTCAAGTG 600
CAGCGGCGGCG AAGGGACCGA CGCGAGGAGC GCACGATGAA ACGAGTGGGA TGAGTGCGCAG 660
GCGTGGAAAG AGTGGGTTGGC GGAAGAACGCT GAACACCGGC TGGAGGTTG ACCGAGTACG 720
AGCAGCGTGA CCCGCGATCC TCCGCGCCCG GCTAGCGGAG GGGCCCGCCAG ACTGCTCTTA 780
TTTGACCAGT GATCGCCGCT CTGGTGTGGC CGCGCGCCCG CGTCGAACAG ATGCACAGTG 840
CATGACAAAGT TACAGGTATT AGGCTCAGGT TCAACAAGGA GACAGCAGAC ATGCGAACAC 900
GTTTTATGC GGAATCAGGAC GGATGGCGGC ACAATGGCGCC CCGTTTGGAG GTGCACGCCG 960
AGACGGTGGA GGAAGGGGCT CGCGCATGAC TGGCGAGCCGC GCAAAACATTC TGGGNGCCGG 1020
GCTGGAGTGGG CATGGCGAG GCGACCTCGC TAGAC 1055

(2) INFORMATION FOR SEQ ID NO:192:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 359 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:192:
CCGCCCTCCTT GTTGGCATATC TCCCCGGCGG CCACCTCGAC CGCAGCTGCC GTGGCGGTGTTG 60
TCGGGGCTGTA CCACCGGAT CGCGGAACCA TCCGAGATCA CTGGCAATAT ATGCCAATCG 120
CGCAGCTGGT CACCCAGCCT CGGCCGCGTG TGGCAGACGG CCTGACATCT CCTGGTATAG 180
CGTCGGCGCC CCAGCGGGAG GAAGTGAGTA TACTGGCCCC CCACCTGGTT ACCGGGAGCG 240
GAGAAGTTCG GGCTGAGGTT CGCCGATGCT GCAGCGGAGGT ATGTGACCCG GAAAAACAGA 300
TCCCCGCGAG GGTGCTGGGG CGGCACCGCC ACGACAACAG CGACGCCGGG ATAGGTCAG 359

(2) INFORMATION FOR SEQ ID NO:193:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:193:

AACGGGCCCC TGGGCACCGC TCTCTAAGG GCTCTCGTTG GTGCGATGAA GTGCTGGAA
GATGCATCTT GCCGAGATTCC CGCCAGACCA AAACAGCCGC TAGTCTAGTT CGAGTGGCC
CGCAAAGTTC CTGGAATAAC TCCGTACCCG GAGGGCCAAA CGGGCTCTCC TCGCTAAGC
TGCGCGAAAC ACTTGAAGTTC CGGGACTCC TTGACGGTCCA GACCGATTGC TCGAGTGGC
TGATCGGTTC GCGGCGCTGG CGCGAATCGG CGCGCGAGCG GGGTGATGTC AACCCAGTG
GTCGCGCTGGA AGAGGGCTCT CACGAGCTGT CTCCGATCGA GGACTTCTCC

(2) INFORMATION FOR SEQ ID NO:194:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 679 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:194:

Glu Gln Pro Lys Gly Pro Phe Gly Glu Val Ile Glu Ala Phe Ala Asp
 1    5    10
Gly Leu Ala Gly Lys Gly Lys Gln Ile Asn Thr Thr Leu Asn Ser Leu
 20   25
Ser Gln Ala Leu Asn Ala Leu Asn Glu Gly Arg Gly Asp Phe Phe Ala
 35   40
Val Val Arg Ser Leu Ala Leu Phe Val Asn Ala Leu His Gln Asp Asp
 50   55
Gln Gln Phe Val Ala Leu Asn Lys Asn Leu Ala Glu Phe Thr Asp Arg
 65   70
Leu Thr His Ser Asp Ala Asp Leu Ser Asn Ala Ile Gln Gln Phe Asp
 85   90
Ser Leu Leu Ala Val Ala Arg Pro Phe Phe Ala Lys Asn Arg Glu Val
100  105
Leu Thr His Asp Val Asn Leu Ala Thr Val Thr Thr Leu Leu
115  120  125
Gln Pro Asp Pro Leu Asp Gly Leu Glu Thr Val Leu His Ile Phe Pro
130  135  140
Thr Leu Ala Ala Asn Ile Asn Gln Leu Tyr His Pro Thr His Gly Gly  
145 150 155 160

Val Val Ser Leu Ser Ala Phe Thr Asn Phe Ala Asn Pro Met Glu Phe  
165 170 175

Ile Cys Ser Ser Ile Gln Ala Gly Ser Arg Leu Gly Tyr Gln Glu Ser  
180 185 190

 Ala Glu Leu Cys Ala Gln Tyr Leu Ala Pro Val Leu Asp Ala Ile Lys  
195 200 205

Phe Asn Tyr Phe Pro Phe Gly Leu Asn Val Ala Ser Thr Ala Ser Thr  
210 215 220

Leu Pro Lys Glu Ile Ala Tyr Ser Glu Pro Arg Leu Gln Pro Pro Asn  
225 230 235 240

Gly Tyr Lys Asp Thr Thr Val Pro Gly Ile Trp Val Pro Asp Thr Pro  
245 250 255

Leu Ser His Arg Asn Thr Gln Pro Gly Trp Val Val Ala Pro Gly Met  
260 265 270

Gln Gly Val Gln Val Gly Pro Ile Thr Gln Gly Leu Leu Thr Pro Glu  
275 280 285

Ser Leu Ala Glu Leu Met Gly Gly Pro Asp Ile Ala Pro Pro Ser Ser  
290 295 300

Gly Leu Gln Thr Pro Pro Gly Pro Pro Asn Ala Tyr Asp Glu Tyr Pro  
305 310 315 320

Val Leu Pro Pro Ile Gly Leu Gln Ala Pro Gln Val Pro Ile Pro  
325 330 335

Pro Pro Pro Gly Pro Asp Val Ile Pro Gly Pro Val Pro Pro Val Leu  
340 345 350

 Ala Ala Ile Val Phe Pro Arg Asp Arg Pro Ala Ala Ser Glu Asn Phe  
355 360 365

Asp Tyr Met Gly Leu Leu Leu Ser Pro Gly Leu Ala Thr Phe Leu  
370 375 380

Phe Gly Val Ser Ser Pro Ala Arg Gly Thr Met Ala Asp Arg His  
385 390 395 400

Val Leu Ile Pro Ala Ile Thr Gly Leu Ala Leu Ile Ala Ala Phe Val  
405 410 415

 Ala His Ser Trp Tyr Arg Thr Glu His Pro Leu Ile Asp Met Arg Leu  
420 425 430

Phe Gln Asn Arg Ala Val Ala Gln Ala Asn Met Thr Met Thr Val Leu
Ser Leu Gly Leu Phe Gly Ser Phe Leu Leu Leu Pro Ser Tyr Leu Gln
435 440 445
Gln Val Leu His Gln Ser Pro Met Gln Ser Gly Val His Ile Ile Pro
450 455 460
Gln Gly Leu Gly Ala Met Leu Ala Met Pro Ile Ala Gly Ala Met Met
465 470 475 480
Asp Arg Arg Gly Pro Ala Lys Ile Val Leu Val Gly Ile Met Leu Ile
485 490 495
Ala Ala Gly Leu Gly Thr Phe Ala Phe Gly Val Ala Arg Gln Ala Asp
500 505 510
Tyr Leu Pro Ile Leu Pro Thr Gly Leu Ala Ile Met Gly Met Gly Met
515 520 525
Gly Cys Ser Met Met Pro Leu Ser Gly Ala Ala Val Gln Thr Leu Ala
530 535 540
Pro His Gln Ile Ala Arg Gly Ser Thr Leu Ile Ser Val Asn Gln Gln
545 550 555 560
Val Gly Gly Ser Ile Gly Thr Ala Leu Met Ser Val Leu Leu Thr Tyr
565 570 575 580 585
Gln Phe Asn His Ser Glu Ile Ile Ala Thr Ala Lys Lys Val Ala Leu
590 595 600 605
Thr Pro Glu Ser Gly Ala Gly Arg Gly Ala Ala Val Asp Pro Ser Ser
610 615 620
Leu Pro Arg Glu Thr Asn Phe Ala Ala Gln Leu Leu His Asp Leu Ser
625 630 635 640
His Ala Tyr Ala Val Val Phe Val Ile Ala Thr Ala Leu Val Val Ser
645 650 655
Thr Leu Ile Pro Ala Ala Phe Leu Pro Lys Gln Gln Ala Ser His Arg
660 665 670
Arg Ala Pro Leu Leu Leu Ser Ala
675

(2) INFORMATION FOR SEQ ID NO:195:

(1) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:195:

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

(2) INFORMATION FOR SEQ ID NO:196:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:196:

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

(2) INFORMATION FOR SEQ ID NO:197:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 119 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:197:

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

(2) INFORMATION FOR SEQ ID NO:198:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 116 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:198:

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

(2) INFORMATION FOR SEQ ID NO:199:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 811 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:199:

TGCTACGCAG CAATCGCTTT GTGTGACAGAT GTGGATGCCG GCGTCGCTGC TGGCGATGCC 60
GTGAAAAGCCG CCAGGTGTTT CGCGGATCTC GGGGAGAACAC TCGAATGCTCT CAAAAGGCTG 120
GTGCGGCGCG CCATCGATCG GTGCGCGCGA AGCGACACTG GCACACCTGC TCGACACAC 180
GCCGTTTGTTT CGTTGCGGTG TGAGGTGGCT GTGGACCGGC GCGGCGGGCTG 240
TCATCGGGTG CGCGTGTTGAT GCCGGCGTAC GGGCGCGGTC GACGACGCTG GTGCGGTGCTG 300
ACGCCGCTGCT GCACCCCGCGC CACGCCAACCA ACCCGGTTCT GCCAGCGCAG TGCCAGCGGC 360
TCGACGTGCC CGACGCCGAC GCCGCTGGGC CGTTGTTGCT GGGTGTGCGAT CGGGTGCTGC 420
ACCGGCCGAT GTGTTGGGCT GCCCGCGCTAC CGCCCGCACA GCACCGCGCC TATGCACGCC 480
ACAACGATTT CGCCACCGAC GTGCAGGCGG CGCGGATGTT CGCGGCGGCG GTCCGGCGTTT 540
TGGTGCTGCG GGGGTTACG GCAAGTGGTG TTGTGCTGATG CGAAGCGGCT CTAGCGAGTG 600
GACGGTGTCGA CGCGTGCGGC CGCCCGCGCA CGACACCTGGT CAAATGGGTC TTCGACGACC 660
GGGCGGGGCC GTGCGGGCGA CGCAGCTCTT GCCGAGTTGGT CGACGAAGAT GCCCGGTTGC 720
GCCGCGCAG CCGTACCGGC GCACCGAGAG CGCCCGAGG CACTACCGGC TTCCGGTGCT 780
GGAAACGGAAT GCGGCTGTTG CGTGGCGGCCGTTG 811

(2) INFORMATION FOR SEQ ID NO:200:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 966 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

```
GTCCGGCAGT GTGGCCAGGC ATGACTTTCG GCAACACCGG GTAGTATCG 60
GACTTTGTTG TCCCGTGGGC GGGATAGACG ACCGTGCAGC GTGGTCAGC 120
GTCGGCGGCC CGAACCCATG CTTCAACGT AGCCTGTCCG TCACACAGGT 180
ACGTACCGGTCAA AATATCGC GTGGAATTTCC GCGTGAGCTG TCCGCTCGC 240
GCACACGGCA TTACATGGCA GCAATTGGGA CGGGTTGCAG CAGCTCTTG 300
TGCGGTCAAA GGTGACACG GCACCGCAG ACCACGCAAT CAGTTTCAAC 360
CCAGGACAA CGGAGCCATC CGGACAACGC GGTGCTGCCA GGGTGTGSC 420
ACTACCCCGCA CTTGCGGGGG CCCTACGAGT CGGCAGGCGC CCAATATGGG 480
ACGAGACCAT CGCCATCTTG CCTGAGAAC GCCCGCGGGA GATGAGGTCG 540
TCAGGGGGCG CGACGTCGAC CGAGTGAATG TCGACCGCAG CAATCTTTTG 600
CGAGTGCAGC GAAATCGATG GTGGCTGCTGG CTAGAGCAGT CGGCGCAGG 660
CGTAATGCTGC TCGGCCCCACG GCGTGAATTC GCAAGAATAA TAAAGCCGC 720
TCGGGGTCTG CACCACTACG CTGAGATCG AGAACATGGC GGTGAGGCT 780
GCAGGGCTAT AGAACGCCGT CGCCAGTATT ATCAAACGCA AATACGCTT 840
TCGGCGGTCG TCACCGGATG CGGCTTTTTT GCCACGCTCC ACCGGTGGC 900
AAGTTGGCAGT GCAATCCACG GCTTCCCCGG AACCAGTTCC GGGGTGGC 960
CATCCT
```

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 2367 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

```
CCGCACCCGG GCAATACCG CGACCGCCAC CGTTACCGCC GTGGCGCCGT 60
TGCGGCGCGTCG CCGCCTGCGCC CGCCCGATGG ATTTCTCATC GCACAAATGA 120
CACCAGGGCC GCGTTGCGGG CGTCACCGCC CAGCCCCGGC GACTCCACCG GCACCCAAG 180
```
CTCCGCCGCT GCCACCGTTG CGGCCGTTGTC CGATCAACAT GCCGCTGGCG CGACCCTTTCG
CACCCACGCC ACCGGCTCCG CCCACCCCGC CGACACCAAG CGAGCTGGCG CGGGAGCACC
CATCACACC TACGCCACCG ACCGCCAGAGA CACCAGGAC CGGGTCTTCCG TGAAACGCTG
CGGTGCCACC ACCGGCGCGG TTACCGCAA CCCCACCAGG AACGCGCGCG CGCCCATCCC
CGCCGCACCC GCCGCTTGCCG CGTTGCAGCG CGTGCAGCAGA CAAACCCCGG CGGCGCGCCG
CGTGTGGCGGC CGCGCGCGCG GTCCCCCGCG CGCCGCCGAC GCACAGGCGG CTGCCGCCCCG
TGCCGCACTC ACCACCTTGG CGCAGCCAGCG CATCGGTGTC TGCCCTGGCG TGGGCGGCTG
CAAAACCTGGC GATGGCCAGCG TTGGCCGGCG TTCCCGCGGG CCCCACCGTG GCCGCGCAGC
CACGGATACC ACCGGCCAGA CGGGCCGCCC CGTGGCGCAG ATCGACGAAAT AGCAACCCGC
CGGCCCAAC ACCGCGCGCA CGCAGGCGTT CACCACCGCTG GCCCCGGGAG GCCGCACCTG
CAGCCCCGTT ACCACCGAGA CGCCGCGTAAC CACCGGTAGA GGTGCAGAGT GCCGATGCTG
CGAAAGCCGC GCCGCGGGCG CGCCGCCGTC ACCCCCGAGT GCCCGGCGACT GCCGCGCGCT
ACCCGTTGCC ACCATCACCG CCAAGGGCGC TGCGAATGTC GCCCTGGCGC ACTCCCGCGT
CGCCGCCGTT GCCGCGCGCG CGACCACGCG CGGCCGGTAC GCCGTCAACCA CGGCGACCCG
CGTGGCGCTTT GCCGAGGCGT GCCGCTGGCG TGGCACCAGTC GCCCGCGGTG ACCACCGTCG
GGGCCGGGCC AGTCGAGATGG CCCGGCGTGC CGGCGGCGCC GCCGGGGTTGA TCACCGATGC
CGGACACATC TCACCGGGTC TCCCCGGTG TGGCCCGCGG GCCGCGGGTG GGGATGACCC
GTTTGGCCCG GCCGGAGCGCG GCCGCGCGCG TACCACCGGC GCCGCGATCG GCCGACACCG
CGCGCGCGCC GCCGTGAGCG CGCCGCGCGG GCCACGCTGC GCCACGCTG GCCGCGCGCG
CACCCCGGTGG GCCGGCGGCGT CCCCCAACC ACCCCCGGGT CCCCCCGSAC CGCCGGCGCG
CGGCGCGGCT ACCGGCGCCGG CGTGGCGGAT CACCGCGCGC GCCGCCCGCG GCCGCCTCCG
TGCCGCCGGT TGACGCCAAC CGCCAGCGCC CGCGGTGCGC ACCCGCGCCA AACAGCAACC
CGCCGCCGCG GCCGAGGCGTG CGGGGCGTGC TCCCGTGCGC GCCGTTTCCG ATCAAGGGGC
GCCCCAAAAG GCCCTGGGCG GCCGAACATTCA CGGCCACCGAG CAGACTCCGC TCAACAGCGG
CTTCAGTGGT GGCAATCCGA CCCCGGCGCG CAGTCACAGGC CTGGCACAAC TGGCTGCTGAA
AGGCTGCCAC CTGTACGGTG AGGGCCGAGT ACTGCCGAGC ATGGGGCGCG AACAACCCCG
CAATGGCCGC CAGACTTCCA TGCCAGCGCG CAGCCACGAC TCCCGCGCTG GCCGATCGCG
CGCGCGGATT AGGGCGGCTC ACCTGGCGAA CAAATAGTCGA TAAATTCCAA GCCGCGAGGTG
CCAGCAGCTG CGCGTGCGG ATCACAAGG ACACCTCGCA CCTCCGCGA CCCCATATCG
CGCACCAGTG TCCCCAGCGG CCACGTGACC TGGGTTGATT GCTGGCGGCG CCGTGAATCG
GCCGGCAGCG CCCCCCTTC TATCCCCC CCATGCGACG TTGGTTGCGG AGTTGAGAC
GGAGGAGACGG CCGAGCTTG GCTGTAGAAGT GGGTCAGTAG GGAATCAGAC GTCCGGCGCG
AGATGAAATAG GCGGACGCCT ATCTCCTTGT TGCTGAGTTCC CTCCACGACC AGTAGAGCCA
CCTCAAGGTCG TCTCGTGTTAC AACGGCGCCCC AGCCACTTGT CGGGGCTTTC CGTGCAAGCG
GCCCTCAGTG CGGTAGCCG ATCGCTCAT CGATCGATAA CGCGATTCTC TGCGCAGGA
CATCGTCGAA CTGCTGTTCA CCCATGGATT TTGGAAGGGG GGCTAGCCGAC GAGTTACAGC
CCGCTGGGTA GATCCCCGAAG CCGACCG

(2) INFORMATION FOR SEQ ID NO:202:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 376 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:202:

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

(2) INFORMATION FOR SEQ ID NO:203:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2852 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:203:
GGCCAAAACG CCCCGGCCAT CGCGGCCACC GAGGCCGCTC ACGACCAGAT GTGGGCCCAG 60
GACGTTGGCC GAGATTGGTG TACGACATGG CGGGCTTCCG CGCCCGTCCT CGCCTTGGAC 120
CCCTTCGGCC AGGCCGCTCC GAGCCGCGCC CGCCGCGGCTG CGTGGGTGAC CGCGGGCCGC 180
GTCAGGTGAC CGCCGCGGTT CTCCGCGAAC CTGGGCTTTGG CGAACGTCCG CGAGGGCAAC 240
GTCGCGAAGC GTGAAGCTCG GAACCTCAAT CTGGGCTCGG CCAACATGCC CAACGGCAAC 300
ATCCGAGCGC GCAACATCAG CAGCCTCAAC ATCCGGTGGG GCAACGTGGG TCTGCGGTTG 360
ACCGCAAGGC TGAAACACAT CGTTTGCACC AACACCGGCC GAAACCAACAT CCGTGTGGGC 420
AACACCGCA GCAAACACAT CGGGTCCGGA AATACGGGAG ACGCGAAAGC AGGTATCCGG 480
CTCAGGGGTC GCAGTTTTGG GCAGTTTCCG GCCGCTGAAC CGCAGCCGGCA CAACATCGGT 540
CTGGTTCAACT CGGGCCAGGG AAAGGCGTCCG ATCCGGCAACT CGGGTACCGG GAACCTGGGC 600
ATTGCGAACG CGGGCAACGT CTACACACCG GGTATCTGGCA ACTCCGGCGA CGCGCAACACG 660
GGCTCTTCTCA ACTCGGGAAT GCGCGAACC CCAGCTCCTCC CAAGCGGGCA ACAGCGGCAA CTACAACACC 720
GGTAGCCTCA ACCCGGCCAA CAGCAATCCC CGCCGGCTTC GCATGGGCCCA ACATCGGGCA GTACAACAG 780
GGCTACCTCG AGACGCGGCA CTACACACCG GGTATTGGGCA ACTCCGGGGA TGCTCAACACC 840
GGCCGCTTCA TTACTGCGAA CTTCAACACCG GTTCTTTGGT GGCGCGGCGA CCAACAGGCC 900
CTGATTGTTG GAGCCCCCGG CTGTTTTCAAC TGACCGATGG CGCCCGTCCT CGGATTCTTC 960
AACCGGCTG CGGCTAGCGG CTGCCGCTCT GCATGATCTCG GGGCAACAAA TGCTGGCTTC 1020
TTCAACTTCT CGTCCGGGCG CACGCCCGAG TCCGCGTACG CGTGCGCTTGG CAAACGGGGG 1080
TGGGCGGCTG TCAACTGGGA CAAACCTGTA TGAGGGTTGG TCAACGCTGG CCTGGGCTGC 1140
ACACAGACGC CGGCCTTGGT CCCTGGGCTTC TCCAAACACCG GAGCGAACAGT GCACGGGATT 1200
TCCGCGGCGC CACCGCTTCT CCATGCGCCGG GCAGCAGTCA CCGCGCGGCGA CAGCGGCAAC 1260
GCGCAAGCCA ACATCGGCAG CACATACATT CTCGCGCCCGA GAAACGCGGC TGACTTCGCA 1320
ATCGTTGGCA CGCGGCAACGT CGCCGCACAA AACATCTGGG CGACGCGGCCA CGTGGCCAGC 1380
TTCAATATGG GCAGTGGAAAT CAGCGGAAGT TCCAAATGCG GTTCGGGAAG CTTGGGAAAC 1440
TACAAATAGC GATCCGGAAA CTCGGGGAAC TACACATGAG GTTGTGGGAA AGGCCGGGAC 1500
TACACAGTCG GCTCCGGGAA CGCCGCGCAC TCCAAACAGA GCTTGGCACA CACCGCAAGC 1560
AACAAACATCG GGTCCGGCAAC CACCGCAAC CGACACCTCG GCATCGGGCT GTCCGGCCAGC 1620
ACCCAGCAGG GCTTCAATAT TGCTAGCGGC TGGAACTCGG GCACCAGCGA CACCGGGCTG 1680
TTCAATTCCG GCACCAATAA CTGTTGGCATC TTTCAACGCAG GCACGGGAAA CGTGGGGCATC 1740
GCAAACCTGG GCACCGGGAA CTGGGGTATC GGGAGCCCGG GTACCCGACAA TACCCGACTC 1800
CTCAATGCTG CGAGCTTCAA CAGGGGCACC TTTCAACGCGG GCCGACTTCAA CACGGGGCTTC 1860
TACCAACGCG GCAGCTTCAA CACCGGGCGC TTTCAACGCGG GTAACCCCAA CACCGGGCAC 1920
TTCAACGTGG GTGACACCAA TACCCGGCGC TATAACCCGG GTGACACCAA CACCGGGCTTC 1980
TTCAATCCCG GACACGTCAA TACCCGGCGC TTTGACACCG GCCGACTTCAA CACGGGGCTTC 2040
TTGGTGGGCGG GCGATAACCA GGGCCAAGATT GCCATCGATG TCTCGGGTAC CAATCCATTTC 2100
ATCCCCATAA ACGACGAGAT GTGTCATTGAC GTACCAACAG TAATGACCTT CGGCGCGCAA 2160
ATGATCAGG TTACCAGGGC CTGGACGGTT TTCCCCCAAA CCTTCTATCT GACCGGGTTTG 2220
TTCTTTCTCG GCCCGGTCAA TCTCAGGGCA TCCAGCCTGA CGTGGGCAG GATCAACCTTC 2280
ACCATCGGGG GACCCGCGGT GACCCGCTCC ATCAAGAATTG TCGTGGCTCT GGAAGACGCC 2340
AGGATTACCCT TCTCTAAGAT CGATCCGGCG CGGCGCATCG GAAATTCGAC CACCAACCCCC 2400
TCGTGGGGGT TCTTTAAGTCG GGGACCGGTT GGCACAGATG GTTCCCAAAA GGTGGGCGGC 2460
GGGAGGTCAG GGGTCCTGGA CAGTTGTTTG AGCCAGCAGC TAGGGAAATTC GGGTTTCCAG 2520
AACCTCGGCT CGTGCGAGTC AGGCTGGGCG AACCTGGGCA ACTCAGGATGC GGCTTTTTTC 2580
AACACCAAGTA CGGTGAACTCT CTCCACGGCG GCCAATGTCT CGGCGCTGAA CACCATCGGC 2640
GCAAAGCCTGG CGGGGCTGGT CGGGCGTGCC ACCGGGCGAA TTTTCAAGGC GGGCTTGGCC 2700
AACCTGGGGG AGTTGAAACAT CGGCACGGGC TCGTGGCAGA TTTGGCAGA GTTAGTACG 2760
GTTTCAACAA TCATATCCGC GTTTTGGCAG AGTGCGATCG ACAGAATCGAA CGCCGGAGAC 2820
GTAAGCGAAT AAACCGAATG GCGGCGCTGTC AT 2852

(2) INFORMATION FOR SEQ ID NO:204:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 943 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: linear
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:204:

Gly Gln Asn Ala Pro Ala Ile Ala Ala Thr Glu Ala Ala Tyr Asp Gln
1      5     10     15
Met Trp Ala Gln Asp Val Ala Ala Met Phe Gly Tyr His Ala Gly Ala
20 25 30
Ser Ala Ala Val Ser Ala Leu Thr Pro Phe Gly Gln Ala Leu Pro Thr
35 40 45
Val Ala Gly Gly Ala Leu Val Ser Ala Ala Ala Ala Ala Gln Val Thr
50 55 60
Thr Arg Val Phe Arg Asn Leu Gly Leu Ala Asn Val Arg Glu Gly Asn
65 70 75 80
Val Arg Asn Gly Val Arg Asn Phe Asn Leu Gly Ser Ala Asn Ile
85 90 95
Gly Asn Gly Asn Ile Gly Ser Gly Asn Ile Gly Ser Ser Asn Ile Gly
100 105 110
Phe Gly Asn Val Gly Pro Gly Leu Thr Ala Ala Leu Asn Asn Ile Gly
115 120 125
Phe Gly Asn Thr Gly Ser Asn Asn Ile Gly Phe Gly Asn Thr Gly Ser
130 135 140
Asn Asn Ile Gly Phe Gly Asn Thr Gly Asp Gly Asn Arg Gly Ile Gly
145 150 155 160
Leu Thr Gly Ser Gly Leu Leu Gly Phe Gly Gly Leu Asn Ser Gly Thr
165 170 175
Gly Asn Ile Gly Leu Phe Asn Ser Gly Thr Gly Asn Val Gly Ile Gly
180 185 190
Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Ser Gly Asn Ser Tyr
195 200 205
Asn Thr Gly Phe Gly Asn Ser Gly Asp Ala Asn Thr Gly Phe Phe Asn
210 215 220
Ser Gly Ile Ala Asn Thr Gly Val Gly Asn Ala Gly Asn Tyr Thr
225 230 235 240
Gly Ser Tyr Asn Pro Asn Ser Asn Thr Gly Gly Phe Asn Met Gly
245 250 255
Gln Tyr Asn Thr Gly Tyr Leu Asn Ser Gly Asn Tyr Thr Gly Leu
260 265 270
Ala Asn Ser Gly Asn Val Asn Thr Gly Ala Phe Ile Thr Gly Asn Phe
275 280 285
Asn Asn Gly Phe Leu Trp Arg Gly Asp His Gln Gly Leu Ile Phe Gly
290 295 300
Ser Pro Gly Phe Phe Asn Ser Thr Ser Ala Pro Ser Ser Gly Phe Phe
Asn Ser Gly Ala Gly Ser Ala Ser Gly Phe Leu Asn Ser Gly Ala Asn
305
310
315
320
Asn Ser Gly Phe Phe Asn Ser Ser Ser Gly Ala Ile Gly Asn Ser Gly
325
330
335
340
345
350
Leu Ala Asn Ala Gly Val Leu Val Ser Gly Val Ile Asn Ser Gly Asn
355
360
365
Thr Val Ser Gly Leu Phe Asn Met Ser Leu Val Ala Ile Thr Thr Pro
370
375
380
Ala Leu Ile Ser Gly Phe Phe Asn Thr Gly Ser Asn Met Ser Gly Phe
385
390
395
400
Phe Gly Gly Pro Pro Val Phe Asn Leu Gly Leu Ala Asn Arg Gly Val
405
410
415
Val Asn Ile Leu Gly Asn Ala Asn Ile Gly Asn Tyr Asn Ile Leu Gly
420
425
430
Ser Gly Asn Val Gly Asp Phe Asn Ile Leu Gly Ser Gly Asn Leu Gly
435
440
445
Ser Gln Asn Ile Leu Gly Ser Gly Asn Val Gly Ser Phe Asn Ile Gly
450
455
460
Ser Gly Asn Ile Gly Val Phe Asn Val Gly Ser Gly Ser Leu Gly Asn
465
470
475
480
Tyr Asn Ile Gly Ser Gly Asn Leu Gly Ile Tyr Asn Ile Gly Phe Gly
485
490
495
Asn Val Gly Asp Tyr Asn Val Gly Phe Gly Asn Ala Gly Asp Phe Asn
500
505
510
Gln Gly Phe Ala Asn Thr Gly Asn Asn Ile Gly Phe Ala Asn Thr
515
520
525
Gly Asn Asn Asn Ile Gly Ile Gly Leu Ser Gly Asp Asn Gln Gln Gly
530
535
540
Phe Asn Ile Ala Ser Gly Trp Asn Ser Gly Thr Gly Asn Ser Gly Leu
545
550
555
560
Phe Asn Ser Gly Thr Asn Asn Val Gly Ile Phe Asn Ala Gly Thr Gly
565
570
575
Asn Val Gly Ile Ala Asn Ser Gly Thr Gly Asn Trp Gly Ile Gly Asn
580
585
590
Pro Gly Thr Asp Asn Thr Gly Ile Leu Asn Ala Gly Ser Tyr Asn Thr
595
600
605
Gly Ile Leu Asn Ala Gly Asp Phe Asn Thr Gly Phe Tyr Asn Thr Gly
610 615 620
Ser Tyr Asn Thr Gly Gly Phe Asn Val Gly Asn Thr Asn Thr Gly Asn
625 630 635 640
Phe Asn Val Gly Asp Thr Asn Thr Gly Ser Tyr Asn Pro Gly Asp Thr
645 650 655
Asn Thr Gly Phe Phe Asn Pro Gly Asn Val Asn Thr Gly Ala Phe Asp
660 665 670
Thr Gly Asp Phe Asn Asn Gly Phe Leu Val Ala Gly Asp Asn Gln Gly
675 680 685
Gln Ile Ala Ile Asp Leu Ser Val Thr Thr Pro Phe Ile Pro Ile Asn
690 695 700
Glu Gln Met Val Ile Asp Val His Asn Val Met Thr Phe Gly Gly Asn
705 710 715 720
Met Ile Thr Val Thr Glu Ala Ser Thr Val Phe Pro Gln Thr Phe Tyr
725 730 735
Leu Ser Gly Leu Phe Phe Phe Gly Pro Val Asn Leu Ser Ala Ser Thr
740 745 750
Leu Thr Val Pro Thr Ile Thr Leu Thr Ile Gly Gly Pro Thr Val Thr
755 760 765
Val Pro Ile Ser Ile Val Gly Ala Leu Glu Ser Arg Thr Ile Thr Phe
770 775 780
Leu Lys Ile Asp Pro Ala Pro Gly Ile Gly Asn Ser Thr Thr Asn Pro
785 790 795 800
Ser Ser Gly Phe Phe Asn Ser Gly Thr Gly Gly Thr Ser Gly Phe Gln
805 810 815
Asn Val Gly Gly Ser Gly Val Trp Asn Ser Gly Leu Ser Ser
820 825 830
Ala Ile Gly Asn Ser Gly Phe Gln Asn Leu Gly Ser Leu Gln Ser Gly
835 840 845
Trp Ala Asn Leu Gly Asn Ser Val Ser Gly Phe Phe Asn Thr Ser Thr
850 855 860
Val Asn Leu Ser Thr Pro Ala Asn Val Ser Gly Leu Asn Asn Ile Gly
865 870 875 880
Thr Asn Leu Ser Gly Val Phe Arg Gly Pro Thr Gly Thr Ile Phe Asn
885 890 895
Ala Gly Leu Ala Asn Leu Gly Gln Leu Asn Ile Gly Ser Ala Ser Cys
 900    905    910
Arg Ile Arg His Glu Leu Asp Thr Val Ser Thr Ile Ile Ser Ala Phe
 915    920    925
Cys Gly Ser Ala Ser Asp Glu Ser Asn Pro Gly Ser Val Ser Glu
 930    935    940

(2) INFORMATION FOR SEQ ID NO:205:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 53 base pairs
    (B) TYPE: nucleic acid
    (C) STRA NDEDNESS: single
    (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:205:

GGATCCATAT GGGCCATCAT CATCATCATT ACGTGATCGA CATCATCGGG ACC

(2) INFORMATION FOR SEQ ID NO:206:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 42 base pairs
    (B) TYPE: nucleic acid
    (C) STRA NDEDNESS: single
    (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:206:

CCTGAAATTCA GGCCTGGTT GGCSCGGCCT CATCTTGAAAC GA

(2) INFORMATION FOR SEQ ID NO:207:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 31 base pairs
    (B) TYPE: nucleic acid
    (C) STRA NDEDNESS: single
    (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:207:

GGATCTGTCA GGCTCAGAAC CACCGGCGGG T

(2) INFORMATION FOR SEQ ID NO:208:

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 31 base pairs
    (B) TYPE: nucleic acid
    (C) STRA NDEDNESS: single
    (D) TOPOLOGY: linear

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:208:
CTCTGAATTG AGCGCTGGA AAATCGTCCGGA T

(2) INFORMATION FOR SEQ ID NO:209:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 33 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:209:

GGATCCAGGC CTGAGATGAA GACCGATGCC GCT

(2) INFORMATION FOR SEQ ID NO:210:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:210:

GGATATCTGC AGAATTCAGG TTTAAGGCC ATTTGCGA

(2) INFORMATION FOR SEQ ID NO:211:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:211:

CCGCATGCCA GCCAGTGCC CACAACGGCC

(2) INFORMATION FOR SEQ ID NO:212:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:212:

CTTCATGGAA TTCTCAGGCC GGTAAGTCC GCTGGCGG

(2) INFORMATION FOR SEQ ID NO:213:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7676 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:213:

TGCGCAATGG GACGCCGCTT GTAGCCGCAG ATTAAACGCG CGGGTGTTGG TGTTTACGCG 60
CAGCTGACC GCTACACTTG CCAGCCGCTT AGCCGGCCGCT CTTTCTGCTTCTTCTCTTC 120
CTTCTCGCC ACCTGCGGCC CGTTTCCCCG TCAAATCTTTA AATCGGGGGGC TCCCTTTTGG 180
GGTAGATTT ATGCTTTCGCC GCACCTCGGA CCCCCAAAAA CTGTATTAGG GTGATTGTTC 240
ACGTAGTGGG CCATCGCCAT GATAGAGGTT TTGGCGCCCC TGGACGGTTG AGTCCAGCTT 300
CTTAAATAGTT GGACTCTTGG TCCAAAGTGG AACAACACTC AACCCTATTCTGG GCTTCATTTC 360
TTTTGATTTTT TAAAGGGATTT TGCCGATTTGC GCCTATTGG TTAABAAAGTT AGCTGATTTA 420
ACAAATAATT AAGCCGAATT TTAAACAAATT ATTAAGCCTTT ACAATTTCAG GTGGCACTTT 480
TCGGGAAAT GTGGCCGGAA CCCCCATTGTT TTTATTTTTC TAAATACATT CAAATAGTA 540
TCGCCCATGT AATTAATTGT TAAAGAAAAT CACTGAGCAT CAAATGAAAC TGCAATTTAT 600
TCATATCAGG ATTATCAAATA CCAATTGTTT GAAAAAGCCTT TTTCTGTAAT GAAGGAGAAA 660
ACTCCCGGAG GCAGTTCCAT AGGATGGCAA GATCCTGTTA TGGTCTGCGG ATCCCGACTC 720
GTCCAAACTCT AATACAACTG ATTAATTTCC CCGTAGTCAA AATAAGGTTA TCAAGTGAG 780
AATACACATG AGTGAGCAGC CTGACCCGCCG AGAAATGGCAA AAGTTTAATG CTATTCTTTTCC 840
AGACTTTCAG AACGCGCCAG CCATTCAGCT CGTCATCAA AATCAGTGC ACAAACAAAC 900
CGTTATTCTG TGGATGTGCG GCCTGAGCGCA GACGAATATC GCGATCGCTG TTAAGAGGAC 960
AATACAAAC AGGAATCGAA TGCAACGCCC GCAGGAAACCG TGGCAGCCGA TCAACAAATAT 1020
TTTGACCTGA ATCAGGATAT TCTTCTATAA CCTGGAATGC TGTTTTCCCG GGGATCGCAG 1080
TGTTGAGTAA CCAGATCATC TCGAGGTATC GCAAATAATG TTTGATGGTC GGAAGAGGCA 1140
TAAATTCCGT CAGCCAGTTT AGCTCTGACCA CTCGATCTTG AACATCAATG GCAACGCTAC 1200
CTTTGCAGTG TTTACAGAAC AAATCTGCGC CATCGGGCCT CCAATACAAAT CAGATGATTG 1260
TCGGACCTGA TGGCGCGACA TTATGGCGAG CCCATTATAA CCCATATAAA TCAGGATCCA 1320
TTTTGGAAATT TAATCGGCGGC CGAGGACAGC ACGTTTCCGG TTGAATATGG ACTTAACAC 1380
CCCTCTATT ACTGGTTATG TAAGCAGACA GCTTTTATTGT TCAGGACCAA AATCCCTTAA 1440
CGTGAACCTTT CGTCCACAGT AGCGTCAGAC CCGTACAGAA AGATCAAAGG ATCTTCTTGA 1500
GATCCTTTTT TTCTGCGCGT AATCTGTGTC TTGGCAACAA AAAAAACCACC GCTACCAGCG 1560
GTGTTTTTGTT TCCCGGATCA AGAGCTACCA ACTCTTTTTTC CGAAGGTAAAC TGGATCTGAC 1620
AGAGCGGAGA TACCAATAAC TGTCCTTCTA GTGTAACGGTT AGTTAGGCCA CCACCTTAAG 1680
AACTCTGTAG CACCCGCTCTC ATACCTCGCT GTGCTAATCCC TGTTACAGCT GGTCTGGCTCC 1740
AGTGCGGATA AGTGCTGTCT TACGGGTTTG GACCTAAGAC GATAGTTACC GGATATAAGCG 1800
CAGCGGTCCG GCTGAACCGGG GGGTTCGCTGA ACACAGCCCA GCTTGAGAAGC AAGCAAGCTAC 1860
ACCGAATCAGA GATACCTACA GCGTGAGCTA TGAAAGAAGCG CCAGGGCTTC CCAAGGGAGA 1920
AAGCGGGACA CGTATCCGCT AAGGCGCCAGGG TGTCAAACAG GAGACCGCAG AAGGGAGCTT 1980
CCAGGGGAAA ACAGCCTGGTA TCTTTTATAGT CCTGGGCTGTT TTCCGCACCT CTTGACTTGA 2040
GGTGGATTTC TTGTGATGCTC GTCAAGGCTGC CGAAGGCTAT CGAAGAAACCC CAGCAAGCCG 2100
GCCTTTTACT GGGCTCTGGC CTTTTGCTGC ACGATGCTTC TCTGTGGCTTA 2160
TCCCCGTGAT CTGTGGGATA CCGTATACCGC GCTTTTAGGT GAGCTGATAC CGCTCCGGGC 2220
AGCCGAACGA CGACCGCGAG CGATCATAGT AGCGGAGAAG GGAAGAAGCG CCTGATGGCG 2280
TATTCTCTTC TTAACGATCG GTGCGCTATT TCACACCGCA TATAAGTGTC ACTCTCAGTA 2340
CAATCATGCTG TAATGCAGCA TAGTGAAGCC AGTATAACCT CGCATATCGC TACGTGACTG 2400
GGTCATGGCT GGCGCCGCGA ACCGCGCAAC ACCGCGCTGAC GCGCCCTGAC GGCTTTGCT 2460
GCTCCGGACA TCCGGTTACA GACAACTGCT GACGGCTCTCC GGGACTGCGA TGTTGACAG 2520
GTTTTCACGG TCATCCGCGA AACCGCGGAG CCCAGCTGGC TAAAGCTCAT CAGCGTGGTC 2580
GTGAAGCGAT TCACAGATGT GTGCGCTGTTC ATTCGGCCTCC AGCTGGTGAG GTTTGCTTCG 2640
AAGCGTTTAAG GTCTGGCTTC TGATAAGGCG GCCCATGTTA AGGCGCGTTT TTTCTGTTTT 2700
GGTCACCTAG GCTCCGCTGGT AAGGGGGATT TCTGGTTCTG GGCTTAATGA TACCGATGAA 2760
ACGAGAGAGG ATGCTCAGCA TACGGTTTTAC TGATGATGAA CATGCCCGGT TACTGGAAAG 2820
TTGTGAGGTT AAACAATCGG CGGTATGGAT GCAGCCCAGG CAGGAAAAAA TCACTCGAGG 2880
TCAATTGCCAG CGTTTGGTTA ATACAGATGT AGTGTTCCTCA CAGGGTAGGC AGCAAGCATCC 2940
TGGCATGAG ATTCGGAACA TAAAGGGTA GGGCGCTGAC TTCCGCGTTT CCAGACTTTA 3000
CGAAACAGGG AACGCGAGGA CCTACGCTGT GGGTTCGCTAG GTGGCGACAG GGGTGCAGCA 3060
GCAGTGCTTT CACGTTGGCT CGGTATGCGG TGATTTTTCA TGCTAACCAG TAAGGCAACC 3120
CGCCGAGGCT AGCCGGGTCC TCAACGACAG GAGCAGCTGC ATGCGCAACC GTGGGGCCGC 3180
CATGCGGGCG ATAAATGGCT GTTCTCGCC GAACAAGTTTG GTGCGGGGAC CAGTGACGAA 3240
GGGTTGAGCG AGGGCGTGCA AGATTTCCGAA TAACGCAAGC GCACCGCGGAG TCATGTCGSC 3300
GCTCCACCGA AAGCGGTTCCT GAAGGGAAAT GACCACGAGCC CTTGCGCAGCG CTTGTCGCTC 3360
GACTGACATG ATAAAGAAGA CAGTCTACAAG TGCCCGGACG AGATGTGATG CCGCGGCGCA 3420
CGGAAGGGGC CATGACTGGST TGAAGGCTCT CAAGGGCATC GGTGAGATCG CCGTGTGCCTA 3480
AGAGGTGACG TAATTTACAT TAATTGGGTTT GCCTCCACTG CCGGCTTTCC AGTCCGGGAAA 3540
CCTGTCTGGC CAGTCGCAAT AATGAACTGG CGAACCGCGC GGGAGAGGCG GTTTGGTGAT 3600
TGCGGCCCCAG GGTGTTTTTT GCTTCTGCACCA GTGAGACGGG CGAACCTCTGA TTGCCCTTCA 3660
CGGCCTGGCC CTGGAGAGGT TCAGAGCAAGG GCACACAGCT GTTTGCGGGGG AGCAGCGGAA 3720
AATCTCTTTT GATGGTTTTT AAGGCGGGGA TAAATACATG GCTGCTCTCTGG GTATGCTGAT 3780
ATCCCCACTC CGAGATAATC GCACCCAGCG CGACGGGGCA CGTGGTAATG CCCGGCAATTG 3840
CGGGCGCGCC CATCCTGATG TGGACCAACCA GCATCGACGT GGAGCACGATG CCACTCATCA 3900
GCATTTCGAT GTTTTTGTTA AAACCGGAGCA TGGGACTCCA GTGCCGCTTCC CGTCCGCTTA 3960
TCGGCTGAAT TTGATTGGCA GTGAGATATT TATGCCCAGG AGCCGACGCG AGACGGCGCG 4020
AGACAGAACT TAAAGGCCC GCTAAACAGCG CGATTGGCGT GTGACCCTGAT CCCGGACGAT 4080
GCTCCACCGC CAGTCGCTGA CGCTCTCATG GGGAGAAAAAT AATCTGTGGT AATGGGTGCT 4140
GGTCAGAGAC ATCCGAAATG AAGCGCGGAA CATTAGTGCA GCACGGTTCC AGACAATGG 4200
CATCTCTGGTC ATCCACGGCA TAGTAAATGA TCAGGACCAC GCACGCGTTGC CGAGAAGATG 4260
TGCCACCGGC CGGTCTTACG GTTCTGACGG CGCTCTGCTGC TACCATCGAC ACCACACGCG 4320
TGCCACCCAG TGGATCGGCG CAGAGATTTAA TGCCCGGCGG AATTTGGCGG CCCGCTGCGA 4380
GGGGCGACTG GGAGGCGGCA AGGCGCAATCA GCAAGCGACTG TTGGCGCGGG AGTTGGTGCT 4440
CCACGGCGTT GGAGAATGAA TCAGGCTCCC CCATCGGGCG TTCCACTTTT TCGCGCTTTT 4500
TCCGAGAAAC GTGGCTGCGG TGCTTACCA CGGCCGGAAG AGTCTGATGAA GAGACCGCG 4560
CATCTCTGCG GACACTTATG AAGGGTACTG GTTCTGACAT CACCACTCTGG AATGGAACCT 4620
CTTCCGGCGG CTATCATGCC ATACCCCGAA GGATTGTGCG CCATTGCTAG GTGCGGGGGA 4680
TCTGACGGCT CTCCCTTAAG CGAATCTTGC ATAGGAGAAG AGCCGGATAG TAGTTGAGG 4740
CGTGTGGAGA CGGCCGCGGC AAGGAATGGT CATGCGAAGG AGATGCCGCC CAACAGTCCC 4800
CGGCCCCAGG GGCCTGCCAC CATACCCAGG CCGAAACAAG GCTCATGAG CCGAAGTGG 4860
CGAGCCCCAT CTCTCCCCATG GGTGATGTCG GCCGATATTAGG GCCCGCCGAC CGCATCGTGTG 4920
GCCCGCGTGA TGCCCCGCA GATGGCGCTGG CGGGCAGAGA TGACTAGTCG CTAGCCCGGA 4980
AATTAATAC ACTCTACATA GGGGAATTGTG GAGCAGGATCA CAATACCCCCT ATGAGAAATA 5040
TTTTGTTTAA CTTTAAAGAG GAGATATACA TATGGCCCAT CATCDOCATC ATCAGCTGAT 5100
CGACATCCTC GCCAGCCACC CACATCTCTG GGAACAGCGG GGGCGGGGAGG CGGTTCAGCGG 5160
GGCGCGGATG AGCGCAGTATG ACATCCGGTTG CGCTCCGGTG ATGGACAGGG ACATGGCCGT 5220
GGAGAGGGC GCCAAGATCG CTAACCCGACG CATAGCTGAA GTTGTGGTCA AGATGAGGCC 5280
GGGCACCGGAG AGGGGCTGCA AACCCACCGG CGGTTGGCCT GAAACGCGCG GGGCGGGCGG 5340
TACCTGCGCG ACCTCCCCCG GTCTGCTGCC GGGTCAGGGT GGCGAGACC CGTAAGACGCT 5400
GCTCTACCCG CTGTTCAACC TGGGGCTGTC GCCCTTCAC CAGAGTATCC GAACAGTCAC 5460
GATCACCGCT CAGGGCCACG GTTCTGGTGCG CGGGATCGCG CAGGCGCCGG CCGGGACGCTG 5520
CAGCTGCGAC GCCCTAGGCG CCACTGCTGC GAGAGGTGAT ATGGCCGCCG ACAAAGGCGT 5580
GATGAAACTG CGCTAGCCCA TCTCCGGCTCA GCGAGTCAAC TACAACTGCG CCCGAAGTGAG 5640
CGACACCCCT AAGCTGAAGG GAAAAGCTCT CCGCGCCACTG TACAGGGCG TCCAGAAAACC 5700
CTGGGACGAC CCGCAGATCG CTCTGGCTCA CCCCCGGGCTG AACCTGCGCG GCACCGGGCGT 5760
AGTCCGCTG CAGCCCTCCG AGGGCTCCGG TGACACCTTC TTGGTCACCC AGTACCTGTC 5820
CAAGCAAGAT GCCAGGGGCT GGGGCAATGC GGGCAGCTTT GCACCCACCG TCGACTTCCC 5880
GGCGGGTCCG GGTGGCTCTCG GTGAAGACGG CAACGGCGGG ATGGTGACCG GTTGGCCCGA 5940
GACACCGGGC TGCTGGCTGC ATATCGGCGT CAGCTTCTCC GACGACGGGCA GTCAACGGGG 6000
ACTCGGCGAG GCCCAAATCAG GCAATAGCTC TGCGCAATTTCC TGTTGCGGCC AGCCCGAAAG 6060
CATTGACGCC GCGGCGGCTG GCTTCCGATC GAAAACCCCG GCGAACCGGG GATTCTGGAT 6120
GATCGACGGG CCGCCCGCCC ACGCCATCACG GATCATCAAAT TAGAGATACG CAGCCTTCAA 6180
CAACCGGGCA AAGGCACGC CGACCCGGCA GACCTTGCAG GCATTTCGAG CTTGAGTCCAG 6240
CAGCCGAGGC AACACGGCGT CGTCTCCGGA CCAGGTTCCAT TCCGAGCCGC TGCCGCCGCC 6300
GGTGGTGAGA TGTTCTGCAG GTGGATGCGG GACGATTCCC AGCGTGAGGA TGAAGAGGCC 6360
TGCCCCTACC CTGGGGCCAG AGGGAGGTAA TTTGAGAGGG ATCTCGGGCG ACCTGAAACC 6420
CCAGATCGAC CAGGTTGGAGT CGACCGCCAGG TTTCGTTCAG GCCCAGTGGC GGGCGGGCGG 6480
GGGGACGGCC GGCAGGCCG CGGTGGTCGG CTTCCAAGAA GCAGCCAATA AGCAGAAGCA
6540
GGAACTGAC GAGATCTGCA GGAATATAGC TCGGCCGCGC GTCCAATCTG CGAGGCGCGA
6600
CGAGGAGCCG CACGGGCGC GTCTCTCGCA AATGGGCTTT TGCCCAACAA CGGCGGCTTC
6660
GCCGCCGTAG CACGCTGACCC CGGACCGGCA CGGCGGCACTA CCGTGGCCTC CCCCACCCAC
6720
GGCAGCCGGC AACAGCGGGA ATGGCAGGCC GGGCGATCCCA AACGAGCAGC CTCCGGGCGG
6780
CGACCCAARCG GCAGCCGCGC CACCTGTCAT TGCCCAACAC GCACCCCAAC CGTCCGGGAT
6840
CGCAACACCG GTCAGCAGATT TCAGCTTCGC GCTGCTGCTG GGCCTGCTGG AGTCCTGAGC
6900
CGCCCCATTTC GACTACGGTT CAGCACTCTT CAGCAAAAACC ACCGGGACC GGCATATTCC
6960
CGGACAGCGC CCACGGGTGG CCAAATGACAC CCGTATCTGT CTCCGGCGGC TAGACCAAAA
7020
GCTTTACGCC AGGCCGGAAG CCAACGACTC CAAAGCACC CGGCCCGGG GGGCAGTTGG GCTCGGACAT
7080
GGGTGAGTTTC TATATGCCCT ACCCGGGGAC CGGAGTAACCG CAGGAACCG TCTCGGCTGA
7140
CGCCAACGGG GTCGCTGGGA GGGCGCTGTA TTAAGAAGTC AAGTTACAGC ATCCGAGTAA
7200
GCCGAACGCC CAGGATGCA CGCGCTGCCA CCGCTCGGCC CGGCGGCAAG CACCGGACGC
7260
GGGCCCCCT CAGCGCTGTG TTGGCATTG GTCTGGGACG GCCAACAAAC CGTGGCAACA
7320
GGGCGCGGCC AAGCGGCTGG CGCAAATGCG CGCGCCTTGG GTGCGCCCGCG CGCGGCGGCG
7380
GGGACCGGGT CCAGGAGGAC CGGCTCCGCG GCCGGCGCGG GCCGGGGAGA TGCTCCTTAC
7440
CCGAGACCAC CGGACCAGGC AGGCCGCGCT ACCGGCGCTGA GAATTTCTGA GAAATACATC
7500
ACACTGGCGG CGGCTGCAGC ACCACCCACA CACCCACTCA GATCCGCGCT GTCACAAGAC
7560
CGAAGGAAA GCTTAGGTTG CTGCCTCCGAC GCTGGGGCCT AATATAGGAT AACCCGTTGG
7620
GGCTCCCTAAA CGGTTCTTGA GGGTTTTTT GCTGAAAGGA GGAATATAT AT CGGAT
7676

(2) INFORMATION FOR SEQ ID NO:214:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 802 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:214:

Met Gly His His His His Val Ile Asp Ile Ile Gly Thr Ser 1 5 10 15
Pro Thr Ser Trp Glu Gln Ala Ala Ala Glu Val Gln Arg Ala Arg
Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Glu Gln Asp Met
35  40  45

Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Glu Val
50  55  60

Ser Phe Lys Met Arg Pro Ala Gln Pro Arg Gly Ser Lys Pro Pro Ser
65  70  75  80  85  90  95

Gly Ser Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Thr Pro

Ala Ser Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr
100 105 110

Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn
115 120 125

Val Thr Ile Thr Ala Gin Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln
130 135 140

Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser
145 150 155 160

Glu Gly Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala
165 170 175

Ile Ser Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His
180 185 190

Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile
195 200 205

Lys Thr Trp Asp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn
210 215 220

Leu Pro Gly Thr Ala Val Pro Leu His Arg Ser Asp Gly Ser Gly
225 230 235 240

Asp Thr Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Gly
245 250 255

Trp Gly Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val
260 265 270

Pro Gly Ala Leu Gly Glu Asn Gly Asn Gly Gly Met Val Thr Gly Cys
275 280 285

Ala Glu Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp
290 295 300

Gln Ala Ser Gln Arg Gly Leu Gly Glu Ala Gin Leu Gly Asn Ser Ser
305 310 315 320
Gly Asn Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala
325  330  335
Gly Phe Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp
340  345  350
Gly Pro Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile
355  360  365
Val Asn Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala
370  375  380
Phe Leu His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp
385  390  395  400
Gln Val His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp
405  410  415
 Ala Leu Ile Ala Thr Ile Ser Ser Ala Glu Met Lys Thr Asp Ala Ala
420  425  430
Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile Ser Gly Asp Leu
435  440  445
Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly Ser Leu Gln Gly
450  455  460
Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Gln Ala Ala Val Val Arg
465  470  475  480
Phe Gln Glu Ala Ala Asn Lys Gln Lys Glu Asp Glu Ile Ser
485  490  495
Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg Ala Asp Glu Glu
500  505  510
Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe Val Pro Thr Thr Ala
515  520  525
Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro
530  535  540
Val Ala Pro Pro Pro Pro Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro
545  550  555  560
Gly Asp Pro Asn Ala Ala Pro Pro Ala Asp Pro Asn Ala Pro Pro
565  570  575
Pro Pro Val Ile Ala Pro Asn Ala Pro Glu Pro Val Arg Ile Asp Asn
580  585  590
Pro Val Gly Gly Phe Ser Phe Ala Leu Pro Ala Gly Trp Val Glu Ser
595  600  605
Asp Ala Ala His Phe Asp Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr
610
615
620
Gly Asp Pro Pro Phe Pro Gly Gln Pro Pro Pro Val Ala Asn Asp Thr
625
630
635
640
Arg Ile Val Leu Gly Arg Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu
645
650
655
Ala Thr Asp Ser Lys Ala Ala Ala Arg Leu Gly Ser Asp Met Gly Glu
660
665
670
Phe Tyr Met Pro Tyr Pro Gly Thr Arg Ile Asn Gln Glu Thr Val Ser
675
680
685
Leu Asp Ala Asn Gly Val Ser Gly Ser Ala Ser Tyr Thr Glu Val Lys
690
695
700
Phe Ser Asp Pro Ser Lys Pro Asn Gly Gln Ile Trp Thr Gly Val Ile
705
710
715
720
Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Glu Asn Arg Trp
725
730
735
Phe Val Val Trp Leu Gly Thr Ala Asn Asn Pro Val Asp Lys Gly Ala
740
745
750
Ala Lys Ala Leu Ala Glu Ser Ile Arg Pro Leu Val Ala Ala Pro Pro
755
760
765
Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro Ala Pro
770
775
780
Gly Glu Val Ala Pro Thr Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu
785
790
795
800
Pro Ala

(2) INFORMATION FOR SEQ ID NO:215:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 454 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION:- SEQ ID NO:215:

```
GTGCAGCGGC
TCGACGATGA
GGCCAGCGCA
GCCGCGGAC
ACGAAGAAGC
TCAGGAAGCC
GTCGAGACCC
GCCAGATCAT
GCCACAGCGCT
GATCGGCCGC
ACACCGCCCA
GGCCGCGCG
GGAACCCAGC
GATGCGCTCC
```
<table>
<thead>
<tr>
<th>Sequence Details</th>
<th>Sequence Description</th>
<th>Length</th>
<th>Strandness</th>
<th>Topology</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) SEQUENCE CHARACTERISTICS:</td>
<td>(i) SEQUENCE CHARACTERISTICS:</td>
<td>470 base pairs</td>
<td>single</td>
<td>linear</td>
<td>nucleic acid</td>
</tr>
<tr>
<td>(ii) MOLECULE TYPE: Genomic DNA</td>
<td>(ii) MOLECULE TYPE: Genomic DNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(xi) SEQUENCE DESCRIPTION:</td>
<td>(xi) SEQUENCE DESCRIPTION:</td>
<td>279 base pairs</td>
<td>single</td>
<td>linear</td>
<td>nucleic acid</td>
</tr>
<tr>
<td>(i) SEQUENCE CHARACTERISTICS:</td>
<td>(i) SEQUENCE CHARACTERISTICS:</td>
<td>219 base pairs</td>
<td>single</td>
<td>linear</td>
<td>nucleic acid</td>
</tr>
<tr>
<td>(ii) MOLECULE TYPE: Genomic DNA</td>
<td>(ii) MOLECULE TYPE: Genomic DNA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ACACGGTCTGA ACTCGAGAGG CCCCTCGTGGA AGGTGTCGAC CGACAAGGTGACGACCGAAA 60
TCCCTGCGCC GCGCCGAGTGG TGCTGACCAA GTATCGGCC CAAGAAGAGG ACACGGTCTGA 120
GGTGGGCGGCG AGATCTGAGC TCAAGTGCGGA CGCACCAGGTG GCCGCGGAGG CGCAGGGTCCC 180
GGCACCACGAG AAGACTCTGAC CGGCGCGCAC CGGGAATCGCA 219

(2) INFORMATION FOR SEQ ID NO:219:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 342 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:219:

TCGCGTGGCCG CATCGGCAGG GCCGCGGCGGC GCAGACCCAGGC ACCCAAGCCG ACCCAAGGGGC 60
CAGCGCGAGG GCCGCGGACG GACCCACCGGC ATGCGCGCGGC CGCAGCGCGGC CGCAGCGGTG 120
CGCGCGGGGG GCACCCGAGG TGGTGCGCGGC TGCGGCGCCG CGCAGCGCGGC CGCAGCGGTG 180
TGGCTTGCCG GCCGGTGGCCGG GCCGCGGCGGC CGGCGGCGGC CGGCGGCGGC CGGCGGCGGC 240
TGCGCGCGGC TGGACAAAAA AGGCGCGGCGG AGCCCGCGGC AGCCCGCGGC AGCCCGCGGC 300
CGCGCGCGGC CCCGCAAGGCG GGGCGCTGAG ATCCGATGCC GC 342

(2) INFORMATION FOR SEQ ID NO:220:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 515 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:220:

GGGTCTTGTG CATATCCGAC GCCGAGCGAGG AGCGAGCAGG GCCGCGTCGGC GGGCAGTTGG 60
CCCGGATCAGG TTGGCCTGCC GACATCAGGC CGCGCCCGGC CCCCAAGCCCGG GACACCAAGG 120
CCGCCCGAGG AGCGGAGCAGG AGCGGAGCAGG GCCGCGGCGGC GCCGCGGGCGG AGCCTGAGGG 180
AGCCAGCGGG GCCCGCGGAGG GCCGACCGGT AGCTGACGCC GCTGGTGGCGA AGCCTGAGGG 240
CCGAAACCAA CATCGACCTC GCCGCGGGTG GCCGACCCCGG AGTGGTGGTGG CGCATCGGCA 300
AACAGGATGT GCTGGCGCAG GCTGAACAAA GAAGCGGGGC GAGGCAGCGG GCCGGCGGAGG 360
CCGCTCATCA CGCGTTAAC AGCGTTGCCA CGAGACCGGC TTGGCAAGCTC TTGGCGGCGGTC 420
TTGGCTCGGCT GAGCGCCGTC GCCGACCGAG TTAGGTTTAG GGGCGCGGAAA GTTTTCAGTT 480
CGCCAGGAAAG GACACCGCGA ACAGGAGCGG CACCC 515

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 557 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:221:

CCGACCCCCA GGTGCAGATT CAACAGGCAC ATTGAGAAGC ACGCCGACAC CACAAAGGCC 60
TGACTCACAAC GCCGCGCCGAA GTGACCGGTA ACCAGGACTCA ATTGAGAGTG CGACTCAACC 120
GACAGCTGACC GGACATCGGA AAAGCTTCAAG TGAAATGAGCC CCAACGCCCCT ACAGTGCGCC 180
ACCAGGCACC CGCCGCGGGA AGGTCCGGTA AAGGCCAACGA ATACACAAAC GCGCCGGCGG 240
CGTTTCGAGCC CCAAGCTGCTG ACCGGCGCGAC AGAGCTCCTGA AGAACCTGCAAC AGCCTCGATG 300
ACCAGGCTCGT TAGCAGCCGGA GCTCGCCAGCA AGAAGGCCTG CAGAAGCAAAT GCGAGTGCCG 360
TGCAAGCAGAAG ATCGCGCGAG GAAACCATAGG CTCTCAGGCA GCTCGAGAGC GCAGAAGATGC 420
AGGAAGCAGCT TAGCGCCTGCA TTGGCGTCTG TAAGTGAAGT CGCGGCGGCA GGCACAACCG 480
CGAGCCTCCTGA CGAAGCTGTGG GACGGATCTC AGCCGTCCTA CGCACCACCGG ATCGGTCTGG 540
CTGAAGCTTGG CTGAGATG

(ii) MOLECULE TYPE: Genomic DNA

(ii) INFORMATION FOR SEQ ID NO:222:

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:222:

CAGGATAGGTT TTCCGACATCC ACCTGCGGCT GGCAACGCGT GCCCGACGGG GGTGATAGCC 60
AGGAGTGGAC CTGACGGCCAC CGACGATCGA TGAGGAGCTC AACAGAAATG GCCCTTGCGG 120
TTCAAGTACC GCCACTCGGCA GAGGAGCTGA CGGAGGGGAC GTTACCCCGC TGGCTCAAAC 180
AGGAAGCCGA CACGCTCGGAA CTCGACCGGAC CCCTCGTGGGA GGT 223

(ii) MOLECULE TYPE: Genomic DNA

(ii) INFORMATION FOR SEQ ID NO:223:

(A) LENGTH: 578 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:223:

AAGAAGTACA TGTCGCGCCGC GATGTGCAGGG AACCAAGCGA GCCAAAGCGA GCGATAGCCTG 60
ACCCAGAGAC CGCGATAAGC CCACTCGCGG GCCAACAAACA TACGCTCACC CCGGTAGCCTG 120
AGGAAGCGCA CGGGAGCAAC CCAAACACGC GCCGGTGCGGA CCAAGCTCTC GCCCTTGACG 180
CAGGACTGTTG CGGGCGCGGGC TGCAACGCTT TGCTGCTGCG TGGCTGACGG CACCGGCCTG 240
AACAGCATGG GCGACGGTCA GCGTGGGATG TCCCAAGGCTG TGGAGTTGCG TGCAGGAAATC 300
GCTGGCCGCG GCGCGCGCTCG GGAGCTGGTG GGCTGTTGAT GCGAGAGCAG GCGCGCGCTG 360
TGGGCGAGGG GAAACAGCAGC GTGCTACCGA CGGCCCTGAG CGCAATCACG 420
GCCCTCGGCG GACCGAGGCTG ATGGCGAGTGA TGGGTCCGGAG AGAGCGCTGG GCCGCAACCG 480
CCGACGCGAT ACCGCGCTGGG ACCCGAGCTAC CGCGGCGGCTG TTCCCGCGCC CGGTCTTTGG 540
ACTGCTGATG AACGGCGCACCC CGCAGCGCTAC AGCCGCGA 578

(2) INFORMATION FOR SEQ ID NO:224:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 484 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:224:

ACAGCGATCG ATGTATACG ATGAGAGACG AGAGGATCGT GCCCGTCCC CAGTTGACCG 60
ACGACGACCG CGCCGCCGCC TTGAGAAGGG CTTCTGACCG GCGACGAGC 120
TCAGAGTACG GTCTAGCTGG GCCGCCACCA ACCCTACCCA GGTGGTCAAG CAGCGGAGA 180
GGGATGAAAGT CTTGCGGAAA AGTAGGGTGT CTGGCGGTCT TTAGGCTTGG CCCAAGGTTG 240
GGCGATGCTAA GCCCGGAGAG ATAGGACGAG AGCTGGGAAA TGCCGCCCAC CCCGGGTGCT 300
CCGGCGGGCCG ACCGATGCGG CCAGGAAGGCC TGTTAGGGGCT GACCCCTGGC ATACGGGGA 360
GAAGCGCCTT GACACGCGAC GTCTCACACTT CGGGCGGAG GGGGAAACC 420
GCC

(2) INFORMATION FOR SEQ ID NO:225:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 537 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:225:

AGGACTGGCA CGGCGACGCA CCACATCCCG GGCGGCGGGA CCAGCATCTC GCCCTTGACG 60
ACGCGCTGTG CGGCCAGACG TGCAAGCTCT TGCTGGGCTA GGGGTAAGAG CACCGGCGCC 120
AAGACAGTGG GCCAGGTCCA CCGTGGGATG TCCCAAGGCT GGAGTGGGCC TGGGAAATT 180
GTCCAGGCGG CGGCGAAATG GAAGCCCTTG CGGGTGATAT CGGACGGGCA GCCGACGCCC 240
TGCGCGAGCG AACGAAACCA GTGGGGCAGC ACCGCTTGC AGGCGCATA CTG 300
GGCGTGTCGG ACACGGACCA CGGACGCTAG GTGCGCAGAT AGACCGGGA CCGGATCAAC 360
CCCGGGCAT ACCCGTGCGG AAGCAGTCCA CGGGCGACTG TCCCCAGCCA CGGTCTTGGC 420
ACTTGGTACT GACGGCGC CGCCAGATCG AACGCGACGG CCATCGCGCC GAAAGACAGC 480
ACGAAGTCACA CGGCGGACCA TTGGGTCGGA CAGAAGCTAC CCAAACAGCA CCCCGGC 537

(2) INFORMATION FOR SEQ ID NO:226:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 135 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:226:
Gly Gly Ala Ala Ala Gly Gln Gln Ser Val His Pro Phe Ala Asn
1  5 10 15
Leu Ile Ala Val Asp Asp Glu Arg Ala Glu Arg Asp Asp Glu Glu
20 25 30
Arg Gln Glu Ala Val Gln Arg Gly Pro Arg Gly Asp Glu Ala Asp
35 40 45
Pro Val Ala Asp Gln Glu His Pro Gly Asp Gly Ala Asp Gln Cys Arg
50 55 60
Pro Ala Asp Pro Pro His Asp Pro His His Gln Arg His Gln Asp His
65 70 75 80
Thr Gln Gln Gly Ala Gly Glu Pro Ala Glu Ser Val Val Thr Glu
85 90
Asp Gly Leu Pro Asp Arg Asp Glu Leu Leu Thr Asp Arg Val Asn
100 105 110
His Gln Ala Val Pro Gly Val Val His Pro Met Val Val Gln His
115 120 125
Leu Pro Gly Leu Ala Val Arg
130 135

(2) INFORMATION FOR SEQ ID NO:227:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 156 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:227:

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

(2) INFORMATION FOR SEQ ID NO:228:

(i) SEQUENCE CHARACTERISTICS:
(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:228:

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

(2) INFORMATION FOR SEQ ID NO:229:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 72 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:229:

Thr Val Glu Leu Asp Glu Pro Leu Val Glu Val Ser Thr Asp Lys Val
  1    5     10     15
Asp Thr Glu Ile Pro Ser Pro Ala Ala Gly Val Leu Thr Lys Ile Ile
  20   25    30
Ala Gln Glu Asp Thr Val Glu Val Gly Gly Glu Leu Ser Val Ile
  35   40    45
Gly Asp Ala His Asp Ala Gly Glu Ala Ala Val Pro Ala Pro Gln Lys
  50   55    60
Val Ser Ala Gly Pro Thr Arg Ile
  65   70

(2) INFORMATION FOR SEQ ID NO:230:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 113 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:230:

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

(2) INFORMATION FOR SEQ ID NO:231:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:231:

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

(2) INFORMATION FOR SEQ ID NO:232:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 185 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:232:

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

(2) INFORMATION FOR SEQ ID NO:233:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 71 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:233:

Val Ser Thr Ser Thr Trp Val Pro His Pro Val Arg Asp Arg Val Ile
1  5  10  15
Gly Gln Arg Trp Thr Cys Ala Asp Arg Arg Ser Ile Glu Glu Ser Thr
20  25  30
Glu Met Ala Phe Ser Val Gln Met Pro Ala Leu Gly Glu Ser Val Thr
35  40  45
Glu Gly Thr Val Thr Arg Trp Leu Lys Gln Glu Gly Asp Thr Val Glu
50  55  60
Leu Asp Glu Pro Leu Val Glu
65  70

(2) INFORMATION FOR SEQ ID NO:234:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 182 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:234:

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

(2) INFORMATION FOR SEQ ID NO:235:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 160 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:235:

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

218

(2) INFORMATION FOR SEQ ID NO:236:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:236:

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

Pro Arg

(2) INFORMATION FOR SEQ ID NO:237:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 271 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:237:

ATGCCAAGCC GGTGCTGGATG CCAGGGCTCGGG ATCCGGAGGGG ACCGTCATTCC 60
GGTGCTGAA GAGACCTGCGG ATCGGTGGTTC CTTTTGACGA GCACTGAGTG GAGGTGTC 120
CGACAAGGGT GGACCCCCAG ATACCGTCCC CGGTGGCTGG ATGATTACCGG 180
CGACACAGAG CCGCCCCGAGG CCGTGCGGGG CGGAGTTGGC CCGGATCGGT GTGCTGCGG 240
AGATCGGCGC CGCCCCGCCC CCGAAAGCCC C 271

(2) INFORMATION FOR SEQ ID NO:238:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 89 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:238:

Ala Lys Pro Val Leu Met Pro Glu Leu Gly Glu Ser Val Thr Glu Gly
1   5
Thr Val Ile Arg Trp Leu Lys Ile Gly Asp Ser Val Gin Val Asp
10  15
Glu Pro Leu Val Glu Val Ser Thr Asp Lys Val Asp Thr Glu Ile Pro
20  25  30  35  40  45
Ser Pro Val Ala Gly Val Leu Val Ser Ala Asp Glu Asp Ala
50  55  60
Thr Val Pro Val Gly Gly Leu Ala Arg Ile Gly Val Ala Ala Glu
65  70  75  80
Ile Gly Ala Ala Pro Ala Pro Lys Pro
85

(2) INFORMATION FOR SEQ ID NO:239:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 107 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:239:

GAGGTACCGG ATGGCCGGAG GAGCACCAGCA GGACCGCGGC CGAACCCTGGG GTGCGGTC
60
TGATATGTG GACCCGCTTC GTCCGCTCGC CCGAGGTCAT TGACG
107

(2) INFORMATION FOR SEQ ID NO:240:
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 339 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

ATGAAGTTGGA AGTTGCTCG CCTGACTCTG GGCGACGCTG GGGCTTTGTG  
TTGCTGCTG CGCAGGCAAG GCCGACTCCA CTCGCCGCGC ATCAGCGGAG CAGCGAAA  
GGCGATGGG CGGCTGCTCG ATGGTTTCTG GGGCGACCCG CGGCGAGAAG  
TACCCCGACG GCCACAGAAT AGGCAAACGT GGTTCACCG GCCACAGTTT  
TACTTCGATT GTGCGACCG CGCAGGAGCA CTCCCCGAGG CGCAGCCGACC GCCGAGGATG  
GGGGGCGAA TCTGGTCCGA GCGCCCCAAC GCTCCCTGA

(2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

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

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 371 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO: 242:
| GTGACCCACGG TGGGCTGCCC ACCAACCCGG GCACCGCCAG CCGGGCCGCC GCGGGGCGCT | 60 |
|-----------------|-----------------|-----------------|
| CCGGGGCAA CCGGCGGGCC GGGGTAACCC CCAAGGCTTC AGCGGCAAG GCGGGCGGCG | 120 |
| GTGCGCAATGG CGGTGATGGG AGTTGCGCG CTACAGCGCG CCGGGCCTGC ATCGGGGTCG | 180 |
| CGGGGCCCAC CGGGGGCAAC GCGGGGAAAG GGGGGCGGCC TGCGCACAAA GCGGACGGCT | 240 |
| CAGGTGGGGCA CGGGCGCAAA GGGGGCAGCC GGGTGCGCG CGCGAAGGG GCAGTGATCG | 300 |
| GCACCCACAG CGGGATCGTC GCGGGTTCGG GTGGGCGCG CGGGCTCGGC GGGGGCGGGG | 360 |
| GAAACGACAG C | 371 |

(2) INFORMATION FOR SEQ ID NO:243:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 424 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:243:

| GTCCGGGGTCC CACCACCCGG CGGGCGGGCC CCTAGGGGCC GGGCCGACCA GCCCCTTTTC | 60 |
|-----------------|-----------------|-----------------|
| TTTGACTCTTT CAAGAAAGGA GCTCTCTCTAT TGAGCGCTCA TGGGGATGAC ATCGGGAGCC | 120 |
| ATGGGGACAC TCGACGTCGA CATCCTGGCC AAGGCTTAGG TCAGATGGCA CGGGCGGGCC | 180 |
| GTGGGTGAGG ATCGGCTTAC CGGCTATAGG AGGGTTGCGG GTACCTGTGG CGGAGCAGA | 240 |
| CGGGCCTGGG CGGGGGTTTG TGCTCTACGG GGTGCTTGCG AAGGSGTGTG TCCGGACAGT | 300 |
| CGTCGGGGCC CCAATGGGGT GGGAGCGGCT GTCAGCAGAG CGGGTCAGGG CTGTTGGGCT | 360 |
| CGGACTACGG CGGTGCAACA ACTGGGTTGA CATCAGGGTG CCGGAGGTCG ACGGCGATCC | 420 |
| CGAC | 424 |

(2) INFORMATION FOR SEQ ID NO:244:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 317 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:244:

| GCAGTGGGGCG CGGGCGGGTAC CACCGCCAAT GGGGAAGCGT TTCCCCAACCC CAACGATTCC | 60 |
|-----------------|-----------------|-----------------|
| TTGCATCTGG GTCGAAATGG CTTCAAGCCG GGCAATTCTG TACGCGGGG CGCCGCGTCTA | 120 |
| AGGGCGGATG CGTCCTACT GCAGCGGTCC GACGGCTGCG CTGGCAGGCA GCGGGCGGAT | 180 |
| TGCCGACTTG TGCGGACCCA GTGGCGACGC ACCCGTCACG CGCGACAGGT GGGCATCCG | 240 |
| GTGAGTCCG GCAGATGGCG TCCGGAAGCT GCAGCGGGGT CGAAACTCAC GAAATGGCTG | 300 |
| CGGCTTACAG ATCCTGCT | 317 |

(2) INFORMATION FOR SEQ ID NO:245:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 422 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:245:

TGGCGTATGCC GCTTCGACGC CCGTGGCCGG TCAACGCAGC CGAAGGCACGC GCTTCGCTGC 60
CGAGGAATGG TTGTAGAGGC ATCGCAGTTT GCCCGTGCGC ACCGACACCG CGGCAACCAG 120
TGAACTGAGG GCAGGAATTC GCGGAAATG TGGGTCGTGC TGGGCGTGGC AACGCTAAGC 180
GTACCTGAGG TTGGGTGGGC CTTACTCCGC GACCCGCAGC GGCTTCTCGC TGGGCGTCTC 240
GGACAGGAAG ACCTGTGATG CGATGTGGGC TGGGACTCTTGG AAGGCTCTGG TGGCCGCGAT 300
GCACTGCGTC TGCCGATGAG ACGCAAGAT GCACCTCAGC GACAGGAGTC GGCTGAGCGA 360
GATGGGCGTG CCGAGTCTTTA GAACTTGTGG CCAAGCCTGG GCCGTGGCCG CAGGCTGGCG 420
AT

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:246:

GCCGTGGCGCT GAACACCCGC CCAGCGCTGGC CAGATCTCCG CAGACTTCGG GTGCCGCGCG 60
TGCCGTCGGTT GCTCTTCTTA CCGGCGGGCGG GACCCAAATG GTCGCTACA GCGAGGTGTC 120
GCGACCCGGT CATGGAGTGC ATGATGTGCC CAGCGCCGGC ACGGGTCTGG ACGAAGGCACG 180
CATCGGGCCT GATCAGGCGC CAGCGGTTGG ACAAGTGCTG CGAAGCTATA CGGGCCGCGA 240
GAGGTGGCGGT GACACCTGGC CGCTTCTCGC GCGAATATGC GTTGATCTTC CTGGCCGGCT 300
AGTGTTCGGG CAGTCGCTGG GAAAATGGAC GCAGCGGCGG GATGGGCGGC ACGGCGGTTG 360
ACCTGAGGTG TGGCGCTTTTG CGCCAGTGGG TCGACAGCAGC GTACTCGGGC TAAATGCTGG 420
CCCCGA

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:247:

AGACCCGGCGA GGGTGGGCTG GCTGGCCGGG CCGATTGTGGA TAAATCGGCG TGGGTCGACG 60
CGCCGATCAA CTAGTGAGGG GCAACGCTAG GCTGGGTGGG GTAACACAGT AAAAAGTTTA 120
TCAAAAGGAC GAAGAAGGTT GCAAGTGCAC CTTGTGCGGC CTACCCGCCG ATGCTCGCGA 180
CGGAAACCTT GACCAAGACG AGATACCGGT GCGCGCGCGA GACGTGGCGA GCAAGCGCAG 240
TGACATCAGA ATCGGCGGCA ATCGTGCTGG CCGACATCCA TACCGTGCCG ACCGAATGTG 300
GCGAAGCCGA TTACCTGTG GTTCCTCG 327
## INFORMATION FOR SEQ ID NO:248:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:248:

<table>
<thead>
<tr>
<th>Asp</th>
<th>His</th>
<th>Gly</th>
<th>Gly</th>
<th>Pro</th>
<th>Ala</th>
<th>Thr</th>
<th>Asn</th>
<th>Pro</th>
<th>Gly</th>
<th>Ser</th>
<th>Gly</th>
<th>Ser</th>
<th>Arg</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Ala</td>
<td>Thr</td>
<td>Gly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td>Lys</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Gly</td>
<td>Gly</td>
<td>Asp</td>
<td>Gly</td>
<td>Ser</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ala</td>
<td>Thr</td>
<td>Ser</td>
<td>Gly</td>
<td>Pro</td>
<td>Ala</td>
<td>Ser</td>
<td>Ile</td>
<td>Gly</td>
<td>Val</td>
<td>Thr</td>
<td>Gly</td>
<td>Ala</td>
<td>Pro</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td>Gly</td>
<td>Lys</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Ser</td>
<td>Asn</td>
<td>Pro</td>
<td>Asn</td>
<td>Gly</td>
<td>Ser</td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td>Ile</td>
<td>Gly</td>
<td>Ala</td>
<td>Asn</td>
<td>Ser</td>
<td>Gly</td>
<td>Ile</td>
<td>Val</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td>100</td>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td>Ala</td>
<td>Gly</td>
<td>Gly</td>
<td>Gly</td>
<td>Asn</td>
<td>Gly</td>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## INFORMATION FOR SEQ ID NO:249:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: protein

### (xi) SEQUENCE DESCRIPTION: SEQ ID NO:249:

<table>
<thead>
<tr>
<th>Met</th>
<th>Ala</th>
<th>Ala</th>
<th>Ala</th>
<th>Gly</th>
<th>Thr</th>
<th>Thr</th>
<th>Ala</th>
<th>Asn</th>
<th>Val</th>
<th>Glu</th>
<th>Arg</th>
<th>Phe</th>
<th>Pro</th>
<th>Asn</th>
<th>Pro</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>10</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Asp</td>
<td>Pro</td>
<td>Leu</td>
<td>His</td>
<td>Leu</td>
<td>Ala</td>
<td>Ser</td>
<td>Ile</td>
<td>Asp</td>
<td>Phe</td>
<td>Ser</td>
<td>Pro</td>
<td>Ala</td>
<td>Asp</td>
<td>Phe</td>
</tr>
<tr>
<td>20</td>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Thr</td>
<td>Glu</td>
<td>Gly</td>
<td>His</td>
<td>Arg</td>
<td>Leu</td>
<td>Arg</td>
<td>Ala</td>
<td>Asp</td>
<td>Ala</td>
<td>Ile</td>
<td>Leu</td>
<td>Leu</td>
<td>Arg</td>
<td>Arg</td>
</tr>
<tr>
<td>35</td>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asp</td>
<td>Arg</td>
<td>Leu</td>
<td>Pro</td>
<td>Phe</td>
<td>Ala</td>
<td>Glu</td>
<td>Pro</td>
<td>Pro</td>
<td>Asp</td>
<td>Trp</td>
<td>Asp</td>
<td>Leu</td>
<td>Val</td>
<td>Glu</td>
</tr>
<tr>
<td>50</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Gln</td>
<td>Leu</td>
<td>Arg</td>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
<td>Thr</td>
<td>Ala</td>
<td>Asp</td>
<td>Thr</td>
<td>Val</td>
<td>Arg</td>
<td>Ile</td>
<td>Asp</td>
<td>Val</td>
</tr>
<tr>
<td>65</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Ala</td>
<td>Asp</td>
<td>Met</td>
<td>Arg</td>
<td>Pro</td>
<td>Glu</td>
<td>Leu</td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
<td>Ser</td>
<td>Lys</td>
<td>Leu</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Ser</td>
<td>Leu</td>
<td>Arg</td>
<td>Leu</td>
<td>Tyr</td>
<td>Asp</td>
<td>Ser</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:250:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 41 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xvi) SEQUENCE DESCRIPTION: SEQ ID NO:250:

Ala Tyr Ala Leu Arg Ser Arg Cys Arg Val Asn Ala Pro Glu Ala Ile
1     5  10  15
Ala Ser Leu Pro Arg Asn Gly Ser Ile Thr Ile Ala Val Cys Arg Arg
20  25  30
Ala Pro Thr Pro Pro Ser Asn Val Asn
35  40

(2) INFORMATION FOR SEQ ID NO:251:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 25 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xvi) SEQUENCE DESCRIPTION: SEQ ID NO:251:

Val Pro Leu Asn Thr Ser Pro Arg Leu Pro Asp Leu Pro Asp Ser Val
1     5  10  15
Val Pro Pro Val Ala Ser Leu Leu Ser
20  25

(2) INFORMATION FOR SEQ ID NO:252:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xvi) SEQUENCE DESCRIPTION: SEQ ID NO:252:

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

(2) INFORMATION FOR SEQ ID NO:253:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 213 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:253:

GCTTTGGACC CTGGAGCCGAC GGTGTGCGTC TGGGGGTCGA TTCGTCTCCTGGGAAATGCA
ACTAAAGACC ACUGTGACAC CCAACCGCGG GCCCGGCTATG GCCGGCTGCGG GCCGTACAGG
TTTGACCGCG GCGCCGAAAGG TCGCCGCTGCTGGGCCCATGCC AGATCGCAGCA CGTCTGCTTG
AACATTGGTT GGAGCGGCGTG GTAAGGGCGAG

(2) INFORMATION FOR SEQ ID NO:254:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 367 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:254:

CCGAGCTCGCT GTTGCGCGGCC GCGCGGTCGCG GCCGGCGCGGC TGGGGGCGCGC ACCGACGCGC
GGGCCTCGGC TACCGCGCGGC ACCGCGCGGCC ACCGCGGCGAGT CGCGCGCGAC GCCGATGCC
TGCGCGCGCG CAGGCGCGGCG GCCGGCGCGGC GCGAGCTCGCT GCCGGCGAGC
ATGGCTCGCGC GTGCCGCTGCT ACCGCGCGGCA CGCGCGCGAC CGCGCGGCGC GCCGGCGCGGC
GCGGCTCGCGC CACACTCGCTT TGGGCGCGTG CGCGACGCGG CGCGCGCGCG CGCGCGGCGGC
GACACGGCGG CAGGCGCGGCG GCCGGCGCGG AATGCGGCCTCG TGGGGGCTGCT CGTGGGACT
GGTGCTGA

(2) INFORMATION FOR SEQ ID NO:255:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 420 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(x) SEQUENCE DESCRIPTION: SEQ ID NO:255:

AAGGCGTGTAT TGGCAAGCGC ACCCGCGCAGC GCGCGGTAAGC CGCGCGACGG CGCGGCGCCC
GACCGCGAGC GCCCGGGCTCTT CAGCGCGCTC CGCGCGCGAC CGCGCTGCGT GCCTCGCTCGT
CTACTTCGAC GCGCGCGCGT TCGTCAAGCT TCAGCGCCGCC GAGACAGGGA GCTCGCTGGCC
GTCCGCTCCTA TGGGACGGCT GCGACGCCGC ATTGTCCAAC CCCTGGGCCCT ACCCAGAATT 240
CCGCGCGCGCA CTCGTCGCAA GGCGGCACCA TCAGCTCACTA ACCGAATCGG AGCTGGCCGA 300
CGCAGCGAGCT GACGCGGAGG ACTTCTGGGC CGCACCAGC CGTCTGAACT CACCAGGACG 360
GTTAACGAGCCGCGGCCGCA CTGCGCCGGCA ACACATGCGT TAGCGCGAGG CGCACCAGT 420

(2) INFORMATION FOR SEQ ID NO:256:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 299 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:256:

CTCTTGTCGG TGTCATCGGC GTTACCAGGG GAACCCCGCC CGAACCGCGGT ATGCTGCGGG 60
GCACCGCGCC GGCCCGCGCT GCGCGCGGCT TCATTTCAGG CAGCTCGCAG GCAGCTGCGG 120
GGCAGCGGCG GGCAGGGGGT CGGTGACCAA CGCGCGGCGG CTGGCAGCGG CGGCGGGCGG 180
GTGCGCGCGG CGCGCGGCGG GGAGGGCGGG GTGGCTGAGG GCAGCGGCGG GCAGCGGCGG 240
TGGCGCGCGG GGGGGCTTCG GGGGATACCG GAACGCTACGG CACCAGCGGG GCAGCGGCGG 299

(2) INFORMATION FOR SEQ ID NO:257:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 20 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:257:

Leu Glu Pro Trp Ser Asp Gly Val Gly Leu Gly Val Asp Ser Phe Ser  
1  5 10 15
Ala Lys Val Asn  
20

(2) INFORMATION FOR SEQ ID NO:258:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 121 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:258:

Glu Leu Leu Phe Gly Ala Gly Gly Ala Gly Gly Ala Gly Gly Ala Gly  
1  5 10 15
Thr Asp Gly Gly Pro Gly Ala Thr Gly Gly Thr Gly Gly His Gly Gly  
20
<table>
<thead>
<tr>
<th></th>
<th>Val</th>
<th>Gly</th>
<th>Gly</th>
<th>Asp</th>
<th>Gly</th>
<th>Trp</th>
<th>Leu</th>
<th>Ala</th>
<th>Pro</th>
<th>Gly</th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
<th>Gly</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Gly</th>
<th>Gly</th>
<th>Gln</th>
<th>Gly</th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
<th>Ala</th>
<th>Arg</th>
<th>Ser</th>
<th>Asp</th>
<th>Gly</th>
<th>Gly</th>
<th>Ala</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Gly</th>
<th>Thr</th>
<th>Gly</th>
<th>Thr</th>
<th>Gly</th>
<th>Thr</th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Gly</th>
<th>Arg</th>
<th>Thr</th>
<th>Leu</th>
<th>Leu</th>
<th>Leu</th>
<th>Gly</th>
<th>Gly</th>
<th>Gln</th>
<th>Gly</th>
<th>Gly</th>
<th>Leu</th>
<th>Gly</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Gly</th>
<th>Ala</th>
<th>Gly</th>
<th>Glu</th>
<th>Gly</th>
<th>Gly</th>
<th>Thr</th>
<th>Gly</th>
<th>Gly</th>
<th>Gly</th>
<th>Arg</th>
<th>Arg</th>
<th>Arg</th>
<th>Trp</th>
<th>Arg</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:259:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:259:

Gly Val Ile Gly Lys Ala Thr Ala Gln Arg Pro Val Ala Ala Gly Arg
1    5    10    15
Pro Arg Pro Arg Pro Gln Arg Pro Val Ser Asp Arg Val Ser Asp Gln
20   25   30
Arg Arg

---

(2) INFORMATION FOR SEQ ID NO:260:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 99 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:260:

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

(2) INFORMATION FOR SEQ ID NO:261:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 282 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:261:

TCCTGTTCGG CGCCGCGCGG GTGGCCGCTG TTGGCGGTGA CGTGCTGCGA TTCTCTGCGA
CGGCCCCCGG GGGGCGCGGT GGGCGCGCGG GGGCGCTGCG CCGTGCTGCG GTGGGGTGGGG
CGGGGGCGCC CGGGCGGATC GGATGGTGCG GGAACAGCAG TGGCGGGGGG TGGCGGGGGG
CGCCCTGCGT CTGGGGCGAC GGGGTGCGCG GGGCGGGCGG TGGGTGCAGG TCCACACCGG
GCGGTGGCGG CGGGCGGCGC GCCAACGCGA GCCCTGCTGCG AA

(2) INFORMATION FOR SEQ ID NO:262:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 415 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:262:

CGGCACGAGC CGTGCTACTG GTCAACTGAT GCCCTGATGT TGACCTCCCC GGGCGGGCGG
CAGTGCTTTC CAGGACCGAC GTAATTACG AAAACCAAAT CGGCCGCGGA GGGAGATG
AAAGCACAGC CAGCGGGCGA CGCGGAGGAA GACCAACAGC GATGCGCGCG CGTGGCACCC
GAGCGGGAAC AGCGGACGCG CAGCTATCGGA GACTGAAA GAAGACACCG
CCTGCGCCGT CGCTGATTCAG CGCTCGTCG TAGCTCTCCG GGGCGGAATC TAACGCCGCG
GCCCACAAACC GGAAGAAGGG GGGCGATACG AAGCCCATGC GGGCGGTGAG GGCCGAAGGC
ATGGTGCCAG CCCACTCGAC ACCACGCGTG GCCAAGCACTGC AGGTCAACAC GCCGT

(2) INFORMATION FOR SEQ ID NO:263:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 373 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:263:
TCACCGGTG AACCGTTCGT AACACTGATA CTGTAATGCTT CGTACCGGCA GATCAAGTCC
AGTCCGACCA ATGCGTCAAG ATGATCGCTG AGGTCACGG TTTTCCTGCG GACGAGACGG
TATGGAGTTC TGGGTTGAGC ATGTCGGTGG CGTGGTTGAGA AGTCCGACGC GGCATAGAGA
CATGTTGCAA TACCCGCTT GGTGGGCATAT GCTGTACTTG GTGTGCGTCG GGCCTCAGCAC
CGGATCCCTT TTGGACCCCT GCSCAGGCSC GGTCCAGTTA CGGCTGCTGT ACTGCGGGCT
GGCGTAGCGT CGCGCGGCGG ATGCGTGTGA ACGCATCGCC AATCCCAGCA ATGCGCTAGT
ACCTGACGCTT CCT

(2) INFORMATION FOR SEQ ID NO:264:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 423 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:264:

CCAAACCGGA CAGGCGGCAC GCGACCTCGG GAAGTGCGAC CACCGGTGCC GCTCCATGTA
GCCAACCCTG GACCCAGGCG TAGACAGCAAT ACCGTTGGAAT CGCGCGTGCG TGGTCAGGC
GCCGAGTAC CCGCGGGCGG AACCGCAGCG ACCCAAGGAA CGCGTGCGAT AGGGGTGACG
CCTTGGGAAA TTTTTGGCTG AAGGAGGAAA CATATCGCGG ATCAGTAGGA GTAGTACCCG
AGGGCGGCGG ACCGATCGCC CTCCCAAACA GTCGCGAGC GTGAATGCCG AATGGACCGG
CCC

(2) INFORMATION FOR SEQ ID NO:265:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 404 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:265:

AGTGCCACGC CGTGGCCTGA ATGGCATTAG CTCCCGGTGAC GTGACGGCAG CATCCGCCCA
ACTGACCGCC ACCGATCAG GCTTGCGCGC AGCGATTTCG CGCGACCGCC TATCGCCAGC
GGTGCGCCGC GTGCTCACAT CGGCGACGCG GGTTGGCGCG GGATGCCTGC CGCGGCCTGC
AGATGCTCG CGTCCGCGGC CCGCCGATCG CACCACCGCT ACCAAGGCC GTACACCCGC
CAGCCACCGCC CTGCCGAGGC TTTCCGGCGC CATGTCGGCG ACCGCGCGCG CGAACCGAGC
CATACGACG GTCGACGCTAC CGTGTCGGCG GTGCCCGCAG ACCTGACCG GAAATGCGCA
CAGGCTTCCTT ACGCCACCGC GACGGAAGCT CTCCCGGTGG GCAG

(2) INFORMATION FOR SEQ ID NO:266:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 421 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:266:

```
GTCCCTGGTGC CAGGCTGTTC TTTGGAACCG TGTCGTAATCT TCGCACCAGG GGTATCCGGCC  60
ACCATCGAAC CGGCCCAACC GGGCGCTGTA CGGCAACTTT GCACAAGACC GGGCGATGCA  120
GTGCCGACCAC TGGCCGGGCA CCCGCCGGCG GCGACTATCG ACAACGGGAC CGCCCAATTG  180
GTGGCTCTGC GCCCGGGGCC CGATTCGGCG GCACCCCGCA GCATCATGTG CTTCGATGAC  240
ATGCACGGTG CACACCGGGCT CATTTTTGCG CGGCGCCGCG GACGCCCTGT GACGACGGAC  300
GACGCAGGCA CGGCGTTCTT TGGCAGGCGC GCCGGCTACT TCGTGCGCGA CGCGCTCTCC  360
GGTCACACCG CACGAGTGAA TGTGCGTGGC GCCGCGCGCA CGGATTTCAC CGGCGATGCG  420
C                                      421
```
CGCCGCTCCG AGGGCAAGCT GGTGCTGGGC AGGCCAGATG GGCAGCTCAG CACGCTTTGCC
AGGAACCAGC AGTGACACCG CGTGCCCCGC GCCACGGTAG CC

(2) INFORMATION FOR SEQ ID NO:269:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 739 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:269:

GCTGGGGCCG ACCGCGGTCC GGGGCCCCCA GCCCCCTGCG CCAGACCCCG CCACAACCA
ACCCTGGGCG CCTAGTGCCG GGCAGCGCCG CGGCTGGCT CGT CCTCTGTT TTGGGCGCCA
TCGCGACTCT GATGCCATC GGAGCCAAAGC CGGCTGACT CCGCAGGCG CCCTCAGCTT
AGGCGCTTAG AGGGCTACTG CTGAACTCTG CAGAAGTCAG GCGGCTGTG TAGGCGCCTG
CGCGCAGCGA GGGCGCGCTG TAATCCAGCC AGGATCCCGT GTATCGCGCC ACCGCTGATC
CGGCGAGCTG AGGCGCGCTG GGTGCGCGAT CTCCGCGAT CTAGCAACAT TGCGGTGAACC
AAGCCGCTGCT GCCTGTTCCG ACCGCCGACA AAGCCCGGCG GTGTGCTGAG ACTTCCGCGG
ACAATGGAAA GAAGCCGCGT TAGGTGACG TGGAGGACCG TGAATGGGCA AAGGACCCG
GGTGACGCGT GGGCGGCTG GCAGCGCAGCG CGGCGAGAT CAGGCTGATG GACACGGC
AACCGCGCGA TGGGGCTGAC CAGATTTAC ATCAAGCGAG CCAGAAGGAC CCAAGTTTGC

(2) INFORMATION FOR SEQ ID NO:270:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:270:

AGACGTCGTC GAGGCGCGCA TCGCCCGCGC CGAAGGGCGT AACGCCGAC TGAACCCGT
GGCTATGCG

(2) INFORMATION FOR SEQ ID NO:271:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 523 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:271:
ACTGCAACCG GCAGGGCCGGA CCAACGGATC GGTTCAACTA GCACGGCGGC TGAGGGCCAC 60
CCCGGCTTCT GTGCGCTTCAG ACGAGGAAGC CTGGGGCGAG GCATTGGCAG AAGGTTTGGA 120
GGGAGATTCT GAGCAAGGATG CAGACCCGAC GTGGGGTCAC CTCAGGAGTC CGTGGTCCGC 180
CGAAGGCTGTC CGGGCCCAGC TGACCCCGAT GACGGTGGAT GTAGTGGTGA TGGAGCTGGA 240
CCGGGCGGCT CGGGCGAGCT GTGGGATGAT GGGGATGCTGG CCCGTGGGCT CGGAGTGATT 300
GGAGAAAGACT CAGATCGCGG CGGTACCACG TTGTTCTGAG AGTTCAGTTG CGGAAGTTTG 360
TGTTGCGCCGC GGCCAAAGGC CGGTGGTGGA CGCAGAGCCG TACGAGCGGT GACCCGCGGC 420
CGAGCAACCG CAGGTCAGCG AGCTGCTGCC GTGGTGCTGA GGTGAGGTGA CCGCGCGCAC 480
GTCGGCTGCG GTCCGAGAGG TGTCCTGGAC GGGCAGGCTGC CTG 523

(2) INFORMATION FOR SEQ ID NO:272:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 224 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:272:

GTGTCCGGTG GTCCCGGCTA GAGGGACTT CCCGGCGCGG CGCGCGCGCC GGACGGCGCT 60
CTGCGAGAAC CGGTGGCGCG CGCCGGCGCG GGACGGCGCA AGGGCGGATC GATCCGGCCA 120
GCCGAGTTGC GGGCCCGCTT GCCCATCGGT TTGTTGGCGGT CCGGAGCCAT ACCAGCAACC 180
AACGCTTGAAG GGCGCCCGCG TCGCTGGCTG TCGCCGCGGG CGAC 224

(2) INFORMATION FOR SEQ ID NO:273:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 521 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:273:

TGAAGCTGAC GCCCGCGCTCG ATCGGGCGCG GCCGCGCTGTC ATAGCTGCGC CGCCAGGCA 60
TGAAGCTGCTC TTGGGCCTATG GGGGCTTTGG TTCCGGCCTT GTCCAAACCC TGGAGCGGCC 120
CGTATGTTGGG TTGGTGAGGG CCGCAGCTAC GCCGCGAGGG AAATCCAGAG CGATCGGCCG 180
TGGTCCAGGG CAGATCGCGG GGTGGATGAC GCCGGCGCGC CAAAGCGGTG TAGAACGCTG 240
CGGCGAATAG GTGGCCCTCG GCAGTCAGCT GGGCGTCTCG AAGCGCGCTTCC GGGTGCGGGC 300
TGCTGCTCCGC GCAGACAGCG GCCAGCGCGG TGAAACAGGTT GAGGGGATT CGATCGCTCT 360
CGCGGTGGCG CAGGCAAGCC AGGGTGGCGG TTGGTGGCAT ACCGGCAAGT CTCTAGAGCA 420
CTGGCGAGCT CTCATCGTGG GACCAATGTT CCGGAATTTT CTCAGTGCGC TGCCGCGCGC 480
CGTTCATACC GCCGAGTGGG TGGCCACGCT AACGCGGCGGT T 521

(2) INFORMATION FOR SEQ ID NO:274:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 426 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:274:

CTCCAGGCTC ATTCGCTCGA ACAAGACCCAC CGGGCGCTAC AGCGGAGCCC CCGTATCGT 60
GTGCTGATAG TGGCGGTACA GCTGGGSCATG GGGCTGCTGA CGAACCCTGC CCCAGGCGCA 120
GCAGGAACGC CGCGCCGCCG TCAGCCGGGG TCAGAACGGT AGTCGCAAGC AGTCCTCGGC 180
CGCGAGGGGT TTAGGGGCGT AGACCTGCGG TGGCGGTCTT CGGCCGACGC GTGGATGCGC 240
CGCCAGCTGAG AGCTGACGCC CTGGAGCTCA CGGGCGAGCC GTCCGCAAAC CGAGTGCCAC 300
TCGCCGCTGTT GGGGGGCGGC CAGGACGCAG GTGAAGCTGA CGGCGGTGCG ATCCGCGGCG 360
GTGACCGCGA AGCAGCAGGAG CTGCCGCTGCC GTGCCGCTGCC GCCGGCGGGG GCCGCAACTGC 420
TCGCCA 426

(2) INFORMATION FOR SEQ ID NO:275:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 219 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:275:

CGCGACACGG CGCAGAAAGCC GCAATCGCC TGGCCGCGGG CGGGCGCGCC GCCGGGGG 60
AGGGCAGGGC CGCGCCGGGA CTTGGGGTGA CTGGCCGGGC CGCGCGAGCA CGCGGGGCAA 120
GGCGGTGGTA CGGGGGGGCC ACCGGCGGCC GTCAGCGGAC GGTGCAGGCC CGCGGGGTGGC 180
GCCGGTGCCCC TGACGCGCAA CGCGCGGGGC GCCGGCGCAC 219

(2) INFORMATION FOR SEQ ID NO:276:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 571 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:276:

AAGATCATCG GCGCGCGCTC TTAGCATCGC TGCGCTCCTG GTCGCTCGCG GGGCGGATCA 60
CGGAGGTGCG GGCTTGGTGG CCACTGCTGC AAGGTCAGCG ACCAGATGCG GGTCTTCGGG 120
ATCCTTCTTG CGCGCGCTCG GCAGGAGCTG GCTGACAGGG TCCACGAGG TGAGATCGCC 180
TGGTGCGCGG AATCAGCGATA CTTGGGGTGA CAGCAGATCA CGAGTAACAC CCGGGCGGCG 240
GTTGCCGAGGT ACCCCTGGCG CGCGCGAGTC GTTTCTCTCG TGCGCGAGGC ATCCAGATCTC 300
GCCGGTGATC TGGCGCGCCCC GCGCGAGCGC CGCGCGCCAC CGCGAAGACA CGGCCGTGGG 360
CTGGAAGCCG TTACCGAGATG AGTCGGCGGCGG CCGCGGACTC AGACCGAGCA CGACACTCAC 420
GOTGCTCGTC TTTGGGGTTG GCTACAGGAG CGAAGCAGGC GAATCGGCCG CCAGAAGCGC 480
GCACAGCTG GATCGCTAAC TACGGGGCCA A 540

571
(2) INFORMATION FOR SEQ ID NO:277:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 93 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:277:

1      5 10
Phe Leu Gly Thr Ala Pro Gly Pro Gly Gly Ala Gly Gly Ala Gly
20     25 30
Gly Leu Phe Ser Val Gly Gly Ala Gly Gly Ile Gly Leu
35     40 45
Val Gly Asn Ser Gly Ala Gly Ser Gly Gly Ser Ala Leu Leu Trp
50     55 60
Gly Asp Gly Gly Ala Gly Gly Val Gly Gly Ser Thr Thr Gly
65     70 75
Gly Ala Gly Gly Ala Gly Gly Gly Ala Ser Leu Leu Val
85     90

(2) INFORMATION FOR SEQ ID NO:278:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:278:

Met Pro Pro Val Ser Ala Asn Ala Met Val Pro Ala His Ser Thr Pro
1      5 10 15
Pro Val Ala Asn Ile Glu Val Asn Thr Pro
20     25

(2) INFORMATION FOR SEQ ID NO:279:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 26 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:279:
Lys Pro Asp Arg Pro Ala Ala Thr Val Gly Ser Cys Thr Thr Thr Val Arg
1 5 10 15
Ala Pro Cys Ser Gln Pro Val Thr Thr Ala
20 25

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(x) SEQUENCE DESCRIPTION: SEQ ID NO:280:

Trp Pro Ala Gly Arg Pro Met His Pro Ala Pro Gly Thr Ser Ala Asp
1 5 10 15
His Pro Pro Asn
20

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 140 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(x) SEQUENCE DESCRIPTION: SEQ ID NO:281:

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

(ii) MOLECULE TYPE: protein
(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 142 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:282:

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

(2) INFORMATION FOR SEQ ID NO:283:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 163 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:283:

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

(2) INFORMATION FOR SEQ ID NO:284:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 240 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(x) SEQUENCE DESCRIPTION: SEQ ID NO:284:

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

(2) INFORMATION FOR SEQ ID NO:285:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:285:

Asp Val Val Glu Ala Ala Ile Ala Arg Ala Glu Ala Val Asn Pro Ala
1 5 10 15
Leu Asn Ala Leu Ala Tyr
20

(2) INFORMATION FOR SEQ ID NO:286:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 174 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:286:

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

(2) INFORMATION FOR SEQ ID NO:287:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 61 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:287:

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

(2) INFORMATION FOR SEQ ID NO:288:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 133 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:288:

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

(2) INFORMATION FOR SEQ ID NO:289:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:289:

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

(2) INFORMATION FOR SEQ ID NO:290:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 72 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:290:

Gly His Gly Gly Gln Ser Ala Ile Gly Leu Gly Gly Gly Ala Gly Gly
  1   5   10   15
Asp Gly Gly Gln Gly Gly Ala Arg Gly Leu Trp Gly Thr Gly Gly
  20  25   30
Ala Gly Gly His Gly Gly Ala Arg Arg Trp Tyr Arg Gly Pro Thr Ala
  35  40   45
Ala Arg Ser Gly Arg His Gly Arg Gly Trp Arg Arg Trp Ala Asp
  50  55   60
Arg Gln Arg Arg Gly Arg Arg Arg
  55   70

(2) INFORMATION FOR SEQ ID NO:291:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 74 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:291:

Asp His Arg Arg Arg Ser Leu Ala Ser Leu Arg Ser Ala Ser Ser Pro
  1   5   10   15
Ala Arg Ile Thr Glu Val Arg Pro Cys Thr Pro Leu Leu Glu Arg Ser
  20  25   30
Ala Pro Gln Ser Gly Ser Arg Asp Pro Phe Arg Pro Trp Pro Ala Asp
  35  40   45
Ala Gly His Ala Arg Ser Pro Ala Trp Tyr Arg Leu Gly Ala Gly Asn
  50  55   60
Pro Ile Pro Val Arg Ala Ala His His Glu
65
70

(2) INFORMATION FOR SEQ ID NO:292:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 174 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:292:

CGCAGCTCAA CACCGTGATG TGAAGGGAGC CGCTGGTCAT GGGCGGATTC TATCGTGGG
60
CGAACGGTTA TTAGACGCCCC GGGGCCAATC CGCTGCACCAC CAAGTTGCTGAC ATCGGCCGCT
120
TTTCAAGCCA AGAAGCGCAG GACACACCAC TTAGACATTGC ACGACACGCC GGGC
174

(2) INFORMATION FOR SEQ ID NO:293:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 404 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:293:

TCGCAAACGG GGTGACGTTG CTCGGCGTG CGCTAGAGAG TTTTGCCGCAG TTTCCGGTGA
60
CCGTGGCCGC GCACGCGCCG ACCGGGAGC TCAAGTCTGC ATGGAGAGTT GCGAGCGGCT
120
CGCTTGGGCAC GTAGGGCCGC GGAAGCTGCG GCAAGCAGGT GCTGGCCGCT TTACAGGGCT
180
TGCTCAGCGGC GTGGATGGCG CGTGTCGCGG AGCGAGACAT TCTCTGGGAC GGGCCACGG
240
ATCCGCAAGCC ACCGGGGCTCT CGGATGCTGA CGAAGTCCGC CCCCGGGGTT CATAACCGCT
300
TCGGGGGAAT CGCTGGCGCA CAGCTTCACT CGGTTGGCG GTGATGGGCG GATGGTCAGC
360
TGAGTACGGG GGAAGGCAAG TATTGGGGGC ACCGGCGCTCC CACT
404

(2) INFORMATION FOR SEQ ID NO:294:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 134 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:294:

Ala Asn Gly Val Thr Phe Arg Pro Val Ala Leu Glu Ser Leu Ser His
1     5
8     10
15
Phe Pro Val Thr Val Ala Ala His Arg Ser Thr Gly Glu Leu Thr Leu
20    25
30

10
Leu Val Glu Val Leu Asp Gly Ala Leu Gly Thr Met Ala Pro Glu Ser
35 40
Leu Gly Arg Arg Val Leu Ala Val Leu Gln Arg Leu Val Ser Arg Trp
50 55 60
Asp Arg Pro Leu Arg Asp Val Asp Ile Leu Leu Asp Gly Glu His Asp
65 70 75 80
Pro Thr Ala Pro Gly Leu Pro Asp Val Thr Thr Ser Ala Pro Ala Val
85 90 95
His Thr Arg Phe Ala Glu Ile Ala Ala Ala Gln Pro Asp Ser Val Ala
100 105 110
Val Ser Trp Ala Asp Gly Gln Leu Thr Tyr Arg Glu Leu Asp Ala Leu
115 120 125
Ala Asp Arg Leu Ala Thr
130

(2) INFORMATION FOR SEQ ID NO:295:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 526 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:295:

GCTTCGACGG CTACGAGTAC CTGTTCGAGG TGTTTGTCGC GGGGCTCTAC GACGACAAG
60
CCAGAGAAGC CAGACAAGGC GTGCGCAAGG TTGTCGGCTG CGGCGGGCTG AAATACTGG
120
TGCTGGCGGG TGCCGAAAAC TGGAGGCGGC ACCTCGGCCG CGCTGCCGGG AAGAGGTCC
180
TCTGCCAGCA CTGCCAAGCAG GAGACCGTCTG GGGTTCTGCA GAGCTGCGAT
240
AGAGCGAGCT CAGCAGACGT TGAGATATCGC CCGCGCCCTG CTGGACCTG
300
AACATGCAGC AGTGGGCGGC CAAGCTTACA TTTGCGACCA CCTGTCCGGT
360
TGCTGGCGGG CAAGCCGGCC GTCCGGGGCA GTAACCCGCAG CACAGCCAGG
420
ACCCGTGGCTA CCTGAGTCCG CACACCAAGG TCTACAGGGC ACCACGGGAG
480
CCGGGGGGGG CACCTGAGCG GAGAGCCCGG GCCATGCGCG CGCAG
526

(2) INFORMATION FOR SEQ ID NO:296:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 487 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:296:

CTGCCGCGCG TAGATCTGGCC GGGAAACTTC GTCAGTGCGAT CCAGACCCCA ACGATCATCG
60
ATCAGCGGCG TGCCCATGAT CACCGCAGCG GCCACCCCGA CCGCGGGCAT CCGGCGGGAA
120
TAGACGAAA CCGGGGGTAG TGGCGGAAAGT GGGCGGCAA GAAAGACGG GGGAGCAATT
180
CCGGAAACA TCAGCCACAC ACCAGGCGCA GGGTAAAGCG TGAGTCAGAC ATCCTGGCTC
240
CGCGGTTAGC CGAGCAGCGC CGAGGGCGAG GCCAGCATCC GCCAGGACG GTGCGGAAG
300
TAGATGTAGA TGCCCGCGGT CAGCCGCGCC AGGGCAAGCT CACCCGCGCG GACCGGCGG
360
CCGGGATAGG ACAGGGCGAG CAAGGCCACG GCAACGCGCGG CCACATCGCT GGACACCTCG
AGACCGTACT GCACCCAACCT GAAGAGCTGA ACACTCGGCGG AACGTGCAAC AGCTGGGCGA
AATTTGGG

(2) INFORMATION FOR SEQ ID NO: 297:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 528 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:
ACGAAGCCCG AGAATAAGAG CCCGGGCAAC CGCGCATGTA CGAGCTTGA GTCACCAGGC
CTCACCGTGC GTGCCCGCGG GGGCTGGTCG CTCGCTGCTG TGGAGGCTTT GAAAGTTTCT
CCGAGCCCCG CCATCGGACT CGGCTGCGCG CCAGCCCACCT CAAGAGGCGG CTGAGAACAG
AGCTGGGTCG GCCCTCTCGG ATCGTGAAGC TACGTGGCTA CCCGCTCGGG CGGCGCATTA
TGACCTTCAA GACGGCAGTC TCCACCCACT CCGATGATCC TGAGCTAACG CGTGATAGC
TGCCGCGAG CAGCTGGCAC CCATTTTGCG TGGCTGGCGG TTTGAGCGCC GACCTGAGT
GGAGCCGCTT CATCACGCCG GTCGGATGCG TGGCCGAGCC CTTTGCTGTA CGCGAGAAC
ATCGGCCCTG GACCGCTGCC GATGGCGCTG CGGCCAACAC GACCGAGACG GACGGATCG
GATTCCCAA AAGCGGAGCT ATCCCGGATT TCTCAACCCGT CGATCTCC

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 610 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:
CCAGGCGCGT CAAGGAGCGG GTGCGCGCGT TGCTCCCGGT GCCGCAGCGC CCACGCTGGC
CGCGCGTGCC GCCCGGTGC CGCGATCCC GGTCCGTCTAC GTGCGCGCGC CGCGCGAGCC
TGCCCGCGCT GTTACGCGCG CCGCGGCGCC ACGCGTCTAC CCGCGTATGAG CACCGCGCA
CGCGCGTGCC GCCCGGTGCC GCCCGGCAAC GTGCGCGCGC CCGCGGCGCC ATGCGCGAAG
CCGAGCGAGT GATGGCGCCT ACCGCGACCT GCCCGGGCGT GGGGAGGCCG ACAGCGTATG
CAGCCGCGCG CAGCGGCGCT GCCGCGCGGT GATGGCGGCCATGCGGCGC CGCCGTCCGG
CCAGGCGCGT CAGCGGCGCT GCCGCGCGGT GATGGCGGCCATGCGGCGC CGCCGTCCGG

(2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 164 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:299:

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

(2) INFORMATION FOR SEQ ID NO:300:

(i) SEQUENCE CHARACTERISTICS:
1. LENGTH: 161 amino acids
2. TYPE: amino acid
3. STRANDEDNESS: single
4. TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:300:

Arg Arg Arg Asp Leu Ala Gly Glu Leu Arg Gln Cys Ile Gln Thr Pro
1, 5 10 15
Thr Ile Ile Asp Gln Ala Asp Ala His Asp His Arg Thr Gly His Gln
20 25 30
His Arg Gly His Ala Gly Ile Asp Glu Pro Pro Gly Glu Cys Arg
35 40 45
Lys Leu Gly Gly Lys Asp Gly Ala Asp Asn Ala Gln Glu His Arg
50 55 60
Gln Pro Thr His Pro Arg Gly Arg Arg Asp Val His Ile Ser Leu Pro
65 70 75 80
Arg Val Gly Asp Gly Ser Gln Ala Thr Gly Glu His Pro His Arg Thr
85 90 95
Gly Arg Lys Ile Gly Asp Asp Arg Arg Gly Gln Pro Asp Glu Arg Lys
245

Leu Thr Gln Arg Asp Thr Gly Ala Ala Ile Gly Gln Gly Glu Gln Ala
100 105 110
Thr Gly Asn Ala Gly His Ile Ala Gly His Leu Glu Thr Val Leu His
115 120 125
Gln Pro Glu Glu Leu Asn Thr Arg Arg Thr Cys Asn Ser Cys Glu Gln
130 135 140
Leu
145 150 155 160

(2) INFORMATION FOR SEQ ID NO:301:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 175 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:301:

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

(2) INFORMATION FOR SEQ ID NO:302:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 178 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:302:

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

(2) INFORMATION FOR SEQ ID NO:303:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 921 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:303:

AATTCGGCAC GARCAGCAC AACCACGCTG TCTTCACTCC CGGCGATGGT AATACGGTA  
TGCGCAACAC CGCAGCTTCC AACACCGGCA GCTTCAATCC GGGCGATTCG AACACGGGG  
ATTCAACCC ANGCAGCTAC CACACGGGCA CTGGAACACA CGCGGATTT TACACCCGCS  
GGTCAGCTCC CGGACGTACT AGCAACCGGGT CTTGATAGAT GGAATTATAC AGGGGTCAAT  
GNTGCACCC GGCTTTCGCG ATCCCTCGKG CCAAATCCAC TCCCTCAAA GCTGGCGGCGC  
GCACCTGAGG CCGGTTGAAT GATGCTATT AACCCTNAN CAAATACATT CATAACCCCTT  
TGGGCCTTCG AAACCGGTCT TGAAGGTTT TTTGCCTGAA GGANGAACTA TATCCGGAATA  
ACTGGCATAN TACAGAAAAG CGCAGCCAGTC GCCCTCCCAA CAGCTGCGGCA CGCAAGATCG  
AATGGACNCC CCKTTTACCG GSCATTACN CCGGGGTGTN GGKTTTACCC CCACGNTACC  
GCATACCTTCG CANNSSCCTN RSGGCCTGTC TCSTTTCTCT CTTCCCTCTC CCMTTCCGCG  
GGTTCCTTCG AGCTCTAAAT CGGGGNCCC TTTMGGGATC CAAATTGGC TTACNGSCCC  
CCACCCCAAA AAYMNATTG GGTAAAGTGC CTTTTMTTGGC CNTCCCTCCTA WTNANNGTTT  
TCCCCCTTNA TCCCTTATCC TCCYTATAW NTGAMNCTNT TCCCYGGA AAAMNCTCCA  
CCNTYSSSIS TTTCCCTTGG WTTATMGGR AATTSCAAT CCGYTTKG G TMAANNTTAA  
CYTATTCCNA AATTTCGGCG TTTTMNATR TTNSNCKCGM KNCTCCNRKA SSGNTTTCCT  
CCCCYTTSSG GKYCCCRGN G  

921
### INFORMATION FOR SEQ ID NO:304:

**(i) SEQUENCE CHARACTERISTICS:**
- **LENGTH:** 1082 base pairs
- **TYPE:** nucleic acid
- **STRANDEDNESS:** single
- **TOPOLOGY:** linear

**(ii) MOLECULE TYPE:** Genomic DNA

**(xi) SEQUENCE DESCRIPTION:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCGGCAC</td>
<td>GAGATANGGG CGCACCCGGG CTGCCGACCGG GCGGGACCCT GGGCCAGCAG 60</td>
</tr>
<tr>
<td>ACCGGGCTCA</td>
<td>AGACCAACAC GGTGCCCTCC ANGCCAGAGG CGCGGTGTAT GGGCCGCCAG 120</td>
</tr>
<tr>
<td>AGCCGCAAAC</td>
<td>CTCGGCTGTA CAGTCGGTGG GACTGCGCGC TGCTCCGAAAT CAGCGGCGCG 180</td>
</tr>
<tr>
<td>GGCGCTGCCG</td>
<td>CGGGGANGCC TGCGGATCCT TGCAGCTAAG AAGCAGATAC AAGCGAACCA 240</td>
</tr>
<tr>
<td>GGAGTCCGNAA</td>
<td>CNGGCGGCAG GGTGCGCTGCT GCGCGCGCTG GCGCGGCGGT GCGGCGGTGG 300</td>
</tr>
<tr>
<td>GGGCCGGCAAC</td>
<td>CGGAAACCCG NGAACACGGG AAGCGATATC AAGCGAACAG CAATTTGTTCA 420</td>
</tr>
<tr>
<td>GGCTTAAAGC</td>
<td>CTTCACTACCG AGGGAGTCCG GCGGCGGACA CGGGGCCGTC GCGGGCG CG 480</td>
</tr>
<tr>
<td>CCCNTCTCCTN</td>
<td>GGCGGGGACG TCTCAAAAGG TCGCGATGCA CAGTTGATGG CTTCCGGCGA 540</td>
</tr>
<tr>
<td>TAGTSGAAGGNN</td>
<td>CCGCAACGGNT TTAACCGGCC SAAAAAATCT TCCACTGAGA TAAAGATACG 600</td>
</tr>
<tr>
<td>CCCGGANCCC</td>
<td>CCCCCGSGSM NGTGYGGKCG ATTTNTTCAAC MGGTTTNAGC GGGKGCCNG 660</td>
</tr>
<tr>
<td>GCCAATTGCG</td>
<td>CAAAMTTAAG KNTGGGGNTY CGGGGCGGA ACGGGCINTK NCGCCCTTAA 720</td>
</tr>
<tr>
<td>AAAACCSCGNC</td>
<td>YTTTTCTGAT TAMMACGGGN CCCCCGWTTG GGCGTGKTCC CANNTYAAAC 780</td>
</tr>
<tr>
<td>ACMCCCSSS</td>
<td>MnGSGGTGCGA SAAACCTTCC CGNGGCGGTG NTGTTTSCYT AMWCCCGCGG 840</td>
</tr>
<tr>
<td>AAACCGGKXJG</td>
<td>GGKGGCGRTN WASSAMNNCC CMNGGYTTCT TAAAGGCSCAN KNGAANGKYT 900</td>
</tr>
<tr>
<td>CTTTGGGAAW</td>
<td>CCTTCAATYC AAAAYICTC CTYMMGSGCN CTTKCMRTTN NNGGGAACS 960</td>
</tr>
<tr>
<td>AMWTNYCNC</td>
<td>GWTTCAWTCG GTGCCGAMN AAACXCTTTT TTTTTCGSSC STCCMGGSGNC 1020</td>
</tr>
<tr>
<td>SGTTXKXANAN</td>
<td>AAASATTTCM YCCHNNWXX YYCCSSGCCT CYMGRRRNKR SGMAACCGR 1080</td>
</tr>
<tr>
<td>GS</td>
<td>1082</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO: 306:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 223 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:

<table>
<thead>
<tr>
<th>AATTCCGGGTG</th>
<th>GCAACGCAGG</th>
<th>CCTGTTCCGC</th>
<th>AAGC GGCGCGG</th>
<th>CGGGCGTGGTG</th>
<th>CGGGGGCTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGTGGGGGCGG</td>
<td>CCGGGCGCGG</td>
<td>GGGCGGTACCC</td>
<td>GGGGGGTTGG</td>
<td>TTGGCTATGG</td>
<td>GGGCGTGGG</td>
</tr>
<tr>
<td>GGCGTGGGGGG</td>
<td>GTGGATGCTG</td>
<td>GGGCGGGGCC</td>
<td>AACGCGTGTA</td>
<td>GGC CGCGTCA</td>
<td>GATGGGGGG</td>
</tr>
<tr>
<td>GGGTTGTTGG</td>
<td>CCGGGGCGGA</td>
<td>CRACRCTGGT</td>
<td>GGGCGTGGTG</td>
<td>CCGG</td>
<td></td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 418 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

<table>
<thead>
<tr>
<th>AATTCGCACG</th>
<th>GANGCGGCCA</th>
<th>CGGTGGGCAGC</th>
<th>GCACCGCGAGGC</th>
<th>GGCGGCAAGG</th>
<th>GGATGGCATC</th>
<th>GGGGGGCGG</th>
</tr>
</thead>
<tbody>
<tr>
<td>GGGAGCGGCG</td>
<td>GGCGCGCCGG</td>
<td>GGCGCGCCGG</td>
<td>GGCGGCGCGG</td>
<td>TGATCGCACA</td>
<td>GGCGCGCAA</td>
<td>120</td>
</tr>
<tr>
<td>GGC GGCGTGG</td>
<td>GGCAATGGG</td>
<td>CATTGCCCAG</td>
<td>GGCGGCGAGG</td>
<td>GGTGNCGGG</td>
<td>CATC RSTGGG</td>
<td>180</td>
</tr>
<tr>
<td>CTGGTTGTTG</td>
<td>GTTGGACRGG</td>
<td>GGCAGGCAGG</td>
<td>GCAGAGCCAG</td>
<td>GAACAGGGG</td>
<td>GAACAGGGG</td>
<td>240</td>
</tr>
<tr>
<td>CAGCAGCAGC</td>
<td>GTGGGGCCAG</td>
<td>GGGAGCAAGG</td>
<td>GGATACGGA</td>
<td>GGCGGCTGAT</td>
<td>GGCGGCTGAT</td>
<td>300</td>
</tr>
<tr>
<td>CGGCAAGCGG</td>
<td>CGGACGGCGG</td>
<td>CGCGCGGCGC</td>
<td>GGGGGGGCGG</td>
<td>GGGTCGCGG</td>
<td>GGGGGGTGTG</td>
<td>360</td>
</tr>
<tr>
<td>TTCGGCGGCG</td>
<td>GAAGGACCGG</td>
<td>CGGGTCCCGG</td>
<td>GTCAACRGGG</td>
<td>GGGGGGCGG</td>
<td>AAATGCGG</td>
<td>418</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO: 308:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1049 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:308:

AATTCGGCAG GAGGGGCGAC ATGGCATACA GGGTGTGGCC CAGACCCGCC CAGATACGCA  
GCTGCGCAAC CGCGAGCAGCA GATACGCCGG TCGTTGTCG GCAGGTTGAR CACCACCCTGG  
TACCAGGCGC CAGCGGCGCG ACGAATCTAG ACAGTCTGGG TAGTCCACCG  
GGCAGATCATC CCCCACCAGC CTTTGCTATT GACATCACAC GTCGTTCTCT CATTCTCCCG  
GCGACACGGCC CAGTGCTTACG CAGAGCCTAC CAGCGTCCGC CCGGATCGCC CTTGGCCCGC  
TCGCGCTTCC GAGAAGTCCCA TGAAGAAGTCGGC CAGGTTTCTG ATNAACAGAT GCAGAATAGC  
CCTCGGATACG CCTCAGCGCT CAGTACGCCG AACTGCGCAT GCTGCTGTCG GCGCCTCGAN  
CTTCCGGAAG TGCCTCGGCTG ACTCCGGCGG NGCTGCCAAC GGAATCNTAT CAGCATGTCC  
ACACSANNCA ACTCTGGCAA GGTAGGAGT TCCCTCCNCT TCCATATCCC GGGAGTTRCC  
CMATTCCTGCG CKNATCGKGC MCAAGMCYCT CKAANAAKCG GGCCTCCTCT NTCNGKGA  
CCAAAMGGKT TGGGCGGAAG GGNMNNCAAM CTTWACCCCTG KTKAAAGNGN TTCGGCGCGC  
GGGAGAKGNGA ATYYCCCSMN ACGCGGCGN GNMARATTC TCCGCGMCTC CTGGKGAWT  
GGGCGGGGCG AAAAAAGCGR CCGGGGCTCN TTTTCTCCRN RTRGGANACN CGIYCTGARCA  
CNSCGCGGAG CCGCENNTT TCCNNMTWCTM WGYTNCRMCC MMSNCKSNG KKGNNRCCNN  
CRGGCNRECM AAINTKGYN KNTATMAGC  

(2) INFORMATION FOR SEQ ID NO:308:

(ix) SEQUENCE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:309:

AATTCGGCAG GAGGGGCGAC ATGGCATACA GGGTGTGGCC CAGACCCGCC CAGATACGCA  
CCACGATCCCA GGGTGCAGCG CAGACCCGCC GAGATAGCGG GGGTCGGGAA GAAAGGCTG  
AGGCTGCGGA TGGDAAACTC GAGGCTTACT GAGGTTGGAA GAGGGACAGT GATATATTT  
GGGCAGATCG ANGGTCTGTG ATGCGGAGA GGCTGAGGGC GACCTTGGG ATTTGGGGA  
TGTTGTGGCA CAGAKCGGAC CGCAGGCGCT GATGGTTGCT TSSANTTCTT GTCGGCGSCCA  
CANGGTTAGT GGCAGTCTGG TCACTGCGGG SATCGAAATA TGTTGGTAGC CNAAGCCGAA  
CCAGATYGCC GGGACCTTCA TGCGGCGGAC AACCASMSGT GCSANGTAKA GGTTCCTTT  
ATNTTGATGCG GATGTTCCGC ACGTMTSTGC GAGGTCTCAY MTGATATATC MTSATSGGCA  
YATTATCNGS CCTNAYGGGA ATBAMRRGAA CAAAYATCTCT CCGGGGGAAC ACCAACMSGC  
CTTGGTNSYC CNCCNCCENC AKAAACTCRT KTGTRSTMC CSMAAATWA CSCCSSCCTTS  
NACTCCNGSG CAAATNSCCGC CCCSCKNNTT ATSSYCCCGC GTTCCGGCMMC CCGTTTNAAMC  
TCCCGGTGA ACCCCWNTT SNDCGCGCC TAYAKMNCRG GTCTSTTNTCT CCCCGYTRMK  
CNCCCGCTCK SAMCWNCCNC CTTCAKANAC CCGCKVKGSM TNCCCAATNT WCMMNCKCNS  
KTNTNTMCTC CCGATNYCC CCGNCTTCCGC CCKSTTSCAM WTATTAACCC MCGWYWNK  
KCNCWMAWTA MGACWCTCNY CCCGNNCNCK NTTKTMWCC KMMCCCCKCSW TWCYCCSAC  
CCMCTMT MACCCYCCCK NKMCNMTTNC CCGCGCGCNM CMMBMKCYTCT YGCGKXWCKNC  
NCYNNTTMCCN CYNANMKCKX KTCCTCTCCN CRNCTCTCCGC CCWCCCWCCV KKTCTCSKKC  
CCNCTTCSC MCMGSC  

(2) INFORMATION FOR SEQ ID NO:310:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1036 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:310:

AATTCGGCAC GAGATCATGA ATAGCGGGCT GGTCAGCACG CACGTGTCG CAGATCTCGC 60
GACGCACTTG CGTGCTGCTCG GCCACCGAGG CGTCGGACTG CATGCGAACA CCTGATGATG 120
CGTGGGATGT GGTGACTGTC AGTGAAGGCG AGGACCCGAC CTTGCTAGCT ACGGTGTCCT 180
GGCACTTGGCC AGTCGAGAGT ACACAAATAT GCCAGCAGA CGAATCGCGG CCGCTGATCG 240
GGTGACCAGG CTGGCCCGGA GGCCGCGGTT GTGCCCCTCG GCCGGCGCAG CAGCTGCTTG 300
GGCAGATGAAG TTAGAAGGCG GCCTACCGCC GTGCGGAACT AGCGATGCAC GAGCGAAGC 360
ACAGATGTCG TGCTCTGCGT GGGGTAATGC GANCCCGAGA AGCAGGAGAG CACAACTACT 420
TTGAGCTTCG GCGACTGACC GACCAAAGCG CTGCGGCAAC CCGAGCGGCA ATGGTGGTCG 480
TTCCGGCGGG CCGGNAACGG AATACNGGAG AGCGCGCTGG GCGAAGAAC CAGTANCCNT 540
ACATANCAACG GGNNTCTCGGC CCCACATTTG GGSTTMTGCG CCTGCGCAAC CSSNAAYNC 600
CCCAATTCGAG AACAAAAAAA TTTGGCCATY ARNGTYCTCM CAAAAAAACN AMTCCCCCTA 660
TCCCCGGGCC GGCGGCCCGY NMMAAAAAGG CCCCCWANCC CCGSSCGCC CGGGTTRMN 720
CCCTCCCCCG TGGGCGCCTG GTTGGTCMWC GGGCMNWNW GGGGNTGCCG CCCCCNAA 780
AAAAAAYCNN NGAAAAATYYA CCKKYMMAA ASKYGGGGSC CCMARCCGG GCGAAKXWAA 840
ANNTAANCCN KAATAAAAAW NCAANMCCCC NGGNGNCTAA GGKTYTANNG GTTSTTNANG 900
ARAAAAATMCT CANIAMNNSK TTNNAAAAAA ASCSWAKCC CCCNNKNNKNN CCAAWAAR 960
SRCCTTCGGG TNNWSGGGGG KKKKTNCMC KMNMTTTWGR CCCCNGCCCN NNTWKCCTTN 1020
TCCNYGGNGC RNCAGN 1036

(2) INFORMATION FOR SEQ ID NO:311:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1060 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:311:

AATTCGGCAC GAGTCAATTC GATCGAACAC GCCCGCACCT GCAGGCGCAA CATGGGCGCG 60
GCCATGCGCA ACAGCTACTC GCCAACCAGG AATTCATTGC GCGCTCTACC GCACCCCGCC 120
AAACCGGCGA CGCGGCGAGT GATCAACCAG GCCACCCGG ATCCGGAATAA GGGTCCACAT 180
AATGAGACAC TGCGCGAAGAG AGTGGACAG GGCGCGCAAC AGCGAAGCTG TTAGACGTTG 240
CGTGATGCAG AGAAAGGCGG TGGCCACCCA AGATACACGG GCCAACAGGG ATCGATCAGA 300
GGTGGCAGCT GTGTCGCGCT AACGTGCGGC GCACGAGAAG ATGACGCGTG GCTACCATC 360
GGCAGCTGCT GTACACTCTTG GCCAATGCGG CACCGAANAT ATGANCACG TTAGCCGAG 420
TCTGCGTGGCC ATCGCGCGAG GCCTGAAAATA CTCTTGTGG AGAAATGGAG TTGAGTGAAC 480
CCGAAATCTT NTRCTTGGCA SSAACATNCA GTGTNCGSGT NAAACAYCTT GGGTNGAA 540
ACANCAATAT TAAGANCCCT ANTCGGGCGA AAACCGACGG CAAAAGGAAA TGGGAAACGA 600
TGKGTCCGCAA AAAATCCCCG NGGTAAAAWW CCCCNATNCG MSAATTTTSC CTNGAACAAM 660
AAAAAGTCGCA AGKYGAACAGG GCGCCCGCCC SGNAATTTG TGACGSCAAY WYNARTCTC 720
WWNWNCAATT MTTNGGCTC KNTTCGCCAGT KNAANGGCCCG CCCCNTCRRG GCGTYYCCTC 780
NNNINNGGGNG CYCCCGCGCA AAAAANMMM MTTTCGKGK SMGGKKCCCG CCCGGTYWG 840
GKXVTTAAAC CGGGGGGTGT AAAAANAN ACCCCCGAMS NGGGGGAAA AATGGNAAWT 900
(2) INFORMATION FOR SEQ ID NO: 312:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1040 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:

AAATTCGGCAG CAGATCTCAC CAAAGGCTGTA CATGCGGCTG GATGCGCATC CGCATCGAG
60
GCAAATCGGG CAGATCATGA CGGAAAGGAN TCTGCGGTTG GCTCAAATGG TAGATCGTTC
120
CCAAAGGTCA AGCGTTTGGG GGCAAGAATG CGACGTCTAA TTGCGGTTTC CACCTGCAGA
180
TGTTGTAAGT GATGCCTGGAA CGGCGCTGAC NGATAANGAA TTGGCTGTTG CGCGCGCAGN
240
ATTGATGCTG CAGTTTTTCC NCCCGCGGTT AATTCGTGTA CGATCATCTG GCCGCTATG
300
TTCCGGCTAC RCTCCAGGCC ATCAGTGTA TCGCGGCTAAG GAAAGCTTAG TGAACACGGGC
360
CGAATGCGGC GCTGGAAGAC NGCCGGCTGT TGATTATACGA CNAACCCCTN CNATCTCTTG
420
CCGGNMTGNC GTGTCCTCGCC AAGCGGCGKG SYTTGACNCCG TAAACCCCA CTNCCCGCGC
480
GGTGGACTTA MRNNTTCGAA AAACMGMTNA AGCGGAATNN SAAACTNCCR TCAAANNAMM
540
SAAATCGGCG TTGCGNRCGC CCCGCNAAYW TKCNCNGCGG GMNTYCTCTN GTTYNGGGGC
600
SAAACNNTGG CRTNCMVNNN TTTCAGAGCC NCMTMNNTGM GGSSCNSNAS GWCCCGGGKK
660
TTTNTTCAAW TCNNSKTTTT TTGGGGGCGG GCCGCGTRMC NCGGGCGCGC GCCCGCNGMA
720
AAAAAMCSMS BRCNCGYGGG KKCCCCCMMC BNATNGGGCG YKCRAAACAAC ACCCGAAHRA
780
TNGNMGCGGC SMCSCGGNN GYNAAAKGGT TSNSTMANM MKNMMNINCT SMNCCCMNNSN
840
NGAAGAAGMN CRRNGANRCC NCNGYGNRNN NGNNNNNNNNN GGRKNNNACN NNMNCWNSMC
900
NSNNMGNNNNS CTGTCNKCSC
960

(2) INFORMATION FOR SEQ ID NO: 313:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 348 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:

AAATTCGGCAG GAGCAAAAGG GTGGAATGG GATCGCGCGG AGCTGGGGCC CGTCGTACGC
60
GACTCTTGGG CCAAGTGGCG GCGCGCGCTT CGGCTCTATG GGGCTGTATT ACTGCGCGCC
120
AGCTTGAACGT GCGCGGCGTA TCTGCGGCTT TCTGCGGCTC TTGCCCGCAG TCGCGCGGAG
180
TGGGGGGAAGG TCGGGGCGGA CAAACCGGAG TCGGATGCTG GCGGCGGCTG CAGGACTGGT
240
ANTGCTGTAA CTTGCGAACG CGCGCGGACAA CTGGCGNCAAC TATGCGCGCG CTGCCCTGTC
300
NGTCCGCTCC CGCGCGGCTA GTGCCGACGC TCTCTTCTTC TCTGTGCGC
360

(2) INFORMATION FOR SEQ ID NO: 314:
(i) **SEQUENCE CHARACTERISTICS:**

(A) **LENGTH:** 332 base pairs

(B) **TYPE:** nucleic acid

(C) **STRAINEDNESS:** single

(D) **TOPOLOGY:** linear

(ii) **MOLECULE TYPE:** Genomic DNA

(xii) **SEQUENCE DESCRIPTION:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCGGCAC GAGAACCCGG GTCGTTGACC AACGGACGGT TGGGCGCGGG CCCCCCGGGT</td>
<td>60</td>
</tr>
<tr>
<td>GCCATCAAGCC CTCTCTCCCT CTTAGCCGCTT AACGGCGCTG TGGGTTGTGG GATCTCTTGA</td>
<td>120</td>
</tr>
<tr>
<td>CCACCCGGCTG ATACCAAGGAG GCGGAGAGTC TCTGCGTACC CAACGGGGAG AGTCTCTTCA</td>
<td>180</td>
</tr>
<tr>
<td>CCCGCCCGGCG GCGACCGAGC ACCATCAAGT GTCTCCGAGT GTGSGGCTCGC TGCCGGGAAA</td>
<td>240</td>
</tr>
<tr>
<td>TAGACCGGCTG TGCTCTCCAGT TGACTGCTGA CTGATCCCGG GCAACCTCGGA GAAGCGCCCA</td>
<td>300</td>
</tr>
<tr>
<td>GTCTCGCTTG GTCCTGCTG GC</td>
<td>332</td>
</tr>
</tbody>
</table>

(2) **INFORMATION FOR SEQ ID NO:314:**

(i) **SEQUENCE CHARACTERISTICS:**

(A) **LENGTH:** 962 base pairs

(B) **TYPE:** nucleic acid

(C) **STRAINEDNESS:** single

(D) **TOPOLOGY:** linear

(ii) **MOLECULE TYPE:** Genomic DNA

(xii) **SEQUENCE DESCRIPTION:**

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCGGCAC RAGTCGCTCT AGACGGATCC AATGCTCCCCG CGAGACCTCCG GCCACCTGCA</td>
<td>60</td>
</tr>
<tr>
<td>ACCCGCTAGC CAAAATGTGTC CAAAATGTGTC AACAGGACCTT TCCAGACGCT CACCGCCGCCG</td>
<td>120</td>
</tr>
<tr>
<td>CCAATGATGC AGCCGCCGAG CAACGGAGCT CTTGGAGCGG ACGCTGGCGG GGGCCGGCCT</td>
<td>180</td>
</tr>
<tr>
<td>GATCCGCTCTG CTGACGCGCG ACGCCGCGGG ACCGGCCGAG TCGCTCGGGA AGGCTCTGCA</td>
<td>240</td>
</tr>
<tr>
<td>AAGGGCCGAG GACGGCGCGG GACGGCAGCG AGCTGGGCCT GCCTGAGCTC GCTGGCAGG</td>
<td>300</td>
</tr>
<tr>
<td>GCCACACGGG CACGCGCGAG GAGCGAGCGA CAGCTGGGGC CTGCTGGCTC GCTGGCAGG</td>
<td>360</td>
</tr>
<tr>
<td>GTNACGCGAG CAGCGCGAG GACGGCGCGA GACGGCGCGA GACGGCGCGA GCCTGAGGAG</td>
<td>420</td>
</tr>
<tr>
<td>AGTATGCTCTG CGCCCGCGCGA AAAACCGCGG GCCCGCCGCGA ACATTACCG GCGCGCGCAGG</td>
<td>480</td>
</tr>
<tr>
<td>GCGGCCGCGG CGGAACGGCG GGVNTCTCCC AACGCGCGGG CGCGCGGCGA GNNMGSGTGCT</td>
<td>540</td>
</tr>
<tr>
<td>CTTNGGSSGA AGGGCGCAAK CTTCGGCTAN YYATCCCGGG ANGGCTGAMC CTTATGSNCA</td>
<td>600</td>
</tr>
<tr>
<td>MYTTMAGGA YCTNCCANNT KTTSSGACCC CCGNGCGGAA ASRAWNNGTG KGGCAACNA</td>
<td>660</td>
</tr>
<tr>
<td>NNTNNCTTTGN NATTGGGNNAA AAAAACTTGY CCWCGRACT NCCCGCCGAGG GRGMCNNTNN</td>
<td>720</td>
</tr>
<tr>
<td>NTPTTNYGNNN CGGCGSNANAM NNTTTAATTC NGGSNGGTCCN GGGTTMNAGA AACCCCAAM</td>
<td>780</td>
</tr>
<tr>
<td>MRRNNKCCSCA ANGGKSNCG NGNNMNNGST TTYCTNNMRA MRRNWTYNKN NTCCNARSRN</td>
<td>840</td>
</tr>
<tr>
<td>NAAACNNSNK NGKKTNNNAAAA ARNNTWTNKN KNSCNNNNNNM GRNNGVNGCG CSHKSGSNMNG</td>
<td>900</td>
</tr>
<tr>
<td>MCIWHAWNRN NNGSNCNXKC NNNMNAAAA AASGGYGCNS NSMHEKKEKKG NRRGSGGGGG</td>
<td>960</td>
</tr>
<tr>
<td>GG</td>
<td>962</td>
</tr>
</tbody>
</table>
(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:316:

AAATCGGCAAC RAGAAAGGC CGCAAGGTTT GGCTGGGCTC TACAAGCTCA TCAARGCGCA 60
GGGGGAACGC AACTTGCCGA AGACCTCACG TTGCTTCCCC CAAACCGGCTC CGATGGGCCA 120
GTACCCGGCC GCACCCGACT GCCAGGGCAG CCAAGATCCG GCGCCGGAAC GCCCGGGTTT 180
GCAGAAGATG TGCGCTGTGG AGGGCTTGAC CAGTTTGCAAG CCAAGCCGCG TGACCAGACC 240
GGGTTTCTTG TGGCGAATGC TGCTACACAC CCGGCAGCACC CGGTGGACCT CGACCAGCCTG 300
CACCACCTGT GCCTGCTGGTC CGG 323

(2) INFORMATION FOR SEQ ID NO:317:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1034 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:317:

AAATCGGCACT GTGTGTGGCG GCGTCCAGAA GAAGATGATC GCGAACATCG CCAAGGCGCGG 60
CCAGGGCTAG GTGCGCGTGA TGCGCCGACA CGCGCATCT ACCGGCGATAC AGCGCGCCCGC 120
CCACCGCCAC ACCACGGTCG CTGACGCTGC TGCTCGACGC CAAAGGCGTC AGACACACAC 180
ATAAAACCCG CGTGTACGAC CGCGCCACGG CCGTGTACCC GCGACCCCGT GCACCATGTA 240
CCAGAAAGAC GAGATGACCG CTCAAGGCTCA GCGATCGGCC AGCGTTTGCG GGTGGGCGACC 300
GCTTCCGGC CGAAAGGGCG CCGCGCGTTT GCCTGCTACG CTTTGTGCACT ATCCGGGTGC 360
GCACCCAGTT GAGGGGTTTG GCGCGCGCGG CGCCACATCA CGCCCGCGAC ANCGGTGTGA 420
GCATGACGAC CGGATGAAAT GCGCCCGGCG TGCAGGGCTG TCTGGGGATAT TCAAGCTGGG 480
CNACACCTTG GCGGCAGCCT GAAACCGCGG TGAATGAWTG AATTAAACCC GSTSAAACANT 540
AACTACATAA CCTTGGGGGG CTCCTTACCC GTTGYTGAAG GGGTTTTGGC TTTAAGGAAG 600
AACAYATTTG CAGTANCTCGT CTTTNTWTCG AAAAAAGGCC CRCCCATNCGG CTCCTACGT 660
TTSSCCCTGA ATGGAATGTT MNCNNCCYKRN CNGGGNCTTT AACRCGSGCG GGNTTTTGKT 720
MCCSNNTCAK CNTTMMTGC ABRNNCNGGC SKCCTTCTCC TNYCTCCCTC NTCCCNNNST 780
TNCGNKTGCC CNAMNMYTNW AGCCGGGCGC YTNRGCGCRM WTGKKTGGTG GCCCCCMCCCC 840
NAAANAAANAN GGGKRGNTY CTTTGGCCNC CCAMAAARGTC NYYCCCGCCAM YTNNRRCMSY 900
CNMTNIGGNM CTTGNCKNCG GAARANMCAC KCCCGCGGNSTS STINGTYWAG GNWKGNSRSG 960
CCSCGGGCGY MNNHAAYAWN WSNATNCNNS STNA MANAKKNN NNNNNNNSCN WNGNGNNTCN 1020
SCNSNGGKES CSCC 1034

(2) INFORMATION FOR SEQ ID NO:318:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 331 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:318:
<table>
<thead>
<tr>
<th>Sequence ID: SEQ ID No:319</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SEQUENCE CHARACTERISTICS:</strong></td>
</tr>
<tr>
<td>(A) LENGTH: 1026 base pairs</td>
</tr>
<tr>
<td>(B) TYPE: nucleic acid</td>
</tr>
<tr>
<td>(C) STRANDEDNESS: single</td>
</tr>
<tr>
<td>(D) TOPOLOGY: linear</td>
</tr>
<tr>
<td><strong>MOLECULE TYPE:</strong> Genomic DNA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequence Description: SEQ ID No:319</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCCGAC GAGACCATG CCGGCCGCCG TCGTGTGTCTG ACTGTGTGCC CAGAGTCGAC 60</td>
</tr>
<tr>
<td>RAGCCACAGT CGCTGGAAGCG CAAAGTGTGTG CTGACGGTGCT TGTCCCGGTT GGGACGGCGT 120</td>
</tr>
<tr>
<td>TCGCGCGTGG TGCGGACCCG GAGATCGGCCG CACGGCGGCC GCTGGCGCAGC 180</td>
</tr>
<tr>
<td>GACGGGTGCG CGCGGTGATTG GAGGACGCCTC AAGGCGATCCGG TTGGCTGGCG CACATGCAC 240</td>
</tr>
<tr>
<td>TGCTGCGCCAG TGAGCGCTCGC CGCATCGACG CGCGTCGACC GAGATGCTCGC ANGAGATCAC 300</td>
</tr>
<tr>
<td>CCGGCGGGCC TGAGTGGGGCC TCCGGGCGAC A 331</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:320:

<table>
<thead>
<tr>
<th>Sequence Characteristics: SEQ ID No:320</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) LENGTH: 324 base pairs</td>
</tr>
<tr>
<td>(B) TYPE: nucleic acid</td>
</tr>
<tr>
<td>(C) STRANDEDNESS: single</td>
</tr>
<tr>
<td>(D) TOPOLOGY: linear</td>
</tr>
<tr>
<td><strong>MOLECULE TYPE:</strong> Genomic DNA</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sequence Description: SEQ ID No:320</th>
</tr>
</thead>
<tbody>
<tr>
<td>AATTCCGAC GAGAAGAGCCG CCGARNGTST GCCGCTGCTGT TCAACTCCTCA TCAARGGGCA 60</td>
</tr>
<tr>
<td>NGGGAACGCC AACCTCAGCA GATGCTACGT TCGCTTCCGG GAGAGGTGTG CGATGGCCGA 120</td>
</tr>
<tr>
<td>GTACCTCGCC CGCAGGCCGG CAGAGTCGGC CGAGATGGCC GCAGCGGCGAC GCTGGCGCAGT 180</td>
</tr>
<tr>
<td>GCAGAGAAGTG CGCTGTCGGG CTGCTGGGAC GATTTGCAG CCGACGCGGC TCAACCAGAC 240</td>
</tr>
<tr>
<td>GGGTCTGCTG TCGCCTGTCCG ACGCTACCAC CGCAGCGGCGAC GCTGGCGGCG CAGAGACGTC 300</td>
</tr>
</tbody>
</table>
GCACACTGCTGCGCTGCTGCTGCC

(2) INFORMATION FOR SEQ ID NO:321:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1010 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:321:

AATTCGGCAC GANCCGCTGCC GTCTNAACACC AGGCCGGCCG TGCCGACATAT CCCGGACTCG
GTCATGGGCT GCCATGGGCTG CTGGCTCTCC TGGCAGGGGCG CGGGCAGCAT AAGGTCCGCTM
ATGCCAGGGT AGGCCGCCAGG GTGCACTGAG TGAGTAAGTG TGCCACCTTC CAGCTCGGCC
ACCAGGAGCT CTTGCTCAGC CGTGATGATGG CAGGACGGCC AGAAGAAGTG GATGAATATG
ATAGGTCGCT TCTAGCTGCT GCCGATGGAT GGGGACCGCA SGAAGGCTGAC CACTGAGCAC
CTACCTCGCG GCCGCTGCTC GTAGCCGCTC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
GAAGCGGCTT CTTAGTGGCT CTTGCGGCTT GCAGGAAATG GCTAGGATTC GCTTCCGTAT
ATATCTCGG CCAAGGGGCTG CGGGGCGGCTG ACCCATATG AAGAAGGAGG TAGGCCGTAGG
CTCCCGGCTG GCCGAGCTCG ACCGCGCTGCA CCGTCAGAGC TGGCAGGGGG CACTGAGCCT
WTAAATCGGA AACCCCTTCC CCGCCCTTGG AGGCTTGGTG TTTTCTTATT TTTCTCTTTT
AAATTTCGGC GAGGGGGTCT CTTGCTCAGC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
AGCTTCTCTT SYTGAGGCGG GAGAAAACAC TTNTYTYYYN ACCSCNNAA MYMTTTCG
MNAACNCTT CACCTGCTGG CTCGTCGCCG GCCGCTGCTC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
GCCCGGCGGT GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG

(2) INFORMATION FOR SEQ ID NO:322:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1010 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:322:

AATTCGGCAC GANCCGCTGCC GTCTNAACACC AGGCCGGCCG TGCCGACATAT CCCGGACTCG
GTCATGGGCT GCCATGGGCTG CTGGCTCTCC TGGCAGGGGCG CGGGCAGCAT AAGGTCCGCTM
ATGCCAGGGT AGGCCGCCAGG GTGCACTGAG TGAGTAAGTG TGCCACCTTC CAGCTCGGCC
ACCAGGAGCT CTTGCTCAGC CGTGATGATGG CAGGACGGCC AGAAGAAGTG GATGAATATG
ATAGGTCGCT TCTAGCTGCT GCCGATGGAT GGGGACCGCA SGAAGGCTGAC CACTGAGCAC
CTACCTCGCG GCCGCTGCTC GTAGCCGCTC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
GAAGCGGCTT CTTAGTGGCT CTTGCGGCTT GCAGGAAATG GCTAGGATTC GCTTCCGTAT
ATATCTCGG CCAAGGGGCTG CGGGGCGGCTG ACCCATATG AAGAAGGAGG TAGGCCGTAGG
CTCCCGGCTG GCCGAGCTCG ACCGCGCTGCA CCGTCAGAGC TGGCAGGGGG CACTGAGCCT
WTAAATCGGA AACCCCTTCC CCGCCCTTGG AGGCTTGGTG TTTTCTTATT TTTCTCTTTT
AAATTTCGGC GAGGGGGTCT CTTGCTCAGC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
AGCTTCTCTT SYTGAGGCGG GAGAAAACAC TTNTYTYYYN ACCSCNNAA MYMTTTCG
MNAACNCTT CACCTGCTGG CTCGTCGCCG GCCGCTGCTC CTAGGAGGCT GCCGGAGGCG ACATGCTACC
GCCCGGCGGT GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG GCCGGCTGCTG

255
ATATTCCAGG CACCCAGGCC AGTTTGGTGG ACAATGCCCC GGCACATTCC TCRAAATTCG 300
TGAACACGG AAACCCCTCA ACCCCCGCGA CACCCSCCNC AACARATGGG WTCGCGCGTT 360
CTCCCGACGG KTTCTCGGGG GNTTGCGCAG AANGCCACCC WTGGWTCTCT TCCCGCCACC 420
GGCCGCCACA NTGGGCTGAT AAATTGTGGA AYGGGGCGGC GGGACCCSCA AAGGGTGCCG 480
GAAACTGTTT YCRAAAMCG CAAKCGCGGA TCTCGCGCGR ANAAATTGCT YGGCACACT 540
GCTTTRTACTT CCCCCAGCGT AAMNATTTT ATGCTGCTNN CTCCTGCGCT TGGGGCAGGG 600
CKAAAYACCC GMTTGGTGGTT GGGGAACTGC GGGCCCAATNC CCAAMCCRCA CTTTTNMAT 660
GGNTGCGAAT TSCCCCGCGT RANAACCCGC NTGGCCTNNT CTGAAASSA NNGGCGCTNT 720
KGCCNSSCCGG ATTAANACCCT TACNNAYTAC CAWCTCTGCT CAAASTTGCG AGCAANNKST 780
GGNTTCCGAGK AYTYYTXTGG TNGNCCNCTN TATNGGCTGN TGGCGCCNCGC NCTSTKSGCA 840
NASGGCNGNG NGNNGGGGTTT ACCCCCTGGM GGGGGTTTTT TSTGGAGMCS TGGACTTCCTR 900
GGCGCCNNGG GNAKAATGTTT MWMTMNCNSG GGGGAWTITT NTSTGGAGMCS TGGACTTCCTR 960
GGGGCTTTCTT TCCCCCNCNA NNNAWNGGGG GGGGGANAYT TNGNNSGNGG GWNNTTATT 1020
YTTYCSCNCT TKAACCGSGT QTTTKNANQ GGGGGAGAAA ANAANAATAA RAMGGYNTNTT 1080
TSKNNCACNCT GKNWNNWANR NAGAGCTTCTG CKCRKNCRCG SNNTTCTTCT MGNSGGYGGG 1140
GNNKNGNAAA ACNKSRMMAK KSYTTYCCGG CGYCTCCTCCG MCNGGGGGYGS NGCSGGNSTYN 1200
GNNKGRKWTARTN TNNMGCNCTT TSCCCCGCGC GCGNNKRTGC TNTCMNYGSG C 1251

(2) INFORMATION FOR SEQ ID NO:325:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1099 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION:

AAAYTCCGACG MGATATCAC CAAKCTGTYT GGGCCCGACCA AGTGGGACCA TTAACACTTG 60
TATGTGATGCG TCRACAGCTT ATCCCCTGCTA KGTTGCGGTT GGTAGGGTGG CTCGCKTGAK 120
TCTTAAACCG TCGCAAGGCG TGGTGTCCAC CCGCCTGAGG GCAAGATACC ATCAKCGGCG 180
AACAGCTGAC GTCGCGGCGG ACCCGCGGC GCNAAATTAC CTACCAACCG GGGCMTGGCG 240
GCTGCGCNYC CCGTGTGCGA ANTCGAAACTC ASCGCGCNMA CAAKMAACKA HCAACGTTT 300
CTGAACACGA GTCGCAAGCG GTCCGTCACG GGGCCTGAGG CTCAAGCGAGC NTCGAATGCKA 360
TCTTAAACCG GGGCGGCGCM TGGCAACGGT TCTTCTGGNGT GTRCAMCCCN AAAMCAAGCA 420
TCTCCGGGTC GCNGCTGCCGA CCGCAGCCCA TTTTCATACG GSGGCSCNAT CAAATTTGCG 480
GGAACACGSM CNCCGCTCKN GCAMAGCGGCC TMCCAAACCAC CYGCAAGGCGG ATCCTTCCKY 540
NACNCGCNGA RNCNCCXSKT TCGCGCTCTC NMSGCAGAATA CCCCNSSNNT CCGAATCCAA 600
TCCCTMKYGG CCCTTYYTCC CCTCGCCCGGC AAYNGGYGC CCTASSNNMK NKCNCXNANT 660
CCNOWCTCGG NGTGGCCNAN KGYYGGGCTT NMAATSMAN HNNRGGGTYG TSCYACMMN 720
AACCAGNNGK KCCCMCCXKX MANAAKATT RATCAMKNGN GGGKCNKCNH NAAANACCSSC 780
CNYNCGXCTC TMCXSSKGC CGSMMXANCA SNNNGGAGG NGSRRMMKCT CMTXCTCNCNCT 840
MGCCGGCTTN TYCKSGKAKT ACASMKKTCG GCGMCGGGGC MAAMANRAKA TAKACGGYGN 900
CCSNSTMTYN CTSSMMMNMM TCCWMMWATN NTYYGGKXKN KCTMKATWNC CSTSCSKCCKN 960
MRAACXCTFYG SNMCTCCXCA TCNTCKCXXK NSMKNXCCX KSCNMCXN CNKNCKCNWM 1020
GGNTRSCRCY CTTMNNNTCS AGCKGCSKNC WACNACACX NG WTCTYTCX WKNKMNKCNKM 1080
TCKCCKACRGM MTTMTCWCCS 1099

(2) INFORMATION FOR SEQ ID NO:326:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 296 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:326:

GNGNTATACA TCGCTGTTYA CCASAGGATCG ANTGGCGGGCG MAAKCTWSTM CASAGATCTC 60
AAAYCTGCAC MGAGCGGCAC AKAKYCTGCT CMRACCACGG CATACCCWGCG CMGCGCCCGCT 120
CTRRGACCGG GCGTCATGCA ACCGGTGGCG CGCCCGCNGCA CCTACACCAAC CGAGCGCGCC 180
AGGCCGCACC TRACCAAGAC ACCGCGKKT TACCCCGCGG GCGCCCGGGG CACCACCCAG 240
CCACCACCGC ACCAGCGGCG CGCGGCTTGC CAAAACAGGG CGCGGCTTGC CACCRA 296

(2) INFORMATION FOR SEQ ID NO:327:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1073 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:327:

NGNGSNTKMY AATCATGTTA TGCCACCSNGG MTTCAWTTGG CCGGCAATCT TSTMNASAGA 60
TCTCGAAATC GGAGMARGCA TCTGCGCGGN GRATGTCAGA ANGTGKWTRA CGGCWATCGG 120
TTTCGCYCGA ACCAACKTTR SCAKATGGCG GCGAMTGYCA AACCRAATT TTTGCGCAGG 180
AAAATTMTMC CTKTGTRACCA ACTGCGAGCG GCTCAASCAG CAGGCTCTRA ACGTAATAY 240
CKTAGGNTKNT YCCGGCGAAGA ASCYCRATAA TSCGCGCGCGC AMCGACAAAGA CCTGAWNTGT 300
TNNTCCNRAA NCCGGTGCYCC CRAGGGTSA ACTGCGTAR TCTTNTMCYCT TCTTRACAT 360
TAACCACCCG CCGNTCWTGCG CCGCGCAGAA ATYCTGCTGC TKGGCNACCA YCCACNGCT 420
CSGTATGGRS RAANCASTSG CRACGCGTM MCCCSTACKC TGGTCTGATYCT KTGNGNCTCC 480
SNAATTGCCG GATTACCGS CCMGGTATAA CCAGGGCCCC TNGCTCCTYCT CNACAGCCG 540
ATCMWCNCGC TACCCKWTA AATCTTCTCT GTGCTGACCC AwYKCGAGAA NMTNTYCCCN 600
TCCAMGGGGG CCGGGAAAKK CNACCNTTCGG NACCCTTCNC YTTGAAWTTT TCGTGNCCCC 660
GCCGCCAAAS ANACCSSGAKC CCGGGAAYSC WTAGGCYTTCM TGCTCCTAAAT TAATCGCNYCG 720
AATCCGKCAAA CGGTCCCCCG GGTGCTSSCMY TAAAMMTWCC CCGKCSASNG GAAYTYCGK 780
GCWGTMTAW CCNCCCTNTT CYGGKNAAAC SCCCWWCGKN GCSTYCCCSN NMTTSSCGGC 840
GGTTSGAMYC AAAWNTGGG MCMNRAGNGC SNGMCCSCN GKKGGSHATM TKAAYYYCYGG 900
GGGGTGTCNYC CCCRCNSAAA AAYGTYCGGC KCCSCTCYYC CMARTTYYTC CMGRMCMM 960
ACCANGGGNG CTCCCGTNCW WGCGTCCCN SNSMAMAAN NKCCKCGGGS CKGARRNMNA 1020
MCTCSNNGGG WTCGCCNKTGC NCSNNGCNGS YGGNSAASWCC YNYCNCCACA ANC 1073

(2) INFORMATION FOR SEQ ID NO:328:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1166 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:328:
| CGCCCCGGTTC | TTMMTTTTCAY | TCATTCACCG | GGCTTCTAGTG | CGGCCGCAAK | CTGTCTACAA | 60 |
| GATCTCGAAAY | TGCGAMAGAS | ACAATSTCGG | GTGKGGCAGAT | GTNGGGTGSG | GCAACTTTGG | 120 |
| GCTCGGAACT | YGGGGCTTAA | CGGCCGGCTT | RATGTTGCTG | GGAATATCAG | GGTGTTGTTA | 180 |
| TGCCGGGCAG | TACATAATTCG | TACATTGGCA | ATAGTGGGTG | GGGCTACCGA | GGTYGCGCTA | 240 |
| ACAAGCGGAS | TGAGATTYTC | GTAGTGSCTG | NACGGTGART | AAYCGTGAGC | GGTNGGGTGG | 300 |
|TTYCAGTACCC | GGTAGCGGGA | AGTGGTGGGT | YYAGGYCGGS | GSAACGGNNW | YTTNGGTCTT | 360 |
| TCMNNCTTSM | CCKSAMAAMTS | RGGGSTTYCTT | MTYCNNGGAS | TAMTYNNMMM | CGGWYCKCS | 420 |
| WAYCCCTCTGT | CATYCCMCAC | SGSGGYCTCAG | NGYGGCCYTCC | MMNTCCYCAT | 480 |
| CMNCTCCGTTW | CTCNTMMGCC | CSGSNCCTCCT | AMCCTKSGK | CACCNANMC | 540 |
| MTMMCCAINK | MTMMCTCTCN | CCTYNNMCMA | MCMMCTCTT | TCMAACTCCK | CGGYGGGCM | 600 |
| MCTCTCCKCC | AYNMAACCKC | TYGCGCNMYC | CCYGGCKCGC | WYKNMCTCWC | ACTCTMNTT | 660 |
| TCCCTCCKNC | CMKACACKNTT | CTCWCCSCCC | CCACAKAYMC | YMCCMNMTT | MCTCGRACCC | 720 |
| CYYCNYCNCM | NNCMMCTMCN | TNWAKACANCN | TTCCTTCTTC | MYMMTAMAKC | WCMNTCNCWX | 780 |
| SGACAYTCCT | ACTDCMKCCKM | TCTCCTTMCX | CCGYMCNTCX | MCKCMNTMCX | YMCNTMCXCYT | 840 |
| CTTCCNMBRY | CYYAKCNCAC | NCMTCCCAN | KMACAKCTCX | CCCCACMCX | ACNCGGCNCW | 900 |
| TCTCTCCTCC | WCCCTCCKCT | AATCCTCCTCWX | CMYCNYMCMX | ACNICYCAYT | CNACTMMMN | 960 |
| CGANMCTCCT | CTNYCTCNCK | ACYTYCKCKX | CTMCNCTMCX | NRWNCTCTCCT | CCKCGCNCRCN | 1020 |
| CKCMNKTCTM | TCCTCCCMXCM | TCCCMCCCAT | TMKMKSTCTC | WCCMNTCNC | CNCKCYNYNT | 1080 |
| KCTYCCMMYG | CTCCXKNTMT | MCCNWCTYATC | CTMCCKTCTT | CWCACYMCAC | MMNTACWNC | 1140 |
| ACTCTCTRWC | CKKCCMKCMCR | MTCTCB | 1160 |

(2) INFORMATION FOR SEQ ID NO:329:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1230 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:

NGGGGNNNT TCTCATACMN TCTNACCSSG NGMTCWATTG CGGCCGCAAW NCTTGTNNAS 
AGAAATCCTCN AAYTCGGCAC ANATGCCTTT TSTMATKGTG GGCGGGGNGCG CACGCGGCTAT 
GTGYGCCCTGG GYTRACAAAC CCCGCGCGCS CGGGCGCRA CGGGCGGGRCA TSCAGCGCGC 
GGCGCGCSCGG GGGTATATG RAAGGCGGGY TTTTKRTARA AGCGTCSCGG CGGCCGGTrA 
TTAGGGCACA AAYCGGKKT TTGGGRTAT AACGCTAAAT GCACACCAWT TTTYCGGSGT 
AAAAYACCCG CWWGCANATC NGCGGNCCTN RACCYCGCATT YMGCGCAAAA NTNGGGCGG 
AAAACCXCTT TSYTATTXNT TGGGCTATSC GGYTCGCTGCG CGAACGCTC GGCGGGTATT 
TCCTCGCGCC CGCGCGCGCG AAAAACCCG AATYCCGYTG GGGTGKYYCG CMCGCGCGST 
TGTGGCNXNGY CACCGCNCGA AYAAYAAMNT ATCNCTCCTNT TGGCTGYYCA YCTTMTTGG 
SAAAAGGGGSG ATGCGSGGSGG TTYYTTAACCT CAAYCXYCNA ATCNAYATGC CNGAAATTG 
GGGGCNCAA CSGTTCNGTGT CGSNGNCAKXT TGTKMTCCGGN ATCNAAAGGC CGNAATTTG 
TYYSTYNNCN AAATWYKKXY CCCCCWNTG YAAAAXKCA AAASSAACKC YCNAMMYKT 
NGGGGTYTSSG GCCGYCNYCTT SMNTAAACG CCGCGCNNW YNNSGGGEK TCCGCGNSAT 
KCCACCNCCX GNNSGGGNNJ AAAAAXAY TTTYCGCSCA ATCACCACCTY TGYKTSTKYR 
AMACCCTTXT TYYMKAYCCT CYXSKANNTT SGMMTCWWAA YAYCGGCGCT TNGTHHCC 
CSCGGNGGCC AAAATTTGGT YNCAATNTYCC CCGCNANMCN ATMMGGGGEK ECACTTCTCG 
SCYTAMANTA ANAANAAAGK NKTATGGYCT YMACAAACCN GCAGNCTCNY NGAAMAASN 
AKMAAAAKAN KKKTMTNNSA AANCNCCNCCT CTSTYTNNNT TNNTMNNMKCC CYGKKGNKM 
SWSWYNTTCTT NCCRCRCXCC YNNYKNTGANA AAMGGCGCYS GGSTMCRNAN ASMMNTTTC 
STSTNGMCC KMBASNANAM MCAMWKYCC 

1230
(2) INFORMATION FOR SEQ ID NO:330:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1022 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:330:

<table>
<thead>
<tr>
<th>NNNGGGKNNNA</th>
<th>TMAYCWTC</th>
<th>ACSSGGTCTA</th>
<th>TGCGGCAGCAW</th>
<th>CTMGTMAASA</th>
<th>GATCTCNAAAY</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGCCAMNAN</td>
<td>GCAATMCWMC</td>
<td>CATAATAAAC</td>
<td>CATTTGCTCC</td>
<td>GYWTCGCACT</td>
<td>CRAAWCTGTC</td>
<td>120</td>
</tr>
<tr>
<td>CTTCCSKCGG</td>
<td>TTTAARAGG</td>
<td>GTGGMWMTYW</td>
<td>CNTYCTCTAA</td>
<td>SCCCCTARAT</td>
<td>TCKKCTATYC</td>
<td>180</td>
</tr>
<tr>
<td>CTKGGCGCTAC</td>
<td>ATCCTAACS</td>
<td>RATKSCGCC</td>
<td>TTTKATYCAT</td>
<td>RAGCAAWTA</td>
<td>WTGCGYCRAT</td>
<td>240</td>
</tr>
<tr>
<td>KTTGCAGGCC</td>
<td>RACGGGWCT</td>
<td>TTTYCCCGCR</td>
<td>GRACAATNGA</td>
<td>TGGAGWYCCG</td>
<td>TYCCGGGCGC</td>
<td>300</td>
</tr>
<tr>
<td>CCGCCACCAK</td>
<td>ACGGGCNCC</td>
<td>AAGGYCCGC</td>
<td>GCGAMTTGCT</td>
<td>GGGTCAAAT</td>
<td>TGAGCCACT</td>
<td>360</td>
</tr>
<tr>
<td>AAMNCNATTC</td>
<td>CGGGYTRTAC</td>
<td>CGCAAGTAMC</td>
<td>ACAAKAAT</td>
<td>TCCCTWGGCC</td>
<td>GCCACCAWNT</td>
<td>420</td>
</tr>
<tr>
<td>TTYCRATCWCY</td>
<td>CWYCCCCACC</td>
<td>TTRAACCTTGK</td>
<td>YTGCSGTATT</td>
<td>GCTTCKTCTG</td>
<td>CTCRACAGCM</td>
<td>480</td>
</tr>
<tr>
<td>YCNCCCKTCKA</td>
<td>AACCCTGGGT</td>
<td>GACTCACAAT</td>
<td>GCTCGTGYCG</td>
<td>AASGGGGGYT</td>
<td>CAMCGGACAA</td>
<td>540</td>
</tr>
<tr>
<td>AACCCCAANN</td>
<td>TGCCCAAAT</td>
<td>TCCNCRC</td>
<td>CYCGGGAAAAN</td>
<td>GKTGAMTTT</td>
<td>TCSNAACCSA</td>
<td>600</td>
</tr>
<tr>
<td>CMGGNNVTW</td>
<td>NAACCCTGAA</td>
<td>CSSSSGNKGA</td>
<td>MYNSSCGGGA</td>
<td>ANTTTCCCTT</td>
<td>TVNNGGGCRN</td>
<td>660</td>
</tr>
<tr>
<td>AANANCTTTT</td>
<td>AAGGTCACC</td>
<td>KGGNGGGGGK</td>
<td>CCGYTWGGG</td>
<td>AAAACAAAC</td>
<td>CKATGKGGT</td>
<td>720</td>
</tr>
<tr>
<td>TGGAAMNTTI</td>
<td>TKCNCCCCTT</td>
<td>TCCNSGGG</td>
<td>GGGGGCCCMC</td>
<td>CCNMMCTTTN</td>
<td>TCCMCMNTT</td>
<td>780</td>
</tr>
<tr>
<td>YCYYYGGGAAT</td>
<td>TNYTGGGCGS</td>
<td>GAAHYCGGSM</td>
<td>CCGYCTCATA</td>
<td>NCCCMNNWG</td>
<td>GKYTGSGNAR</td>
<td>840</td>
</tr>
<tr>
<td>GGRATMAWTT</td>
<td>TYSTTTTTYMC</td>
<td>CCGGCG</td>
<td>CGCGACGCNT</td>
<td>GNTGAACAM</td>
<td>AAAGGCGGGG</td>
<td>900</td>
</tr>
<tr>
<td>GSNCNMWWFY</td>
<td>YCCNNNGRNTT</td>
<td>TNRGSSSNMT</td>
<td>TMYAAAMMNN</td>
<td>GGGGYKWTY</td>
<td>CKCCNCGCRC</td>
<td>960</td>
</tr>
<tr>
<td>GKTYSGGST</td>
<td>TTTCCNNTTS</td>
<td>GGSSSATYKG</td>
<td>MACCCKMKT</td>
<td>AYCCGGGGGGT</td>
<td>NTKTKYCCCC</td>
<td>1020</td>
</tr>
</tbody>
</table>

(2) INFORMATION FOR SEQ ID NO:331:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1083 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(iii) SEQUENCE DESCRIPTION: SEQ ID NO:331:

<table>
<thead>
<tr>
<th>NNCGNNKNTA</th>
<th>TAMAYCWYCT</th>
<th>NCACCSSGGGA</th>
<th>TCWATTGCGG</th>
<th>CCGGAATCTT</th>
<th>STMAASAGAT</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>CTCKYAYTCCG</td>
<td>GGAMGNCGG</td>
<td>CAMWTTATGG</td>
<td>KTTGRASCGC</td>
<td>ACCAGGCGRA</td>
<td>CCTCGGCGKT</td>
<td>120</td>
</tr>
<tr>
<td>CKTTYCTTGG</td>
<td>AGRGAGGGC</td>
<td>TGGGGCGCRC</td>
<td>GGGTGGCAAT</td>
<td>GCGAACGCCG</td>
<td>CCGAAAACCN</td>
<td>180</td>
</tr>
<tr>
<td>CCGCAATMY</td>
<td>CRAAAAAAAC</td>
<td>CCGGGGGOTA</td>
<td>GKTCCSGGCCG</td>
<td>GCANAAATMAA</td>
<td>TAACGGKTCT</td>
<td>240</td>
</tr>
<tr>
<td>AAACKCGCAGN</td>
<td>ACGCCGACG</td>
<td>GGYCCGCCGC</td>
<td>AACCAGCGNA</td>
<td>CCTCCCCSCC</td>
<td>NATAAGGYYCG</td>
<td>300</td>
</tr>
<tr>
<td>GTGGGCGGGT</td>
<td>CTKATYKCC</td>
<td>AASTGTCGAY</td>
<td>CTCNACGGGM</td>
<td>CGGYCCMCTT</td>
<td>TCCGCCCTAT</td>
<td>360</td>
</tr>
<tr>
<td>CGGTCTTCC</td>
<td>TTMMATTTC</td>
<td>CRTCCACCGYK</td>
<td>GGGGGGAC</td>
<td>TTTTTHNYNC</td>
<td>CTTGSCWAM</td>
<td>420</td>
</tr>
<tr>
<td>CACCNAGGYY</td>
<td>CNAAATAHGC</td>
<td>CCGTCGCGYK</td>
<td>SNNGCAAYGR</td>
<td>GATCGGGGTY</td>
<td>CKKCTTTTNT</td>
<td>480</td>
</tr>
<tr>
<td>TCMCCMMAAC</td>
<td>CCCCTTTNNA</td>
<td>CGGCCCCMATC</td>
<td>CCGTMCWAC</td>
<td>CCCWWMCNNS</td>
<td>ANGKTTGNSA</td>
<td>540</td>
</tr>
<tr>
<td>AAKTNNCCCC</td>
<td>AATRCCCAA</td>
<td>MTTCTTGCGC</td>
<td>MTTTTMWCY</td>
<td>YYCCTTTTCC</td>
<td>CMCCWNNAAA</td>
<td>600</td>
</tr>
<tr>
<td>GGSCRCCTYY</td>
<td>TCGGAANTY</td>
<td>TCCCNCCAA</td>
<td>AWTAMWCCM</td>
<td>TTTCCCNCCCA</td>
<td>AGAAAMTCGSG</td>
<td>560</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:332:

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:332:

NGNGGGGKWK MATACATCTT TCTTCCAAGGG GGAATACATTG CGGGCCGCAC TCTN GTAACCA  60
SAACGTTGCA TTTGCGGCCA AACATCCACC CCAAAATAAC ACCCGAAAGC GCCAGGTGAT  120
CAGCAGCAG AAGCAGGCA TCTATGATGG CACACCACCC TGGCGAGAGC ATGTCGGCGG  180
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  240
AGCAGGCAG TCTCGGTGCA CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC  300
CGGGCGTCAC CCAATGACCG CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC  360
ACACCGGCGC ACCCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  420
TCACGTTCTG CACCAATGAC CGGGGGACCA TCTGATGAGC GGGGAGGAGC TCTGAGCGAC  480
CTACCTGCC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC TCTGAGCGAC  540
GTGTTGTGGG AAGGCGGTGG GGGGAGGAGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC  600
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  660
CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  720
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  780
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  840
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  900
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  960
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC 1020
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC 1080
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC 1140
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC 1200

(2) INFORMATION FOR SEQ ID NO:333:

(ii) MOLECULE TYPE: Genomic DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:333:

NGNGGGGKWK MATACATCTT TCTTCCAAGGG GGAATACATTG CGGGCCGCAC TCTN GTAACCA  60
SAACGTTGCA TTTGCGGCCA AACATCCACC CCAAAATAAC ACCCGAAAGC GCCAGGTGAT  120
CAGCAGCAG AAGCAGGCA TCTATGATGG CACACCACCC TGGCGAGAGC ATGTCGGCGG  180
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  240
AGCAGGCAG TCTCGGTGCA CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC  300
CGGGCGTCAC CCAATGACCG CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC  360
ACACCGGCGC ACCCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  420
TCACGTTCTG CACCAATGAC CGGGGGACCA TCTGATGAGC GGGGAGGAGC TCTGAGCGAC  480
CTACCTGCC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC TCTGAGCGAC  540
GTGTTGTGGG AAGGCGGTGG GGGGAGGAGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC  600
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  660
CAGCGGGAAG CGGCAGACCA TCTGAGCGAC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  720
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  780
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  840
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC  900
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC  960
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC 1020
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC 1080
GGGGCGGCAC CCAAGATGGA CGCCGGCGCG CAGGCGCGCA TCTGATGAGC GGGGAGGAGC 1140
ACACCGCGCG CACCGCGCGC TCTGATGAGC GGGGAGGAGC TCTGAGCGAC GGGGAGGAGC 1200
## INFORMATION FOR SEQ ID NO:334:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1105 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

### (ii) MOLECULE TYPE: Genomic DNA

### (x) SEQUENCE DESCRIPTION: SEQ ID NO:334:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGSSINGNNA</td>
<td>TMATCAYYCT GYACSGGGM</td>
</tr>
<tr>
<td>CGAAATCGGC</td>
<td>AAKANACAC ACCCCCGGTG MTAACACACC CAAATGGTCCT GKTGGCGAAA</td>
</tr>
<tr>
<td>ACCCGAGGAC</td>
<td>GGCACCGCSCG GGGACGTACG GCXGTCCTAC ACCGCGCCAGY TACGTGTTTRA</td>
</tr>
<tr>
<td>AACCCTGGYT</td>
<td>RAGGGCTCCGG CCCAACCTTA AAGCTTTCAC CAAARGATGG GKTGGCCC</td>
</tr>
<tr>
<td>AGCCCACTGY</td>
<td>TGTGGCTGGC CTWYCGGCTG GTAGGGCGGC TTANCGGCCG TTGGGGGTYC</td>
</tr>
<tr>
<td>AMASCSCCTCA</td>
<td>ACGTRATCA CGWNTCCTCC CGSCCMRAC CGACCGGACC TTGRCRCG</td>
</tr>
<tr>
<td>GCCGCSAAATT</td>
<td>CAAYAYCCTY GRWCTTCCTCMAACACACAC AAGGCGCAAC CMSCMAACAC</td>
</tr>
<tr>
<td>ATMGGRACCT</td>
<td>TTAAGGCGCCA GGCAAACTC NTRAKCNCCT CGCGCGCGCGA GGTCCGCGA</td>
</tr>
<tr>
<td>SCRATCCMA</td>
<td>AAAKCCNAT TCCGCCAAGC ACKAACCACAA MMCGGTTTCC TGGCTTCGCCA</td>
</tr>
<tr>
<td>TCTGGAAMCCA</td>
<td>ATTMCGCGKT NCGWGGAAAA ACACSNCCNCNT WATXCCMGG CGCMCGGCCA</td>
</tr>
<tr>
<td>ATTTCSGRA</td>
<td>SAACCCCTHY CCGGGTTTTC YCTGCCTCMM CGCCAAACC CGCGGAAATC</td>
</tr>
<tr>
<td>AAAAASGTC</td>
<td>GGCNAAANGG GCMAAAACSS SACCMMACTT WTTCCRTCTN GGGGGSSWNN</td>
</tr>
<tr>
<td>CCMNGTTTSTA</td>
<td>AKWSSCTCYY CTSSCCAAAY TCGKMCAAAA NNGRRTGTTG TTTGCCCNAC</td>
</tr>
<tr>
<td>NTTTCCGGKC</td>
<td>CGCGGKGGKKG WKGKYTMNMMA CTSTTTTNTTT SCCCYXAAKA NYSCCCCC</td>
</tr>
<tr>
<td>CGSSGCCCGG</td>
<td>CCCGGGGGA NNTTTTAMMA GKTYCCCTCT CCCAMAAAA ANACCCNYC</td>
</tr>
<tr>
<td>CGGGSCGCST</td>
<td>TTTKNAAANN KCTSCCCNGG GNNGGGGKGMK GKTGATTMTN NNNCCCC</td>
</tr>
<tr>
<td>TCGCSCGSSAA</td>
<td>AAAATRMTTT SYCCGCCCNCT CTSCCNXNMR GKMMSMCGC TCCYCTCN</td>
</tr>
<tr>
<td>GCNNTWMAAN ARSNCKCNGN CCCYKCCGS NSNGKCNWCD NCCSTSSNCT NKGNCNCKCN</td>
<td></td>
</tr>
<tr>
<td>KAAANAAAYNC</td>
<td>NGSMSTSSMN CNKCC</td>
</tr>
</tbody>
</table>

## INFORMATION FOR SEQ ID NO:335:

### (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 936 base pairs
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA  

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:335:

| NGSNSNQNNN | TAMAYCWYCC TSCASCSNGGA ACWANTGGCGG CCRMAWCTS TIKASAGATC | 60 |
| TMGAAYTCCG | CAAAGCCGCG ACAGTGGGCT GCATCAGGC CAGTCCTTGAC TCGGCTGCG | 120 |
| CSGGCGKTR | GASCACGAC CAAACCCGTC CCGGACGGCA GCTGGTCCCT | 180 |
| GCMAWYCCCT | CGCAGGCCGCT GCAAGCAGAC TGGCCTGCGC | 240 |
| RACTTCCCGT | CAGTCGGTCG CCGGTCGCGC | 300 |
| RACRCACC | AATTTGCGG AATTCTCAGC CCGCCTGCGC | 360 |
| CGCAMCCTY | CCAGCCTTCC ATGTCGCTCC | 420 |
| KCYTSYCGC | GTAGCCAGATG TATATCGTCG TGCGTCTCG | 480 |
| KTAATTCGTC | TGAGTGCTCAG ATCAGCAGC NCCGTTGGAT CTGTAACCTAAC | 540 |
| CCCCACACCT | AAAAAAGGA TATAATTGAT CGWACAAAC CCCCACCCCGC | 600 |
| GCCGCAACC | CACGNGGGTG ACGTGGGAAT ATYMAACCACC CACCACTTAC | 660 |
| AAAACCCCGN | GGGYMCACAA ACCCTTCTTT GCAGCTCGCG ATCCACCTAACC | 720 |
| TMMTTCGGTTT | CRWAAAAAC GCAGCNCCTCG GAAACTTTT TGGCAACCACAA | 780 |
| CCCCCTCCTC | TGGSSGNNCGC TATCAGTGGT CCGGATGGTCW | 840 |
| SNCCCNAAN | YYCCNAANNG NKCCCGSNMA AAGAGAANNTT YCMKAANACCC CCCCACCCCG | 900 |
| NAAAYAACC | MAAAWTTCC AAASMSCNNG YCCCCC | 936 |

(2) INFORMATION FOR SEQ ID NO:336:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1042 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:336:

| NNNGNGKNNY | ATTAATTCYY CTSCACCCCGG GGNWCCWATT CCGCCGCMAW KCTTTGMAAS | 60 |
| AGATCTCMRA | YTCGCGACAG ASGGCAGCAG ATCAGGGCGG CTATAYMCYC GGTCCTCGAT | 120 |
| GACCAACCG | TCCTCGGCGW GRATAATGGC NCCGCCCGCG CCCCAAACAG YTCCTAATG | 180 |
| TCGCCCAACCG | CATATNCAC CGGACRATA AAASAAAAC CGCSGCGCG GCTGGTGGCC | 240 |
| SGGRGACGG | GTCGCAACCAC AAAACNCCTT GGGAGCGCGK KTSSRATATCT AASGTTAAT | 300 |
| TCGTCTCCTC | CCCGCTCTGG ATGCGCTGGT CGCGTCTCGG CCGGCGCAC | 360 |
| GGYCRTTTT | CATCNCTCAG YCTACACINT TKGCTYCAAC CAAACACACT CAAAACATTG | 420 |
| CCCCCTTCTG | CGGSSGCGCGG GTATGGNNCCGG TAATACCTCT NTGKTCGCTG MYCACCGWGA | 480 |
| CCAATANCCTG | CGGCGCCTTG GCACCTTCC SAAACTCATY CTCTGTCRAC CCCCACAMRC | 540 |
| CTNSAATACCC | GRATACATNC CCCCNGCTTCT NTYCTCTCN GRCCACATAT GTGGGCTTCA | 600 |
| RKTNCYYCA | TSCAATAGGTS YTYCCTRTGC YGGTCTCAGA ATNACAMAS GTTGTTCMCT | 660 |
| ACCAAACACCA | NTGGSCSCNNAA CMNAAANazon AAAANAKGC KCTTYYAAC CCCCCCTAT | 720 |
| TCAWYCCGAT | CMNNWCCCNG NKKAGGKGN AAAAYTHRA CCAAANCMCT ARSTTSGNAK | 780 |
| AAAAACTYG | CGGTSSCMCCAA MDTNTTSCTG CTGCGCTCGT YCCAAATMSA AAAYTTCCKK | 840 |
| KRMNAAAMMC | YGMCCCSSCACA ANAATTTTGT NAAMCCCMNA YTCCTTMCCT WTTTCTCYCC | 900 |
| CCMNCSNSNG | GNTNCCTYTMT VYCTTCAHM MCRNNSGAGNC CCCCMTYTT TWTCTCWNC | 960 |
| MNARGSWNYT | RGMNNMNCNC CCCNCNCNAK MTCCNCNAAK NTTTNAACNN NYKCYCCGCC | 1020 |
| CCCMMNKNCC | CCCMNCTTT TM | 1042 |
(2) INFORMATION FOR SEQ ID NO:337:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1073 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:337:

| NNSGSGMKK | ATATATCWCCT | CTSYACCNSN | GMTGWTWTG | GGCCGMATTC | TNGTMAASAG | 60 |
| ATCTGAAATT | CGGCAAAANK | ACGCMAAYGT | AAGTGGRAY | CGGTCAATATA | TCMTCGCGGNG | 120 |
| TCAACCCAAA | AGCAGGNGTC | CGGYTCOCCT | GGGGGGCGAC | CCCCTAATGGT | RATGCAACTY | 180 |
| GGGGGCCGCC | CGYCAAAAGG | KTCWTRRAGG | CSGTAAAGGT | CAMCAATTCC | TRAGTGCMCN | 240 |
| CACGGTMTT | TGCGGGGCCC | RAWCYTCRAC | CGGAAATWTC | GGTAAATGCG | AATTGGGGC | 300 |
| YCGGCTTGGG | CAATAAAGTN | TTGGGGCAAGC | GGGGWTYCYN | NCTGGCGCRA | ATTTCCCNCA | 360 |
| TGCTTTAACG | GTGRACCGGT | TTYGCGCYGT | GCCGTAAAYTG | YYTCTGGGGG | GCCGYCGGCG | 420 |
| CRNAGCASSY | CRCITAAGGY | CMCCAGGCAA | TACCGTGGG | TTTTTAACAA | CGGRATNAAY | 480 |
| TGKTACCCCA | YTCAASSGTS | CGRTGNTTRK | TCTNTGGRRA | AANMcAACC | AACCAGGTNT | 540 |
| RATCTGGTCTT | MTCANWTTT | SGGGGAATCT | GCCGTTTGST | AAYCTTNTATC | CMTYCAAAAG | 600 |
| GTTTTMAATT | CCANRRAAAT | CGGYTRGGCA | CTTTGCGGGC | GGGGTGGTMM | GMWCTTTTR | 660 |
| AMATCCNCSS | CGGGGSSAAN | AMTSGGGNTT | SGGGGGGTTTC | CCCCNAATAT | XONTGNCCNT | 720 |
| GNAATTTGSS | GGAGGTCGCC | GGNYAYGGCG | CCMTKGGGGG | TNCCAGCGAT | WGAACATTY | 780 |
| WKCCGTCGCC | AACCCGGGNC | CGGGGCTTGGG | GCSCNTTTTT | CCTMYNAAA | AAGKTGGTGG | 840 |
| NYTTTGMGCG | CNRAAANTTC | CCGCNKNTNT | GGNCCNAAAT | YYCACATNTT | CANACCTTAA | 900 |
| AASAAAAYK | YGKTYYYYYC | TTTTMCSCSGS | SANCCCCCM | NMSKCNCGG | AAATAGAN | 960 |
| TYNGCCTTAN | CNSNKTKKTN | TNKTYCCCCC | NMWNNNSNMCY | NBKBCNKKY | NGSNMNCCCT | 1020 |
| MKYSKCNNNNN | SNDRNNKCGN | GSNCASMXMYK | CMNCCNGKMY | NNGKNSSNCCC | MUR | 1073 |

(2) INFORMATION FOR SEQ ID NO:338:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1061 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:338:

| GNSNGMKNTN | TMCAWCYCTT | SCACSAGGGTC | TATTGCGGCC | GCAAATYTNGT | CKASAGATCT | 60 |
| CGATYTGCGG | AMNAARARTG | TCGTGCTCAA | TTTTGATCGGTG | GKCTCATAAAY | GGGCCAGCGG | 120 |
| GNGACCRACA | CCACTGTAAT | CCCCAAANAC | CACAGCCTNCT | AAATWCTCAG | GCRGACCACT | 180 |
| TRTCAATYC | CRASCATKTA | ACCGTACTTGW | TCRACGGGTC | CRAACAGGCG | ACGCAGYCTA | 240 |
| CCGGSGGGGC | AWTTGCGGCTG | CCGGGCGGNT | TCAGGTGTTGAT | TYTGATCSCCT | RWCTGGTSTS | 300 |
| TGGYCMCNTT | GGTGAGGGCC | WAACCGMKMG | AGAACATGGG | GGCAATCTTT | CAGGGACNGA | 360 |
| GRAAACCNGA | GAAGCCGGGG | TAKANCCCGG | CRAAAACRAG | GCGGYGGGNN | ATTCCNATTA | 420 |
| NAMSIGNNTT | CRACCTTGGG | RAACGCTTGT | CTTTGCTGCC | CGTCCGCAAC | CGTCCGACT | 480 |
| TACCCCKNCT | CCGGGMCMAC | CYGGGCTTGGT | TGGYKCCCAT | NTGGCGCGCCG | GRNRAWNTNGG | 540 |
| CNAATTCCAG | GGCCNCANCTT | TTTGCAGGCGG | AAATCCCTGATG | GTTATCAACCG | GGGCCGCNMCT | 600 |
| GGTGTTGGGG | AACCRCNCYS | CTMTITTTAAA | CATTGCGGSC | CAAATGCGGNC | STTGGSSAAAT | 660 |
TCTNTYCGGT GGGGCGSGCR ANMYTTTCTCT YCCNAASAN TCTAMYCCAN TACGSSNCC 720
CGGKCAAANWS NGGGGGGGNA AAGGGGCCCC CGNGTSCCKC GGGGGGCCGG CYGGKTTCGAA 780
AANNTTSCGCG KKTSTMSCGCG NVTCSCCCCC CGSICCAAGRA CGNNGGTTTT TTTTTGAAAC 840
KCMANTCSSA AMCCCGCCSSC CCAGAAGGSG CCTCNAWRG RAYTNKNCRR CNNAACCGSG 900
CCCCCAKTYG SSNGKTDCNCC CNCCSSGGKTG CMOSTTTNMM MRCCTTNTGN QNNKTTTAN 960
MOSCTTTNNC CACCCCTCYCK GGGKSMNNA GAAKTYWYKC CNGGGGNNAN RSCCCCCCN 1020
GSGKGGGGGKG MGAGYSSCCT CTGGCCNCCNN YNNTTTCCTCC C 1061

(2) INFORMATION FOR SEQ ID NO:339:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 986 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:339:

GNNNGNNNKWN ATMCTACWYHY CTSCACCSSG GATCGATGTC GGGGCGCAWKY TNGTMAASAG 60
ATCTMMAYGT GGGCAACANAG CCG/CAGGAG TGGTGCAGTC TTGGTCANAG CTGTCAACGSG 120
GGGGTCCCGCG TGGTTRASCMA CMACTGGCGCR AAGCACAACAC CCGTCGGCGCG GYCCGCGGCG 180
TGGCCTGCBA AAYCTTCGCAG GCCAACACRC ACAAAYWYCTT CCTGCACACSC ARSCGGTCTYC 240
GGGCCTGGAT CTTGGKYSACG YTCGCTCGTGC GGGGGGCCGA GTACTCGCGC CWYCRANAC 300
GCTYGGGRRR ACCNAAAGTC AATCTTGGCNCN AATTTGCTNTT CCCCTTCTCCT TTRTANNAATT 360
GTTAANCCCA CCAGAACCCTY CGGCGKTCCTC CTGCGCGCRA WTCGRWTCGC RAYNCGCGGA 420
TGCGCTNKTC KYCTVCYKCS GMGCCAAAAT CTGCTGTATCC TATATTGGCC CTAATGCGAA 480
ATCTKGCGCTG TGATNTGTCG GGGCTGCCGA TTWAMANCAG NGGTTCCTTY TCTGCCNAAC 540
CCSTTTGGGG CGAAGCAAAGA AATGTNAATA ATAATGTCGG TNGTCAAACGC CGNCGGCTAT 600
CNACGSKCGC AMCCCGCRGN GGGTANKNGG GNAATTCTMM AACCACCAAGG CATAAASNTTG 660
SGANAAACCC YHCMLCGGGYCA CGGGCAANAY NTNTTGGNHY SSNTTGGGMM YCATGCTCTN 720
CMNAAASCAC AAATCTNYGG GYCCAATAAA AAMGGSGGYCY SAMCCGGGAA WTTTYYTTCN 780
KNAANACCCNA AAKCCTTTTT CNAACCDDAN NTVYCCNCC NRCRMAATNG CNSSGGARTK 840
SSSSCTNNCN ATGKYCCMAA AGNNGGRANA CGARCACCAAA TTCCCTNNTN ENKNCCCNST 900
TRXNAAAGGG G0KNTYACMA AASCNCNNCC NCTCTCCCA AAKAMCSCCN AAAGAGKNTCN 960
NAANASKYSN NNNSSCCCCC CCCMNM 986

(2) INFORMATION FOR SEQ ID NO:340:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1074 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:340:

NGNNGGNXRN ATMYAYCNCNT SATYVVACSN GMMNMWATTG CGGCCRMAWT CTNGTMKASA 60
GATCCTMAGAA YTCGCAAGG AGAYAKTGCG GGCGCCAGAT TTTNGGCCCG CAACCGCGCCG 120
AGCTTGGCGYW TCACAKCTCC SGGTTGCGCCA AAAAAMWTCNT ACCCGCTATMC TYYTGCGASM 180
ASYTGCGCSC RAXTRACAC CGCCGCGCGW TGGTGGCGCA GGGTATYCD AAYGYAAY 240
YCTYTTCTAGK TAAATCSCG CSGGCGCSCA CNCCACGGCG CGGTXTAGGY GCYVCTCTCA 300
ATMACCAGCY CGCCCGAGGG CACCTTTGCCC AAAAYCTCTT GGGTCGAGCCA AATTCGCCG 360
CCGCGCAACCM ACCANCAGCA TYCTGCGTNTA AATCYCAAGG GCCGCCCGGTG TAAAMMNANNA 420
GRATCTCCTCT MANCCCCCCTTCAGCCTTNTA CNGCMACTCG CCGCTTCTCTT CAMACCCTGGA 480
RTACCGGCGGT CAACCGGCCS GTCAAAACTC ACGAGGCGNCT ACGCTCTCCC CGGANSAAAG 540
GTCTTACCSCC NNYAANAAA MAAGNTCTGT TTTTCCCCTC CAAASNNAA AANCCCGGCGC 600
CGGCGGCTTCC NMGGGTTTGG GGGMNAAATC AARNCNCGGN CAGAAGNATC CGAAMCTCCT 660
CAAGTNCNMT TWAANACYCEN NNAAANCNNAA ANTTTGTGGGA AAGGNTCCCTC NTNTYNCCCTC 720
TTTNTAGKTS GGGMYYCTYT TAAAANAAATT CCCCCAAAGG CCCCAGGAAG GGCTMCMTGC 780
GGNAAATTTTC CAAMCMNNGK TTTNTYNNTC TMCGGGGGGA AATTCNNTCT CCYNNNGGNGG 840
CSSSGNMMAT TAYGGMNNMNT TTTNNAAWMTT NSGKETTSAMM YNNKCMNNN NNNSMNNANK 900
TNAMCNCCCN CCTCNGNGKY CSCYNCCCSSGS GNMGNNGRAS MKCMNNNAA MKCMNNNNNK 960
CGGAMMCMN ATAKGNNSNC CGGGASCMN NYNNMAAATMT CCCCNNNNNNN AANRGMRACN 1020
CCCNNSNNGMN RGRAGAMTNY YCCCNGGCGKM CGGKNAAAAW YCYYCCCGGCG 1074

(2) INFORMATION FOR SEQ ID NO:341:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1195 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Genomic DNA

(xi) DESCRIPTION: SEQ ID NO: 341:

NGNMGNCNKNT MTACATCWT TCCGCAAGGG GTGCTGWTNGC GCCGCCCAWKC TTGTCTGASAG 60
ATCCGGAAYT CGCCCGAGGG ACHCTCGGCA CGGCCCGCAAC NACTCTCGCG CGTGTACTCC 120
ATTAGNGCCK TCACCGGGGC ATYGNANCAK TNCACTGGGAG TGGCTGTCGG CTGGCGGGCG 180
GGCCTCAGGG CKCTSCGCTTCT NAAAGCTTGG CGGACCGCAG TGGTTTTGCTC AAGGCTCCTG 240
AAATGGGCCA CGCGTCTGGC TCATGGNNTC TACGCAAGGC CNGGCCCAAC CRCTTCCTTA 300
AATCCGYYCC NTCCTGANCN CTTGGATAYCC CGGGSAAAGT AGGTCGTTGC KCCNGGCTGC 360
TGGAGGCMN GCAGTCTTCT NMTCAAMNTC NMTCAAMNTC CGGCGAAAGG NGAACCAACT 420
TTNGGMGNG GCCGGGTNTT CAGGCTGATCA ACAGGGCAACT GGYKNNSTTC 480
KGGGGAAAGG RCTCCTCAGT MTGTYGCTCCK CCGGCAKANC CGCSCCGCTCGK GYCGCAATG 540
GAAGGCMASN CGGTITTAA GCMYCTCRR RSCGGTSTDGA SGGKMTTCCGG MGGANKANNN 600
NKKMAMWTTTK TGCRNNGGGC GATSTCCCGGG CKSTTAKAGA ANACTICTCWW WCCGTNTYSC 660
SAAAGNTKCS CGGMGTITTTS SCKCMGANN YCTGATTTS GGCGGGYKCC CCCCAGGGYT 720
CGAANKMKRY CYYAGGGGGM GNYSAGCCSC CCMNNATNAG AGNAAGGKTTT YRGSTSKNCN 780
TYTNKGGGAC WSNCMNSNAK ANAACNKNKET TGSCSCTNTMS AGNKTNFKRTT YCNKNTSTDG 840
TAAAGGAGGC ATYKMKCGGC CKTTGANGMM GAGWMGCCGC KYCCCSSNKR RTCNNTNWAA 900
TATSKSAGMGX XIICCGMAAMK CKCCTGTGTCT TCGTANANM SMMRKNKGGT TCGGYCTTSG 960
GGGNTTCTTA GGGTATCGSC CSCSSWGGAC WCGMCMNMG AGRKNTNNTS YANTGARGGY 1020
MBNNTKRMTM MScGCGGNN GAANNCGGGC CSANGMNGTGY NGGGNMSNN ARAGKATGAT 1080
GGCCNCNMNN MNFGGAGNMMG SANNMGGMGRG GGGGKTKGK TCCSCGCGMS CSAMGRAGAA 1140
GKTCNGSGCC CGMGGKYGRT KCTTKNNTGG YSTCSCSMMM NAGAAAGAG AGGGC 1195

(2) INFORMATION FOR SEQ ID NO:342:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3572 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(ii) MOLECULE TYPE: Genomic DNA

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:342:

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Description</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCATCTGATC</td>
<td>GTGGGCAACC</td>
<td>60</td>
</tr>
<tr>
<td>TGTTTGGTTG</td>
<td>AAAAAACGGA</td>
<td>120</td>
</tr>
<tr>
<td>TTGGATTTGCG</td>
<td>AGCTGCACT</td>
<td>180</td>
</tr>
<tr>
<td>TAATGGGCG</td>
<td>CCATACATTC</td>
<td>240</td>
</tr>
<tr>
<td>CGATCGGC</td>
<td>AGGACACTC</td>
<td>300</td>
</tr>
<tr>
<td>CATCAGAGA</td>
<td>AACCCGCAAG</td>
<td>360</td>
</tr>
<tr>
<td>CATCCAGCGG</td>
<td>ATAGTATGATG</td>
<td>480</td>
</tr>
<tr>
<td>CGGGTTGAC</td>
<td>AGGGGATGGA</td>
<td>540</td>
</tr>
<tr>
<td>GGCGCAGCA</td>
<td>AGGAATCGG</td>
<td>600</td>
</tr>
<tr>
<td>GCCCGGGCC</td>
<td>CCAAGTCAT</td>
<td>660</td>
</tr>
<tr>
<td>CGCGCCCGC</td>
<td>AGGCGGAGA</td>
<td>720</td>
</tr>
<tr>
<td>CTCTCCTCAT</td>
<td>GGGCGCA</td>
<td>780</td>
</tr>
<tr>
<td>GCCGGCGG</td>
<td>AGGCGCAAG</td>
<td>840</td>
</tr>
<tr>
<td>GCCGCCCA</td>
<td>CCAAGCAGG</td>
<td>900</td>
</tr>
<tr>
<td>GCCGACGTC</td>
<td>GGAGACAGG</td>
<td>960</td>
</tr>
<tr>
<td>GGCCTCG</td>
<td>GGCACCTCG</td>
<td>1020</td>
</tr>
<tr>
<td>CACCTCAT</td>
<td>GCTGCTGAG</td>
<td>1080</td>
</tr>
<tr>
<td>CTGGTATCCG</td>
<td>ACGAATAGG</td>
<td>1140</td>
</tr>
<tr>
<td>ATGGGCGCA</td>
<td>GAGGACAGG</td>
<td>1200</td>
</tr>
<tr>
<td>CTACGACG</td>
<td>GGGTCGGCT</td>
<td>1260</td>
</tr>
<tr>
<td>CTATCGCAT</td>
<td>TGAGTCGAC</td>
<td>1320</td>
</tr>
<tr>
<td>ATGGGCGCA</td>
<td>GAGGACAGG</td>
<td>1380</td>
</tr>
<tr>
<td>CTTGCCATT</td>
<td>TGAGTCGAC</td>
<td>1440</td>
</tr>
<tr>
<td>GACCTCAGATC</td>
<td>GATGGAATG</td>
<td>1500</td>
</tr>
<tr>
<td>AGAGGGTGC</td>
<td>AGCTCAGATC</td>
<td>1560</td>
</tr>
<tr>
<td>GACCTCAGATC</td>
<td>GATGGAATG</td>
<td>1620</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>1680</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>1740</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>1800</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>1860</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>1920</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>1980</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2040</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2100</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2160</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2220</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2280</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2340</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2400</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2460</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2520</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2580</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2640</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2700</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2760</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2820</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>2880</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>2940</td>
</tr>
<tr>
<td>GCTGTGGAC</td>
<td>GTGGCTGAC</td>
<td>3000</td>
</tr>
<tr>
<td>TCAATCAGG</td>
<td>CTACGACG</td>
<td>3060</td>
</tr>
</tbody>
</table>
AAACGTCCTG CGTGAAGGCCA ACAGGAGTGTTC TGGAAAGCGCC TGATATTACG AAGTCAAGTT
CAGCGATCCG AGTAAGCCGA ACGGCCGATT CTGGACGGGCT GTAATCGGCT CCGCCGCGGC
GACGCCGCCG CAGCGCGCCG CCCCTGACGCG CTGGGGGTGGT GTAATGCTCG GGACCGCCAA
CAACCGGCTG AGCAAGGCGG CGGCCGCAAGG GCGGCACCGG CTCACTACCGG AGCACCCCGA CAGGCCACCG
CCGCGCGCGG GCGCGCGCCG GGGAGTGCAC TGCTACCCCCG ACGACCCCGA CAGCGCAGCG
GACCTTAACCC GCCCTGAGAAT CTGGAGAGAA TCCATACACG TCCGCGCGCCG TGGACCGCCA GACACCGCA
CCACTACCGC CACTGAGATC CGGCTGCTAA CAAAGCGCAG AAGGAAGCTG AGTTGGCTGC
TGCCACGCCG GAGGAATAC TAGGATAACC CCTTGGGGCC TCTAAACGGG TCTTGAAGGG
TTTTTGCTGT AAAGGAGAGAA CTATATCCGG AT

(2) INFORMATION FOR SEQ ID NO:343:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:343:

Val  Gln  Phe  Gln  Ser  Gly  Gly  Asp  Asn  Ser  Pro  Ala  Val  Tyr  Xaa  Xaa
     1         5         10        15
Asp  Gly  Xaa  Arg
     20

(2) INFORMATION FOR SEQ ID NO:344:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:344:

Thr  Thr  Val  Pro  Xaa  Val  Thr  Glu  Ala  Arg
    1         5         10

(2) INFORMATION FOR SEQ ID NO:345:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 10 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:345:

Thr  Thr  Pro  Ser  Xaa  Val  Ala  Phe  Ala  Arg
(2) INFORMATION FOR SEQ ID NO:346:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 12 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:346:

Asp Ala Gly Lys Xaa Ala Gly Xaa Asp Val Xaa Arg

(2) INFORMATION FOR SEQ ID NO:347:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 18 amino acids
  (B) TYPE: amino acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:347:

Thr Xaa Glu Glu Xaa Gln Glu Ser Phe Asn Ser Ala Ala Pro Gly Asn

Xaa Lys

(2) INFORMATION FOR SEQ ID NO:348:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 27 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:348:

CTAGTTAGTA CTCAGTCGCA GACCGTG

(2) INFORMATION FOR SEQ ID NO:349:

(i) SEQUENCE CHARACTERISTICS:
  (A) LENGTH: 25 base pairs
  (B) TYPE: nucleic acid
  (C) STRANDEDNESS: single
  (D) TOPOLOGY: linear
(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:349:

GCAGTGACGA ATTCACCTCG ACTCC

(2) INFORMATION FOR SEQ ID NO:350:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2412 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:350:

CATATGGGCC ATCATCATCA TCATACGTG ATCGACATCA TCGGAGCCAG CCGGACATCC 60
TG GGACAGG GGCGCGCGGA GCACGATCGG ATCGACAGGC ATCGACCGG 120
TGCGCTCGGG TACATGACCA GAGATGACCC TCGGAGACGG CCGGACAGAT CACCTACGGC 180
ATCACGCTCG AGGTGCTGT GTGACGACAC GAGGATGTCG CCGGATGTCG CCGGACTGTCG 240
AGCGGCTTCC CTCAAGCCGC CCGGCGCGCC GACTAGCTGG CAGGATACCC CCGGACTGTCG 300
CCGGTACGTG TGGTACCGTG CTGTTGGTCC CGGAGTACCC GAGTACCTTCG CCGGACATCG 360
CGCCTTCCTTC AGGAGAGTGA TCTGAGATCC ATGGCCTGAC CGGAGTACCC GAGTACCTTCG 420
CGGCGATGCTG CGGACGATCTG GGGGACTCGG CTGCGATCGG GGGGACTCGG 480
TGGAGGTGTG ATATGCGGCG CCAACAGGGG CTGATGACGA TCCTGCTGAC CATCTCGCTG 540
CAGCCAGGATCA ACTGACACCT GGGGGAGTCC AGGAGCCAGG CCGCAAGGTC 600
CTGGGAGGATG CCGGAGCGA ATACACAAA AGCTGGGGAGG CCGCAAGGTC 660
AAGCCCGCGA CTGACGCTGAT CGCAGGCGGA ATCTGGGGAGG CCGCAAGGTC 720
GGGACCAGGT CACGATCCTG TCATGACGGC ATCGGAGCAGG CCGGACTGTCG 780
GGGCCGGCCTG CACGATGACAC GGGAGGAGG CCGGACTGTCG CCGGACTGTCG 840
GGGCAGGCGG CACGATGACAC GGGAGGAGG CCGGACTGTCG CCGGACTGTCG 900
ATCGACGGGG CGCAGGTGAG CAGTACGAGC GGGAGGAGG CCGGACTGTCG 960
TCGGCGACTC TTGACGATCG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1020
TCGAAAAACC CGCAGGATCC AAGGATGACCC GGGAGGAGG CCGGACTGTCG 1080
CGAGACCTACG AGCACTAAAA GTACCTGCTG CACGATGACAC GGGAGGAGG CCGGACTGTCG 1140
CGGAGATGGG CTACGGCTGC ATCGGAGCAGG CCGGACTGTCG CCGGACTGTCG 1200
ATATGCGGATG CTACGACCTG GGGAGGAGG CCGGACTGTCG CCGGACTGTCG 1260
AGGAGCTGGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1320
AACTGGGGCG GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1380
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1440
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1500
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1560
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1620
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1680
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1740
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1800
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1860
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1920
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 1980
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 2040
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 2100
GGGATGACGT GGGAGGAGG CCGGACTGTCG CCGGACTGTCG CCGGACTGTCG 2160
ATCGGCTCGC CGCGGCGGAA CGCACCAGGAC GCCGGGCCC CTCAGGAGCT GTTGTGGATA 2220
TGCGTCGGAG CGGCCAACAA CGCGGGGAG AGGGGCGCGG CCAAGGCCT GCACGAATCGG 2280
ATCCTGCGCTT TGGTCGGCCCC GCGCGGCGC CGCGGACGCG CTCTCGGAGA GCCGCGCTCGG 2340
GGCCCGGGCC CGCGGCGGGA AGTCGCTCCT ACCCGACGCA CACCGAGACC GCACGGGACC 2400
TTACCAGCCT GA 2412

(2) INFORMATION FOR SEQ ID NO:351:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 802 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:351:

Met Gly His His His His Val Ile Asp Ile Ile Gly Thr Ser
1      5      10     15
Pro Thr Ser Trp Glu Gln Ala Ala Ala Val Gln Arg Ala Arg
20     25     30
Asp Ser Val Asp Asp Ile Arg Val Ala Arg Val Ile Glu Gln Asp Met
35     40     45
Ala Val Asp Ser Ala Gly Lys Ile Thr Tyr Arg Ile Lys Leu Glu Val
50     55     60
Ser Phe Lys Met Arg Pro Ala Gln Pro Arg Gly Ser Lys Pro Pro Ser
65     70     75     80
Gly Ser Pro Glu Thr Gly Ala Gly Ala Gly Thr Val Ala Thr Pro Ser
85     90     95
Ala Ser Ser Pro Val Thr Leu Ala Glu Thr Gly Ser Thr Leu Leu Tyr
100    105    110
Pro Leu Phe Asn Leu Trp Gly Pro Ala Phe His Glu Arg Tyr Pro Asn
115    120    125
Val Thr Ile Thr Ala Gln Gly Thr Gly Ser Gly Ala Gly Ile Ala Gln
130    135    140
Ala Ala Ala Gly Thr Val Asn Ile Gly Ala Ser Asp Ala Tyr Leu Ser
145    150    155    160
Glu Gly Asp Met Ala Ala His Lys Gly Leu Met Asn Ile Ala Leu Ala
165    170    175
Ile Ser Ala Gln Gln Val Asn Tyr Asn Leu Pro Gly Val Ser Glu His
180    185    190
Leu Lys Leu Asn Gly Lys Val Leu Ala Ala Met Tyr Gln Gly Thr Ile
195    200    205
Lys Thr Trp Asp Pro Gln Ile Ala Ala Leu Asn Pro Gly Val Asn
210    215    220
Leu Pro Gly Thr Ala Val Val Pro Leu His Arg Ser Asp Gly Ser Gly
225    230    235    240
Asp Thr Phe Leu Phe Thr Gln Tyr Leu Ser Lys Gln Asp Pro Gly Gly
245    250    255
Trp Gly Lys Ser Pro Gly Phe Gly Thr Thr Val Asp Phe Pro Ala Val
260    265    270
Pro Gly Ala Leu Gly Glu Asn Gly Gly Met Val Thr Gly Cys
275    280    285
Ala Glu Thr Pro Gly Cys Val Ala Tyr Ile Gly Ile Ser Phe Leu Asp
<table>
<thead>
<tr>
<th>290</th>
<th>295</th>
<th>300</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gln Ala Ser Gln Arg Gly Leu Gly Glu Ala Gln Leu Gly Asn Ser Ser</td>
<td>310</td>
<td>315</td>
</tr>
<tr>
<td>Gly Asn Phe Leu Leu Pro Asp Ala Gln Ser Ile Gln Ala Ala Ala Ala</td>
<td>325</td>
<td>330</td>
</tr>
<tr>
<td>Gly Phe Ala Ser Lys Thr Pro Ala Asn Gln Ala Ile Ser Met Ile Asp</td>
<td>340</td>
<td>345</td>
</tr>
<tr>
<td>Gly Pro Ala Pro Asp Gly Tyr Pro Ile Ile Asn Tyr Glu Tyr Ala Ile</td>
<td>355</td>
<td>360</td>
</tr>
<tr>
<td>Val Asn Asn Arg Gln Lys Asp Ala Ala Thr Ala Gln Thr Leu Gln Ala</td>
<td>370</td>
<td>375</td>
</tr>
<tr>
<td>Phe Leu His Trp Ala Ile Thr Asp Gly Asn Lys Ala Ser Phe Leu Asp</td>
<td>385</td>
<td>390</td>
</tr>
<tr>
<td>Gln Val His Phe Gln Pro Leu Pro Pro Ala Val Val Lys Leu Ser Asp</td>
<td>405</td>
<td>410</td>
</tr>
<tr>
<td>Ala Leu Ile Ala Thr Ile Ser Ser Ala Glu Met Lys Thr Asp Ala Ala</td>
<td>420</td>
<td>425</td>
</tr>
<tr>
<td>Thr Leu Ala Gln Glu Ala Gly Asn Phe Glu Arg Ile Ser Gly Asp Leu</td>
<td>435</td>
<td>440</td>
</tr>
<tr>
<td>Lys Thr Gln Ile Asp Gln Val Glu Ser Thr Ala Gly Ser Leu Gln Gly</td>
<td>450</td>
<td>455</td>
</tr>
<tr>
<td>Gln Trp Arg Gly Ala Ala Gly Thr Ala Ala Ala Val Val Arg</td>
<td>465</td>
<td>470</td>
</tr>
<tr>
<td>Phe Gln Glu Ala Ala Asn Lys Gln Lys Gln Gly Leu Asp Glu Ile Ser</td>
<td>485</td>
<td>490</td>
</tr>
<tr>
<td>Thr Asn Ile Arg Gln Ala Gly Val Gln Tyr Ser Arg Ala Asp Glu Glu</td>
<td>500</td>
<td>505</td>
</tr>
<tr>
<td>Gln Gln Gln Ala Leu Ser Ser Gln Met Gly Phe Val Pro Thr Thr Ala</td>
<td>515</td>
<td>520</td>
</tr>
<tr>
<td>Ala Ser Pro Pro Ser Thr Ala Ala Ala Pro Pro Ala Pro Ala Thr Pro</td>
<td>530</td>
<td>535</td>
</tr>
<tr>
<td>Val Ala Pro Pro Pro Pro Ala Ala Ala Asn Thr Pro Asn Ala Gln Pro</td>
<td>545</td>
<td>550</td>
</tr>
<tr>
<td>Gly Asp Pro Asn Ala Ala Ala Pro Pro Ala Asp Pro Asn Ala Pro Pro</td>
<td>565</td>
<td>570</td>
</tr>
<tr>
<td>Pro Pro Val Ile Ala Ala Pro Asn Ala Pro Gln Pro Val Arg Ile Asp Asn</td>
<td>580</td>
<td>585</td>
</tr>
<tr>
<td>Pro Val Gly Gly Phe Ser Phe Ala Leu Pro Ala Gly Trp Val Glu Ser</td>
<td>595</td>
<td>600</td>
</tr>
<tr>
<td>Asp Ala Ala His Phe Asp Tyr Gly Ser Ala Leu Leu Ser Lys Thr Thr</td>
<td>610</td>
<td>615</td>
</tr>
<tr>
<td>Gly Asp Pro Pro Phe Pro Gly Gln Pro Pro Pro Val Ala Asn Asp Thr</td>
<td>625</td>
<td>630</td>
</tr>
<tr>
<td>Arg Ile Val Leu Gly Arg Leu Asp Gln Lys Leu Tyr Ala Ser Ala Glu</td>
<td>645</td>
<td>650</td>
</tr>
<tr>
<td>Ala Thr Asp Ser Lys Ala Ala Ala Arg Leu Gly Ser Asp Met Gly Glu</td>
<td>660</td>
<td>665</td>
</tr>
<tr>
<td>Phe Tyr Met Pro Tyr Pro Gly Thr Arg Ile Asn Gln Glu Thr Val Ser</td>
<td>670</td>
<td>675</td>
</tr>
<tr>
<td>Leu Asp Ala Asn Gly Val Ser Gly Ser Ala Ser Tyr Glu Val Lys</td>
<td>690</td>
<td>695</td>
</tr>
<tr>
<td>Phe Ser Asp Pro Ser Lys Pro Asn Gly Gln Ile Thr Thr Gly Val Ile</td>
<td>705</td>
<td>710</td>
</tr>
<tr>
<td>Gly Ser Pro Ala Ala Asn Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp</td>
<td>725</td>
<td>730</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:352:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 34 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:352:

GGATCCAAAC CACCGAGCGG TATGCCTGAA ACGC

(2) INFORMATION FOR SEQ ID NO:353:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 37 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: Other

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:353:

CGCTGCCAAT TCACCTCCGG AGGAAATCGT CGCGATC

(2) INFORMATION FOR SEQ ID NO:354:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1962 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:354:

CATATGGGCC ATCATACTA TCATCACGGA TCACAAACGAC CGAGGGTTC GCCTGAAACG
GGCGCGCGCG CGCGACTCGT CGCGACTCC CGCGCGTCTCG CGCGCGGTGAC GTTGCGCGAG
ACCGTACGCA CGCCGCTCTA CGCCGCTTTC AACCTGTGGG GTCCCGCCCTT TCACGAGGAG
TATCGGAACG TCACGATCA CGCTCGGCGC ACCGGTCCTCG GTGCCGGGAG CTGCCGAGGCGC
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>1</td>
</tr>
<tr>
<td>Gly</td>
<td>5</td>
</tr>
<tr>
<td>His</td>
<td>10</td>
</tr>
<tr>
<td>His</td>
<td>15</td>
</tr>
<tr>
<td>Gly</td>
<td>20</td>
</tr>
<tr>
<td>Ser</td>
<td>25</td>
</tr>
<tr>
<td>Ser</td>
<td>30</td>
</tr>
<tr>
<td>Val</td>
<td>35</td>
</tr>
<tr>
<td>Thr</td>
<td>40</td>
</tr>
<tr>
<td>Pro</td>
<td>45</td>
</tr>
<tr>
<td>Pro</td>
<td>50</td>
</tr>
<tr>
<td>Gly</td>
<td>55</td>
</tr>
<tr>
<td>Ala</td>
<td>60</td>
</tr>
<tr>
<td>Gly</td>
<td>65</td>
</tr>
<tr>
<td>Ala</td>
<td>70</td>
</tr>
<tr>
<td>Ala</td>
<td>75</td>
</tr>
<tr>
<td>Ser</td>
<td>80</td>
</tr>
<tr>
<td>Gly</td>
<td>85</td>
</tr>
<tr>
<td>Ala</td>
<td>90</td>
</tr>
<tr>
<td>Ala</td>
<td>95</td>
</tr>
</tbody>
</table>

**Informations for Seq ID No: 355:**

1. **Sequence Characteristics:**
   1. Length: 552 amino acids
   2. Type: amino acid
   3. Strandedness: single
   4. Topology: linear

2. **Molecule Type:** Protein

3. **Sequence Description:**

```plaintext
<table>
<thead>
<tr>
<th>Amino Acid</th>
<th>Position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Met</td>
<td>1</td>
</tr>
<tr>
<td>Gly</td>
<td>5</td>
</tr>
<tr>
<td>His</td>
<td>10</td>
</tr>
<tr>
<td>His</td>
<td>15</td>
</tr>
<tr>
<td>Gly</td>
<td>20</td>
</tr>
<tr>
<td>Ser</td>
<td>25</td>
</tr>
<tr>
<td>Ser</td>
<td>30</td>
</tr>
<tr>
<td>Val</td>
<td>35</td>
</tr>
<tr>
<td>Thr</td>
<td>40</td>
</tr>
<tr>
<td>Pro</td>
<td>45</td>
</tr>
<tr>
<td>Pro</td>
<td>50</td>
</tr>
<tr>
<td>Gly</td>
<td>55</td>
</tr>
<tr>
<td>Ala</td>
<td>60</td>
</tr>
<tr>
<td>Gly</td>
<td>65</td>
</tr>
<tr>
<td>Ala</td>
<td>70</td>
</tr>
<tr>
<td>Ala</td>
<td>75</td>
</tr>
<tr>
<td>Ser</td>
<td>80</td>
</tr>
<tr>
<td>Gly</td>
<td>85</td>
</tr>
<tr>
<td>Ala</td>
<td>90</td>
</tr>
<tr>
<td>Ala</td>
<td>95</td>
</tr>
</tbody>
</table>
```

**Sequence Description:**

- **Met**
- **Gly**
- **His**
- **His**
- **Gly**
- **Ser**
- **Lys**
- **Pro**
- **Pro**
- **Ser**
- **Gly**

**Sequence Id No: 355**
<table>
<thead>
<tr>
<th>100</th>
<th>105</th>
<th>110</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ala</td>
<td>Gln</td>
<td>Gln</td>
</tr>
<tr>
<td>Val</td>
<td>Asn</td>
<td>Tyr</td>
</tr>
<tr>
<td>Asn</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Gly</td>
<td>Val</td>
<td>Ser</td>
</tr>
<tr>
<td>Glu</td>
<td>His</td>
<td>Leu</td>
</tr>
<tr>
<td>Lys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>120</td>
<td>125</td>
</tr>
<tr>
<td>Leu</td>
<td>Asn</td>
<td>Gly</td>
</tr>
<tr>
<td>Lys</td>
<td>Val</td>
<td>Leu</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>Met</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gln</td>
<td>Gly</td>
</tr>
<tr>
<td>Thr</td>
<td>Ile</td>
<td>Lys</td>
</tr>
<tr>
<td>Thr</td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>135</td>
<td>140</td>
</tr>
<tr>
<td>Trp</td>
<td>Asp</td>
<td>Asp</td>
</tr>
<tr>
<td>Pro</td>
<td>Gln</td>
<td>Ile</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td>Leu</td>
</tr>
<tr>
<td>Asn</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Val</td>
<td>Asn</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>150</td>
<td>155</td>
</tr>
<tr>
<td>Gly</td>
<td>Thr</td>
<td>Ala</td>
</tr>
<tr>
<td>Val</td>
<td>Pro</td>
<td>Leu</td>
</tr>
<tr>
<td>His</td>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td>Asp</td>
<td>Gly</td>
<td>Ser</td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>170</td>
<td>175</td>
</tr>
<tr>
<td>Phe</td>
<td>Leu</td>
<td>Phe</td>
</tr>
<tr>
<td>Thr</td>
<td>Gln</td>
<td>Tyr</td>
</tr>
<tr>
<td>Leu</td>
<td>Ser</td>
<td>Lys</td>
</tr>
<tr>
<td>Gln</td>
<td>Asp</td>
<td>Pro</td>
</tr>
<tr>
<td>Gly</td>
<td>Glu</td>
<td>Trp</td>
</tr>
<tr>
<td>Gly</td>
<td></td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>185</td>
<td>190</td>
</tr>
<tr>
<td>Lys</td>
<td>Ser</td>
<td>Pro</td>
</tr>
<tr>
<td>Gly</td>
<td>Phe</td>
<td>Gly</td>
</tr>
<tr>
<td>Thr</td>
<td>Thr</td>
<td>Val</td>
</tr>
<tr>
<td>Asp</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td>Val</td>
</tr>
<tr>
<td>Pro</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>200</td>
<td>205</td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td>Gly</td>
</tr>
<tr>
<td>Asn</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Cys</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>220</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Pro</td>
<td>Gly</td>
</tr>
<tr>
<td>Cys</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>230</td>
<td>235</td>
</tr>
<tr>
<td>Ser</td>
<td>Gln</td>
<td>Arg</td>
</tr>
<tr>
<td>Gly</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td>250</td>
<td>255</td>
</tr>
<tr>
<td>Phe</td>
<td>Leu</td>
<td>Pro</td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>265</td>
<td>270</td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Met</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>275</td>
<td>280</td>
<td>285</td>
</tr>
<tr>
<td>Ala</td>
<td>Pro</td>
<td>Asp</td>
</tr>
<tr>
<td>Gly</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Ile</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Glu</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>290</td>
<td>295</td>
<td>300</td>
</tr>
<tr>
<td>Asn</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Lys</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>305</td>
<td>310</td>
<td>315</td>
</tr>
<tr>
<td>His</td>
<td>Trp</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>325</td>
<td>330</td>
<td>335</td>
</tr>
<tr>
<td>His</td>
<td>Phe</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>340</td>
<td>345</td>
<td>350</td>
</tr>
<tr>
<td>Ile</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>355</td>
<td>360</td>
<td>365</td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>370</td>
<td>375</td>
<td>380</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>385</td>
<td>390</td>
<td>400</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>405</td>
<td>410</td>
<td>415</td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>420</td>
<td>425</td>
<td>430</td>
</tr>
<tr>
<td>Asn</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Glu</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>435</td>
<td>440</td>
<td>445</td>
</tr>
<tr>
<td>Phe</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Trp</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>His</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>450</td>
<td>455</td>
<td>460</td>
</tr>
<tr>
<td>Tyr</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Ser</td>
<td>Lys</td>
<td></td>
</tr>
<tr>
<td>Thr</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>465</td>
<td>470</td>
<td>475</td>
</tr>
<tr>
<td>Gly</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Pro</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>485</td>
<td>490</td>
<td>495</td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Gln</td>
<td>Lys</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Glu</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Asp</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Lys</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>500</td>
<td>505</td>
<td>510</td>
</tr>
<tr>
<td>Ala</td>
<td>Ala</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Leu</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td>Phe</td>
<td>Tyr</td>
<td></td>
</tr>
<tr>
<td>Met</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td>Tyr</td>
<td>Pro</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>515</td>
<td>520</td>
<td>525</td>
</tr>
<tr>
<td>Gly</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Arg</td>
<td>Ile</td>
<td></td>
</tr>
<tr>
<td>Asn</td>
<td>Gln</td>
<td></td>
</tr>
<tr>
<td>Glu</td>
<td>Thr</td>
<td></td>
</tr>
<tr>
<td>Val</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td>Leu</td>
<td>Asp</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>Asn</td>
<td></td>
</tr>
<tr>
<td>Gly</td>
<td>Val</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>530</td>
<td>535</td>
<td>540</td>
</tr>
</tbody>
</table>
Ser Gly Ser Ala Ser Tyr Tyr Glu Val Lys Phe Ser Asp Pro Ser Lys
545 550 555 560
Pro Asn Gly Gln Ile Trp Thr Gly Val Ile Gly Ser Pro Ala Ala Asn
565 570 575
Ala Pro Asp Ala Gly Pro Pro Gln Arg Trp Phe Val Val Trp Leu Gly
580 585 590
Thr Ala Asn Pro Val Asp Lys Gly Ala Ala Lys Ala Leu Ala Glu
595 600 605
Ser Ile Arg Pro Leu Val Ala Pro Pro Ala Pro Ala Pro Ala Pro
610 615 620
Ala Glu Pro Ala Pro Ala Pro Ala Pro Ala Gly Glu Val Ala Pro Thr
625 630 635 640
Pro Thr Thr Pro Thr Pro Gln Arg Thr Leu Pro Ala
645 650
1. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

(a) Asp-Pro-Val-Asp-Ala-Val-Ile-Asn-Thr-Thr-Cys-Asn-Tyr-Gly-Gln-Val-Val-Ala-Ala-Leu; (SEQ ID No. 120)

(b) Ala-Val-Glu-Ser-Gly-Met-Leu-Ala-Leu-Gly-Thr-Pro-Ala-Pro-Ser; (SEQ ID No. 121)

(c) Ala-Ala-Met-Lys-Pro-Arg-Thr-Gly-Asp-Gly-Pro-Leu-Glu-Ala-Ala-Lys-Glu-Gly-Arg; (SEQ ID No. 122)

(d) Tyr-Tyr-Trp-Cys-Pro-Gly-Gln-Pro-Phe-Asp-Pro-Ala-Trp-Gly-Pro; (SEQ ID No. 123)

(e) Asp-Ile-Gly-Ser-Glu-Ser-Thr-Glu-Asp-Gln-Gln-Xaa-Ala-Val; (SEQ ID No. 124)

(f) Ala-Glu-Glu-Ser-Ile-Ser-Thr-Xaa-Glu-Xaa-Ile-Val-Pro; (SEQ ID No. 125)

(g) Asp-Pro-Glu-Pro-Ala-Pro-Pro-Val-Pro-Thr-Ala-Ala-Ser-Pro-Pro-Ser; (SEQ ID No. 126)

(h) Ala-Pro-Lys-Thr-Tyr-Xaa-Glu-Glu-Leu-Lys-Gly-Thr-Asp-Thr-Gly; (SEQ ID No. 127)

(i) Asp-Pro-Ala-Ser-Ala-Pro-Asp-Val-Pro-Thr-Ala-Ala-Gln-Leu-Thr-Ser-Leu-Leu-Asn-Ser-Leu-Ala-Asp-Pro-Asn-Val-Ser-Phe-Ala-Asn; (SEQ ID No. 128) and

(j) Ala-Pro-Glu-Ser-Gly-Ala-Gly-Leu-Gly-Gly-Thr-Val-Gln-Ala-Gly; (SEQ ID No. 136)

wherein Xaa may be any amino acid.

2. A polypeptide comprising an immunogenic portion of an *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative
substitutions and/or modifications, wherein said antigen has an N-terminal sequence selected from the group consisting of:

(a) Asp-Pro-Pro-Asp-Pro-His-Gln-Xaa-Asp-Met-Thr-Lys-Gly-Tyr-Tyr-Pro-Gly-Gly-Arg-Arg-Xaa-Phe; (SEQ ID No. 129) and

(b) Xaa-Tyr-Ile-Ala-Tyr-Xaa-Thr-Thr-Ala-Gly-Ile-Val-Pro-Gly-Lys-Ile-Asn-Val-His-Leu-Val; (SEQ ID No. 137), wherein Xaa may be any amino acid.

3. A polypeptide comprising an immunogenic portion of a soluble *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 1, 2, 4-10, 13-25, 52, 99 and 101 or a complement thereof under moderately stringent conditions.

4. A polypeptide comprising an immunogenic portion of a *M. tuberculosis* antigen, or a variant of said antigen that differs only in conservative substitutions and/or modifications, wherein said antigen comprises an amino acid sequence encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos.: 26-51, 138, 139, 163-183, 201, 240, 242-247, 253-256, 295-298, 309, 316, 318-320, 322, 324, 328, 329, 333, 335, 337, 339 and 341, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos.: 26-51, 138, 139, 163-183, 201, 240, 242-247, 253-256, 295-298, 309, 316, 318-320, 322, 324, 328, 329, 333, 335, 337, 339 and 341 or a complement thereof under moderately stringent conditions.

5. A DNA molecule comprising a nucleotide sequence encoding a polypeptide according to any one of claims 1-4.

6. An expression vector comprising a DNA molecule according to claim 5.
7. A host cell transformed with an expression vector according to claim 6.

8. The host cell of claim 7 wherein the host cell is selected from the group consisting of E. coli, yeast and mammalian cells.

9. A pharmaceutical composition comprising one or more polypeptides according to any one of claims 1-4 and a physiologically acceptable carrier.

10. A pharmaceutical composition comprising one or more DNA molecules according to claim 5 and a physiologically acceptable carrier.


12. A vaccine comprising one or more polypeptides according to any one of claims 1-4 and a non-specific immune response enhancer.

13. A vaccine comprising:
   a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
   a non-specific immune response enhancer.

14. A vaccine comprising:
a non-specific immune response enhancer.

15. The vaccine of claims 12-14 wherein the non-specific immune response enhancer is an adjuvant.

16. A vaccine comprising one or more DNA molecules according to claim 5 and a non-specific immune response enhancer.


18. The vaccine of claims 16 or 17 wherein the non-specific immune response enhancer is an adjuvant.

19. A method for inducing protective immunity in a patient, comprising administering to a patient a pharmaceutical composition according to any one of claims 9-11.

20. A method for inducing protective immunity in a patient, comprising administering to a patient a vaccine according to any one of claims 12-18.

21. A fusion protein comprising two or more polypeptides according to any one of claims 1-4.

22. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and ESAT-6.

23. A fusion protein comprising one or more polypeptides according to any one of claims 1-4 and the *M. tuberculosis* antigen 38 kD (SEQ ID NO:155).
24. A pharmaceutical composition comprising a fusion protein according to any one of claims 21-23 and a physiologically acceptable carrier.

25. A vaccine comprising a fusion protein according to any one of claims 21-23 and a non-specific immune response enhancer.

26. The vaccine of claim 25 wherein the non-specific immune response enhancer is an adjuvant.

27. A method for inducing protective immunity in a patient, comprising administering to a patient a pharmaceutical composition according to claim 24.

28. A method for inducing protective immunity in a patient, comprising administering to a patient a vaccine according to claims 25 or 26.

29. A method for detecting tuberculosis in a patient, comprising:
   (a) contacting dermal cells of a patient with one or more polypeptides according to any one of claims 1-4; and
   (b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

30. A method for detecting tuberculosis in a patient, comprising:
   (a) contacting dermal cells of a patient with a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and
   (b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

31. A method for detecting tuberculosis in a patient, comprising:
   (a) contacting dermal cells of a patient with one or more polypeptides encoded by a DNA sequence selected from the group consisting of SEQ ID Nos.: 3, 11, 12, 140, 141, 156-160, 189-193, 199, 200, 203, 215-225, 237, 239, 261-276, 292, 293, 303-308.

(b) detecting an immune response on the patient's skin and therefrom detecting tuberculosis in the patient.

32. The method of any one of claims 29-31 wherein the immune response is induration.

33. A diagnostic kit comprising:

(a) a polypeptide according to any one of claims 1-4; and

(b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.

34. A diagnostic kit comprising:

(a) a polypeptide having an N-terminal sequence selected from the group consisting of sequences recited in SEQ ID NO: 134 and 135; and

(b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.

35. A diagnostic kit comprising:


(b) apparatus sufficient to contact said polypeptide with the dermal cells of a patient.
36. A diagnostic kit comprising:
   (a) a fusion protein according to any one of claims 21-23; and
   (b) apparatus sufficient to contact said fusion protein with the dermal cells of a
        patient.

37. A fusion protein according to claim 23 comprising an amino acid
    sequence selected from the group consisting of sequences recited in SEQ ID NO: 153, 209,
    351 and 355.
D160 T Cell Proliferation

D160 IFNγ

FIG. 1B
**Proliferation**

![Graph of Proliferation](image)

**IFN gamma production**

![Graph of IFN gamma production](image)

**FIG. 2**
T cell clone 131TbH9 responds poorly to CSP
T Cell Clone PPD 800-10 IFNg Production

FIG. 4B
D131 T Cell Proliferation

Antigen µg/ml

O.D. 450-570

FIGS. 5 A-B
D184 T Cell Proliferation

CPM Incorporated

Antigen µg/ml

D184 IFNg

OD. 450-570

Antigen µg/ml

FIGS. 6 A-B
D201 T Cell Proliferation

CPM Incorporated

Antigen μg/ml

D201 IFNg

O.D. 450-570

Antigen μg/ml

FIGS. 7.A-B
FIGS. 8A-B
FIGS. 9A-B
Tuberculosis: Protection of Cynomolgus Monkeys with Recombinant Antigens of Mtb

FIG. 10
Aerosol TB challenge of vaccinated guinea pigs
DNA Immunized mice challenged with aerosol TB (lung CFU)

**FIG. 12**