<table>
<thead>
<tr>
<th>International Application Number:</th>
<th>WO 97/33909</th>
</tr>
</thead>
<tbody>
<tr>
<td>International Filing Date:</td>
<td>14 March 1997 (14.03.97)</td>
</tr>
<tr>
<td>Priority Date:</td>
<td>08/616,745 15 March 1996 (15.03.96) US</td>
</tr>
<tr>
<td></td>
<td>08/633,840 11 April 1996 (11.04.96) US</td>
</tr>
<tr>
<td>Applicant: CORIXA CORPORATION (US); Suite 464, 1124 Columbia Street, Seattle, WA 98104 (US).</td>
<td></td>
</tr>
<tr>
<td>Agents: MAKI, David, J. et al.; Seed and Berry L.L.P., 6300 Columbia Center, 701 Fifth Avenue, Seattle, WA 98104-7092 (US).</td>
<td></td>
</tr>
<tr>
<td>Title: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY AND IMMUNODIAGNOSIS OF PROSTATE CANCER</td>
<td></td>
</tr>
<tr>
<td>Abstract</td>
<td></td>
</tr>
</tbody>
</table>

Compounds and methods for treating and diagnosing prostate cancer are provided. The inventive compounds include polypeptides containing at least a portion of a prostate protein. Vaccines and pharmaceutical compositions for immunotherapy of prostate cancer comprising such polypeptides or DNA molecules encoding such polypeptides are also provided. The inventive polypeptides may also be used to generate antibodies useful for the diagnosis and monitoring of prostate cancer. Nucleic acid sequences for preparing probes, primers, and polypeptides are also provided.
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></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AM</td>
<td>Armenia</td>
<td>GB</td>
<td>United Kingdom</td>
<td>MW</td>
<td>Malawi</td>
</tr>
<tr>
<td>AT</td>
<td>Austria</td>
<td>GE</td>
<td>Georgia</td>
<td>MX</td>
<td>Mexico</td>
</tr>
<tr>
<td>AU</td>
<td>Australia</td>
<td>GN</td>
<td>Guinea</td>
<td>NE</td>
<td>Niger</td>
</tr>
<tr>
<td>BB</td>
<td>Barbados</td>
<td>GR</td>
<td>Greece</td>
<td>NL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>BE</td>
<td>Belgium</td>
<td>HU</td>
<td>Hungary</td>
<td>NO</td>
<td>Norway</td>
</tr>
<tr>
<td>BF</td>
<td>Burkina Faso</td>
<td>IE</td>
<td>Ireland</td>
<td>NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td>BG</td>
<td>Bulgaria</td>
<td>IT</td>
<td>Italy</td>
<td>PL</td>
<td>Poland</td>
</tr>
<tr>
<td>BJ</td>
<td>Benin</td>
<td>JP</td>
<td>Japan</td>
<td>PT</td>
<td>Portugal</td>
</tr>
<tr>
<td>BR</td>
<td>Brazil</td>
<td>KE</td>
<td>Kenya</td>
<td>RO</td>
<td>Romania</td>
</tr>
<tr>
<td>BY</td>
<td>Belarus</td>
<td>KG</td>
<td>Kyrgyzstan</td>
<td>RU</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>CA</td>
<td>Canada</td>
<td>KP</td>
<td>Democratic People's Republic of Korea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CF</td>
<td>Central African Republic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CG</td>
<td>Congo</td>
<td>KR</td>
<td>Republic of Korea</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
<td>KZ</td>
<td>Kazakhstan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CI</td>
<td>Côte d'Ivoire</td>
<td>LI</td>
<td>Liechtenstein</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CM</td>
<td>Cameroon</td>
<td>LK</td>
<td>Sri Lanka</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>China</td>
<td>LR</td>
<td>Liberia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td>Czechoslovakia</td>
<td>LT</td>
<td>Lithuania</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CZ</td>
<td>Czech Republic</td>
<td>LU</td>
<td>Luxembourg</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
<td>LV</td>
<td>Latvia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
<td>MC</td>
<td>Monaco</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EE</td>
<td>Estonia</td>
<td>MD</td>
<td>Republic of Moldova</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ES</td>
<td>Spain</td>
<td>MG</td>
<td>Madagascar</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FI</td>
<td>Finland</td>
<td>ML</td>
<td>Mali</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FR</td>
<td>France</td>
<td>MN</td>
<td>Mongolia</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>Gabon</td>
<td>MR</td>
<td>Mauritania</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SG</td>
<td>Singapore</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI</td>
<td>Slovenia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SK</td>
<td>Slovakia</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SN</td>
<td>Senegal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SZ</td>
<td>Switzerland</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TD</td>
<td>Chad</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TG</td>
<td>Togo</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TJ</td>
<td>Tajikistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>Trinidad and Tobago</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UA</td>
<td>Ukraine</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UG</td>
<td>Uganda</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>US</td>
<td>United States of America</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UZ</td>
<td>Uzbekistan</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VN</td>
<td>Viet Nam</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
COMPONENTS AND METHODS FOR IMMUNOTHERAPY
AND IMMUNODIAGNOSIS OF PROSTATE CANCER

TECHNICAL FIELD

The present invention relates generally to the treatment, diagnosis and monitoring of prostate cancer. The invention is more particularly related to polypeptides comprising at least a portion of a prostate protein. Such polypeptides may be used in vaccines and pharmaceutical compositions for treatment of prostate cancer. The polypeptides may also be used for the production of compounds, such as antibodies, useful for diagnosing and monitoring the progression of prostate cancer, and possibly other tumor types, in a patient.

BACKGROUND OF THE INVENTION

Prostate cancer is the most common form of cancer among males, with an estimated incidence of 30% in men over the age of 50. Overwhelming clinical evidence shows that human prostate cancer has the propensity to metastasize to bone, and the disease appears to progress inevitably from androgen dependent to androgen refractory status, leading to increased patient mortality. This prevalent disease is currently the second leading cause of cancer death among men in the U.S.

In spite of considerable research into therapies for the disease, prostate cancer remains difficult to treat. Commonly, treatment is based on surgery and/or radiation therapy, but these methods are ineffective in a significant percentage of cases. Three prostate specific proteins - prostate specific antigen (PSA) and prostatic acid phosphatase (PAP) - have limited diagnostic and therapeutic potential. PSA levels do not always correlate well with the presence of prostate cancer, being positive in a percentage of non-prostate cancer cases, including benign prostatic hyperplasia (BPH). Furthermore, PSA measurements correlate with prostate volume, and do not indicate the level of metastasis.

Accordingly, there remains a need in the art for improved vaccines and diagnostic methods for prostrate cancer.
SUMMARY OF THE INVENTION

The present invention provides compounds and methods for immunotherapy and diagnosis of prostate cancer. In one aspect, polypeptides are provided comprising at least an immunogenic portion of a prostate protein having a partial sequence as provided in SEQ ID Nos. 2 and 4-8, or a variant of such a protein that differs only in conservative substitutions and/or modifications.

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 preferred embodiments, the host cells are selected from the group consisting of *E. coli*, yeast and mammalian cells.

The present invention also provides pharmaceutical compositions comprising one or more of the polypeptides of SEQ ID Nos. 1-8, 20, 21, 25-31 or 44-57, or nucleic acids of SEQ ID Nos. 9-19, 22-24 or 32-43, and a physiologically acceptable carrier. The invention also provides vaccines comprising one or more of such polypeptides or nucleic acids in combination with a non-specific immune response enhancer.

In yet another aspect, methods are provided for inhibiting the development of prostate cancer in a patient, comprising administering an effective amount of one or more of the polypeptides of SEQ ID Nos. 1-8, 20, 21, 25-31 or 44-57, or nucleic acids of SEQ ID Nos. 9-19, 22-24 or 32-43 to a patient in need thereof.

In further aspects, methods are provided for detecting prostate cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that is capable of binding to a polypeptide of SEQ ID Nos. 1-8, 20, 21, 25-31 or 44-57; and (b) detecting in the sample a protein or polypeptide that binds to the binding agent.

In related aspects, methods are provided for monitoring the progression of prostate cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that is capable of binding to a polypeptide of SEQ
ID Nos. 1-8, 20, 21, 25-31 or 44-57; (b) determining in the sample an amount of a protein or polypeptide that binds to the binding agent; (c) repeating steps (a) and (b); and comparing the amounts of polypeptide detected in steps (b) and (c).

Within related aspects, the present invention provides antibodies, preferably monoclonal antibodies, that bind to the polypeptides described above, as well as diagnostic kits comprising such antibodies, and methods of using such antibodies to inhibit the development of prostate cancer.

The present invention also provides methods for detecting prostate cancer comprising: (a) obtaining a biological sample from a patient; (b) contacting the sample with at least two oligonucleotide primers in a polymerase chain reaction, at least one of the oligonucleotide primers being specific for a DNA sequence selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43; and (c) detecting in the sample a DNA sequence that amplifies in the presence of the oligonucleotide primer. In one embodiment, the oligonucleotide primer comprises at least about 10 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43.

In a further aspect, the present invention provides a method for detecting prostate cancer in a patient comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with an oligonucleotide probe specific for a DNA sequence selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43; and (c) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe. In one embodiment, the oligonucleotide probe comprises at least about 15 contiguous nucleotides of a DNA sequence selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43.

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

Fig. 1 illustrates a Western blot analysis of sera obtained from rats immunized with rat prostate extract.

Fig. 2 illustrates a non-reduced SDS PAGE of the rat immunizing preparation of Fig. 1.

Fig. 3 illustrates the binding of a putative human homologue of rat steroid binding protein to progesterone and to estramustine.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the immunotherapy, diagnosis and monitoring of prostate cancer. The inventive compositions are generally polypeptides that comprise at least a portion of a human prostate protein, the protein demonstrating immunoreactivity with human prostate sera. Also included within the present invention are molecules (such as an antibody or fragment thereof) that bind to the inventive polypeptides. Such molecules are referred to herein as "binding agents."

In particular, the subject invention discloses polypeptides comprising at least a portion of a human prostate protein provided in SEQ ID Nos. 2 and 4-8, or a variant of such a protein that differs only in conservative substitutions and/or modifications. As used herein, the term "polypeptide" encompasses amino acid chains of any length, including full length proteins, wherein the amino acid residues are linked by covalent peptide bonds. Thus, a polypeptide comprising a portion of one of the above prostate proteins may consist entirely of the portion, or the portion may be present within a larger polypeptide that contains additional sequences. The additional sequences may be derived from the native protein or may be heterologous, and such sequences may be immunoreactive and/or antigenic.

As used herein, an "immunogenic portion" of a human prostate protein is a portion that reacts either with sera derived from an individual inflicted with autoimmune prostatitis or with sera derived from a rat model of autoimmune prostatitis.

In other words, an immunogenic portion is capable of eliciting an immune response and
as such binds to antibodies present within prostatitis sera. Autoimmune prostatitis may occur, for example, following treatment of bladder cancer by administration of Bacillus Calmette-Guerin (BCG), an avirulent strain of *Mycobacterium bovis*. In the rat model of autoimmune prostatitis, rats are immunized with a detergent extract of rat prostate. Sera from either of these sources may be used to react with the human prostate derived polypeptides described herein. Antibody binding assays may generally be performed using any of a variety of means known to those of ordinary skill in the art, as described, for example, in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988. For example, a polypeptide may be immobilized on a solid support (as described below) and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide. Unbound sera may then be removed and bound antibodies detected using, for example, ^125^I-labeled Protein A.

A "variant," as used herein, is a polypeptide that differs from the recited polypeptide only in conservative substitutions and/or modifications, such that the immunotherapeutic, antigenic and/or diagnostic properties of the polypeptide or molecules that bind to the polypeptide, are retained. For prostate proteins with immunoreactive properties, variants may generally be identified by modifying one of the above polypeptide sequences, and evaluating the immunoreactivity of the modified polypeptide. For prostate proteins 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 prostate cancer. 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.

Polypeptides having one of the sequences provided in SEQ ID Nos. 1 to 8, 20, 21 and 25-31 may be isolated from a suitable human prostate adenocarcinoma cell line, such as LnCap.fgc (ATCC No. 1740-CRL). LnCap.fgc is a prostate adenocarcinoma cell line that is a particularly good representation of human prostate cancer. Like the human cancer, LnCap.fgc cells form progressively growing tumors as xenografts in SCID mice, respond to testosterone, secrete PSA and respond to the presence of bone marrow components (e.g., transferrin). In particular, the polypeptides may be isolated by expression screening of a LnCap.fgc cDNA library with human prostatitis sera using techniques described, for example, in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY (and references cited therein), and as described in detail below. The polypeptides of SEQ ID No. 48 and 49 may be isolated from the LnCap/fgc cell line by screening with sera from the rat model of autoimmune prostatitis discussed above. The polypeptides of SEQ ID Nos. 50-56 may be isolated from the LnCap/fgc cell line by screening with human prostatitis sera as described in detail in Example 4. The polypeptides of SEQ ID No. 44-47 may be isolated from human seminal fluid as described in detail in Example 2. Once a DNA sequence encoding a polypeptide is obtained, any of the above modifications may be readily introduced using standard mutagenesis techniques, such as oligonucleotide-directed site-specific mutagenesis.
The polypeptides disclosed herein may also 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.

Alternatively, any of the above polypeptides may be produced recombinantly by inserting a DNA sequence that encodes the polypeptide into an expression vector and expressing the protein in an appropriate host. 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 CHO cells. The DNA sequences expressed in this manner may encode naturally occurring polypeptides, portions of naturally occurring polypeptides, or other variants thereof.

In general, regardless of the method of preparation, the polypeptides disclosed herein are prepared in substantially pure form (i.e., the polypeptides are homogenous as determined by amino acid composition and primary sequence analysis). Preferably, the polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. In certain preferred embodiments, described in more detail below, the substantially pure polypeptides are incorporated into pharmaceutical compositions or vaccines for use in one or more of the methods disclosed herein.
Polypeptides of the present invention that comprise an immunogenic portion of a prostate protein may generally be used for immunotherapy of prostate cancer, wherein the polypeptide stimulates the patient's own immune response to prostate tumor cells. In further aspects, the present invention provides methods for using one or more of the immunoreactive polypeptides of SEQ ID Nos. 1 to 8, 20, 21, 25-31 and 44-57 (or DNA encoding such polypeptides) for immunotherapy of prostate cancer 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. Accordingly, the above immunoreactive polypeptides may be used to treat prostate cancer or to inhibit the development of prostate cancer. The polypeptides may be administered either prior to or following surgical removal of primary tumors and/or treatment by administration of radiotherapy and conventional chemotherapeutic drugs.

In these aspects, the polypeptide 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. The vaccines may comprise one or more of such polypeptides and a non-specific immune response enhancer, such as an adjuvant, biodegradable microsphere (e.g., polylactic galactide) or a liposome (into which the polypeptide is incorporated). Pharmaceutical compositions and vaccines may also contain other epitopes of prostate cell antigens, either incorporated into a combination polypeptide (i.e., a single polypeptide that contains multiple epitopes) or present within a separate polypeptide.

Alternatively, a pharmaceutical composition or vaccine may contain DNA encoding one or more of the above polypeptides, such that the polypeptide is generated in situ. In such pharmaceutical compositions and 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, bacteria and viral expression systems. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter). Bacterial delivery
systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an epitope of a prostate cell antigen on its cell surface. In a preferred embodiment, the DNA may be introduced using a viral expression system (e.g., vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch et al., *PNAS* 86:317-321, 1989; Flexner et al., *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner et al., *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld et al., *Science* 252:431-434, 1991; Kolls et al., *PNAS* 91:215-219, 1994; Kass-Eisler et al., *PNAS* 90:11498-11502, 1993; Guzman et al., *Circulation* 88:2838-2848, 1993; and Guzman et al., *Cir. Res.* 73:1202-1207, 1993. 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 published PCT application WO 90/11092, and Ulmer et al., *Science* 259:1745-1749, 1993, reviewed by Cohen, *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.

Routes and frequency of administration, as well as dosage, will vary from individual to individual and may parallel those currently being used in immunotherapy of other diseases. 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 10 doses may be administered over a 3-24 week period. Preferably, 4 doses are administered, at an interval of 3 months, and booster administrations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of polypeptide or DNA that is effective to raise an immune response (cellular and/or humoral) against prostate tumor cells in a treated patient. A suitable immune response is at least 10-50% above the basal (i.e., untreated) level. 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 0.01 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 and/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/or magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactic glycolide) 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 non-specific immune response enhancers may be employed in the vaccines of this invention. For example, an adjuvant may be included. 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 response, such as lipid A, Bordella pertussis or Mycobacterium tuberculosis.

Such adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI) and Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ).

Polypeptides disclosed herein may also be employed in ex vivo treatment of prostate cancer. For example, cells of the immune system, such as T cells, may be isolated from the peripheral blood of a patient, using a commercially available cell separation system, such as CellPro Incorporated's (Bothell, WA) CEPRATE™ system (see U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). The separated cells are stimulated with one or more of the immunoreactive polypeptides contained within a delivery vehicle, such as a microsphere, to provide antigen-specific T cells. The population of tumor antigen-
specific T cells is then expanded using standard techniques and the cells are administered back to the patient.

Polypeptides of the present invention may also, or alternatively, be used to generate binding agents, such as antibodies or fragments thereof, that are capable of detecting metastatic human prostate tumors.

Binding agents of the present invention may generally be prepared using methods known to those of ordinary skill in the art, including the representative procedures described herein. Binding agents are capable of differentiating between patients with and without prostate cancer, using the representative assays described herein. In other words, antibodies or other binding agents raised against a prostate protein, or a suitable portion thereof, will generate a signal indicating the presence of primary or metastatic prostate cancer in at least about 20% of patients afflicted with the disease, and will generate a signal indicating the absence of the disease in at least about 90% of individuals without primary or metastatic prostate cancer. Suitable portions of such prostate proteins are portions that are able to generate a binding agent that indicates the presence of primary or metastatic prostate cancer in substantially all (i.e., at least about 80%, and preferably at least about 90%) of the patients for which prostate cancer would be indicated using the full length protein, and that indicate the absence of prostate cancer in substantially all of those samples that would be negative when tested with full length protein. The representative assays described below, such as the two-antibody sandwich assay, may generally be employed for evaluating the ability of a binding agent to detect metastatic human prostate tumors.

The ability of a polypeptide prepared as described herein to generate antibodies capable of detecting primary or metastatic human prostate tumors may generally be evaluated by raising one or more antibodies against the polypeptide (using, for example, a representative method described herein) and determining the ability of such antibodies to detect such tumors in patients. This determination may be made by assaying biological samples from patients with and without primary or metastastic prostate cancer for the presence of a polypeptide that binds to the generated antibodies. Such test assays may be performed, for example, using a representative procedure
described below. Polypeptides that generate antibodies capable of detecting at least 20% of primary or metastatic prostate tumors by such procedures are considered to be able to generate antibodies capable of detecting primary or metastatic human prostate tumors. Polypeptide specific antibodies may be used alone or in combination to improve sensitivity.

Polypeptides capable of detecting primary or metastatic human prostate tumors may be used as markers for diagnosing prostate cancer or for monitoring disease progression in patients. In one embodiment, prostate cancer in a patient may be diagnosed by evaluating a biological sample obtained from the patient for the level of one or more of the above polypeptides, relative to a predetermioed cut-off value. As used herein, suitable "biological samples" include blood, sera, urine and/or prostate secretions.

The level of one or more of the above polypeptides may be evaluated using any binding agent specific for the polypeptide(s). A "binding agent," in the context of this invention, is any agent (such as a compound or a cell) that binds to a polypeptide as described above. As used herein, "binding" refers to a noncovalent association between two separate molecules (each of which may be free (i.e., in solution) or present on the surface of a cell or a solid support), such that a "complex" is formed. Such a complex may be free or immobilized (either covalently or noncovalently) on a support material. The ability to bind may generally be evaluated by determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind" in the context of the present invention when the binding constant for complex formation exceeds about $10^3$ L/mol. The binding constant may be determined using methods well known to those of ordinary skill in the art.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome with or without a peptide component, an RNA molecule or a peptide. In a preferred embodiment, the binding partner is an antibody, or a fragment thereof. Such antibodies may be polyclonal, or monoclonal. In
addition, the antibodies may be single chain, chimeric, CDR-grafted or humanized. Antibodies may be prepared by the methods described herein and by other methods well known to those of skill in the art.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding partner to detect polypeptide markers in a sample. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In a preferred embodiment, the assay involves the use of binding partner immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a second binding partner that contains a reporter group. Suitable second binding partners include antibodies that bind to the binding partner/polypeptide complex. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding partner after incubation of the binding partner with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding partner is indicative of the reactivity of the sample with the immobilized binding partner.

The solid support may be any material known to those of ordinary skill in the art to which the antigen may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which may be a direct linkage between the antigen and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent.
in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μg, and preferably about 100 ng to about 1 μg, is sufficient to immobilize an adequate amount of binding agent.

Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group on the support with an amine and an active hydrogen on the binding partner (see, e.g., Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a second antibody (containing a reporter group) capable of binding to a different site on the polypeptide is added. The amount of second antibody that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact
time (i.e., incubation time) is that period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with prostate cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include enzymes (such as horseradish peroxidase), substrates, cofactors, inhibitors, dyes, radionuclides, luminescent groups, fluorescent groups and biotin. The conjugation of antibody to reporter group may be achieved using standard methods known to those of ordinary skill in the art.

The second antibody is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound second antibody is then removed and bound second antibody is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of prostate cancer, the signal detected from the reporter group that remains bound to the solid support is generally
compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without prostate cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for prostate cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett et al., *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (i.e., sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (i.e., the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for prostate cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the antibody is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized antibody as the sample passes through the membrane. A second, labeled antibody then binds to the antibody-polypeptide complex as a solution containing the second antibody flows through the membrane. The detection of bound second antibody may then be performed as described above. In the strip test format, one end of the membrane to which antibody is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second antibody and to the area of immobilized antibody. Concentration of second antibody at the area of immobilized antibody indicates the presence of prostate cancer. Typically, the concentration of second antibody at that site generates a pattern, such as a line, that can
be read visually. The absence of such a pattern indicates a negative result. In general, the amount of antibody immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 µg, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the antigens or antibodies of the present invention. The above descriptions are intended to be exemplary only.

In another embodiment, the above polypeptides may be used as markers for the progression of prostate cancer. In this embodiment, assays as described above for the diagnosis of prostate cancer may be performed over time, and the change in the level of reactive polypeptide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, prostate cancer is progressing in those patients in whom the level of polypeptide detected by the binding agent increases over time. In contrast, prostate cancer is not progressing when the level of reactive polypeptide either remains constant or decreases with time.

Antibodies for use in the above methods may be prepared by any of a variety of techniques known to those of ordinary skill in the art. See, e.g., Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In one such technique, an immunogen comprising the antigenic polypeptide is initially injected into any of a wide variety of mammals (e.g., mice, rats, rabbits, sheep and goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating
one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for the antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Monoclonal antibodies of the present invention may also be used as therapeutic reagents, to diminish or eliminate prostate tumors. The antibodies may be
used on their own (for instance, to inhibit metastases) or coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include $^{90}$Y, $^{125}$I, $^{131}$I, $^{186}$Re, $^{188}$Re, $^{211}$At, and $^{212}$Bi. Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, Pseudomonas exotoxin, Shigella toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (e.g., covalently bonded) to a suitable monoclonal antibody either directly or indirectly (e.g., via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulphydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (e.g., a halide) on the other.

Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulphydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, e.g., U.S. Patent No. 4,671,958, to Rodwell et al.

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of
different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (e.g., U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (e.g., U.S. Patent No. 4,625,014, to Senter et al.), by hydrolysis of derivatized amino acid side chains (e.g., U.S. Patent No. 4,638,045, to Kohn et al.), by serum complement-mediated hydrolysis (e.g., U.S. Patent No. 4,671,958, to Rodwell et al.), and acid-catalyzed hydrolysis (e.g., U.S. Patent No. 4,569,789, to Blattler et al.).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers which provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (e.g., U.S. Patent No. 4,507,234, to Kato et al.), peptides and polysaccharides such as aminodextran (e.g., U.S. Patent No. 4,699,784, to Shih et al.). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (e.g., U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison et al. discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the
precise does of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

Diagnostic reagents of the present invention may also comprise DNA sequences encoding one or more of the above polypeptides, or one or more portions thereof. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify prostate tumor-specific cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for a DNA molecule encoding a polypeptide of the present invention. The presence of the amplified cDNA is then detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes specific for a DNA molecule encoding a polypeptide of the present invention may be used in a hybridization assay to detect the presence of an inventive polypeptide in a biological sample.

As used herein, the term “oligonucleotide primer/probe specific for a DNA molecule” means an oligonucleotide sequence that has at least about 80% identity, preferably at least about 90% and more preferably at least about 95%, identity to the DNA molecule in question. Oligonucleotide primers and/or probes which may be usefully employed in the inventive diagnostic methods preferably have at least about 10-40 nucleotides. In a preferred embodiment, the oligonucleotide primers comprise at least about 10 contiguous nucleotides of a DNA molecule encoding one of the polypeptides disclosed herein. Preferably, oligonucleotide probes for use in the inventive diagnostic methods comprise at least about 15 contiguous oligonucleotides of a DNA molecule encoding one of the polypeptides disclosed herein. Techniques for both PCR based assays and hybridization assays are well known in the art (see, for example, Mullis et al. Ibid; Ehrlich, Ibid). Primers or probes may thus be used to detect prostate and/or prostate tumor sequences in biological samples, preferably blood, semen or prostate and/or prostate tumor tissue.

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

Example 1

A. Isolation of Polypeptides from LnCap.fgc using human prostatitis sera

Representative polypeptides of the present invention were isolated by screening a human prostate cancer cell line with human prostatitis sera as follows. A human prostate adenocarcinoma cDNA expression library was constructed by reverse transcriptase synthesis from mRNA purified from the human prostate adenocarcinoma cell line LnCap.fgc (ATCC No. 1740-CRL), followed by insertion of the resulting cDNA clones in Lambda ZAP II (Stratagene, La Jolla, CA).

Human prostatitis serum was obtained from a patient diagnosed with autoimmune prostatitis following treatment of bladder carcinoma by administration of BCG. This serum was used to screen the LnCap cDNA library as described in Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. Specifically, LB plates were overlaid with approximately $10^4$ pfu of the LnCap cDNA library and incubated at 42°C for 4 hours prior to obtaining a first plaque lift on isopropylthio-beta-galactoside (IPTG) impregnated nitrocellulose filters. The plates were then incubated for an additional 5 hours at 42°C and a second plaque lift was prepared by incubation overnight at 37°C. The filters were washed three times with PBS-T, blocked for 1 hour with PBS (containing 1% Tween 20™) and again washed three times with PBS-T, prior to incubation with human prostatitis sera at a dilution of 1:200 with agitation overnight. The filters were then washed three times with PBS-T and incubated with $^{125}$I-labeled Protein A (1 μl/15 ml PBS-T) for 1 hour with agitation. Filters were exposed to film for variable times, ranging from 16 hours to 7 days. Plaques giving signals on duplicate lifts were re-plated on LB plates. Resulting plaques were lifted with duplicate filters and these filters were treated as above. The filters were incubated with human prostatitis sera (1:200 dilution) at 4°C with agitation overnight. Positive plaques were
visualized with $^{125}$I-Protein A as described above with the filters being exposed to film for variable times, ranging from 16 hours to 11 days. In vivo excision of positive human prostatitis antigen cDNA clones was performed according to the manufacturer’s protocol.

B. Characterization of Polypeptides

DNA sequence for positive clones was obtained using forward and reverse primers on an Applied Biosystems Inc. Automated Sequence Model 373A (Foster City, CA). The cDNA sequences encoding the isolated polypeptides, hereinafter referred to as HPA8, HPA13, HPA15 - HPA17, HPA20, HPA25, HPA28, HPA29, HPA32 - HPA38 and HPA41 are presented in SEQ ID Nos. 32 and 33, 34 and 35, 36, 9 and 10, 11, 12, 13 and 14, 15, 37 and 38, 16, 39, 22 and 23, 17 and 18, 19, 24, 40 and 41, 42 and 43, respectively. The 3' sequences of HPA16 and HPA20 are identical. HPA13, HPA16, HPA20, HPA29 and HPA33 are believed to be overlapping clones with novel 5' end points. Two of the positive clones were determined to be identical to HPA15. Also, HPA15, HPA34 and HPA37 were found to be overlapping clones. The expected N-terminal amino acid sequences of the isolated polypeptides HPA16, HPA17, HPA20, HPA25, HPA28, HPA32, HPA35, HPA36, HPA34, HPA37, HPA8, HPA13, HPA15, HPA29, HPA33, HPA38 and HPA41, based on the determined cDNA sequences in frame with the N-terminal portion of β-galactosidase (lacZ) are presented in SEQ ID Nos. 1-8, 20, 21 and 25-31, respectively.

The determined cDNA and expected amino acid sequences for the isolated polypeptides were compared to known sequences in the gene bank using the EMBL and GenBank (Release 91) databases, and also the DNA STAR system. The DNA STAR system is a combination of the Swiss, PIR databases along with translated protein sequences (Release 91). No significant homologies to HPA17, HPA25, HPA28, HPA32, HPA35 and HPA36 were found.

The determined cDNA sequence for HPA8 was found to have approximately 100% identity with the human proto-oncogene BMI-1 (Alkema, M.J. et al., *Hum. Mol. Gen.* 2:1597-1603, 1993). Search of the DNA database with 5' and 3'
cDNA sequence encoding HPA13 revealed 100% identity with a known cDNA sequence from a human immature myeloid cell line (GenBank Acc. No. D63880). Search of the protein database with the deduced amino acid sequence for HPA13 revealed 100% identity with the open reading frame encoded by the same human cDNA sequence. Search of the protein database with the expected amino acid sequence for HPA15, revealed high homology (60% identity) with a *Saccharomyces cerevisiae* predicted open reading frame (Swiss/PIR Acc. No. S46677), and 100% identity with a human protein from pituitary gland modulating intestinal fluid secretion (Lonnroth, I., *J. Biol. Chem.* 35:20615-20620, 1995). The deduced amino acid sequence for HPA38 was found to have 100% identity with human heat shock factor protein 2 (Schuetz, T. J. et al., *Proc. Natl. Acad. Sci. USA* 88:6911-6915, 1991). Search of the DNA database with the 5′ DNA sequence for HPA41 and search of the protein database with the deduced amino acid sequence revealed 100% identity with a human LIM protein (Rearden, A., *Biochem. Biophys. Res. Commun.* 201:1124-1131, 1994). To the best of the inventors' knowledge, except for LIM protein, none of the inventive polypeptides have been previously shown to be present in human prostate.

Positive phagemid viral particles 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 prostatitis sera (1:200 dilution) 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 1, wherein (+) indicates a positive reaction and (-) indicates no reaction.
<table>
<thead>
<tr>
<th>Antigen</th>
<th>Human Prostatitis Sera</th>
<th>Anti-lacZ Sera</th>
<th>Protein Mass/Kd</th>
</tr>
</thead>
<tbody>
<tr>
<td>HPA8</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA13</td>
<td>(+)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>HPA15</td>
<td>(+)</td>
<td>(+)</td>
<td>50</td>
</tr>
<tr>
<td>HPA16</td>
<td>(+)</td>
<td>(+)</td>
<td>40</td>
</tr>
<tr>
<td>HPA17</td>
<td>(+)</td>
<td>(-)</td>
<td>40</td>
</tr>
<tr>
<td>HPA20</td>
<td>(+)</td>
<td>(+)</td>
<td>38</td>
</tr>
<tr>
<td>HPA25</td>
<td>(-)</td>
<td>(+)</td>
<td>32</td>
</tr>
<tr>
<td>HPA28</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA29</td>
<td>(+)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>HPA32</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA33</td>
<td>(+)</td>
<td>(+)</td>
<td></td>
</tr>
<tr>
<td>HPA34</td>
<td>not tested</td>
<td>(+)</td>
<td>50</td>
</tr>
<tr>
<td>HPA35</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA36</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA37</td>
<td>not tested</td>
<td>(+)</td>
<td>50</td>
</tr>
<tr>
<td>HPA38</td>
<td>(-)</td>
<td>(-)</td>
<td></td>
</tr>
<tr>
<td>HPA41</td>
<td>not tested</td>
<td>(+)</td>
<td></td>
</tr>
</tbody>
</table>

Positive reaction of the recombinant human prostatitis antigens with both the human prostatitis sera and anti-lacZ sera indicate that reactivity of the human prostatitis sera is directed towards the fusion protein. Cloned antigens showing reactivity to the human prostatitis sera but not to anti-lacZ sera indicate that the reactive protein is likely initiating within the clone. Antigens reactive with the anti-lacZ sera but not with the human prostatitis sera may be the result of the human prostatitis 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. Antigens not
reactive with either sera are not being expressed in *E. coli*, and reactive epitopes may be within the fusion protein or within an internal open reading frame. Due to the instability of recombinant antigens from HPA13, HPA29 and HPA33, it was not possible to determine the size of the recombinant antigens.

The expression of representative human prostatitis antigens was investigated by RT-PCR in four different human cell lines (including two metastatic prostate tumor lines LNCaP and DU145), normal prostate, breast, colon, kidney, stomach, lung and skeletal muscle tissue, nine different prostate tumor samples and three different breast tumor samples. The results of these studies are shown in Table II.
Table II
Analysis of HPA clone mRNA expression by RT-PCR in human cell lines, normal tissues and tumors

<table>
<thead>
<tr>
<th>Clone</th>
<th>LNCaP</th>
<th>DU145</th>
<th>MCF-12A</th>
<th>HBL-100</th>
<th>Prostate</th>
<th>Breast</th>
<th>Colon</th>
<th>Kidney</th>
<th>Stomach</th>
<th>Lung</th>
<th>Skel. Muscle</th>
</tr>
</thead>
<tbody>
<tr>
<td>hpa-17</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>hpa-20</td>
<td>+++</td>
<td>++++</td>
<td>NT</td>
<td>NT</td>
<td>+</td>
<td>NT</td>
<td>NT</td>
<td>NT</td>
<td>-</td>
<td>+</td>
<td>NT</td>
</tr>
<tr>
<td>hpa-28</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>

Prostate Tumors (n=9)

<table>
<thead>
<tr>
<th>Clone</th>
<th>Tumor 1</th>
<th>Tumor 2</th>
<th>Tumor 3</th>
<th>Tumor 4</th>
<th>Tumor 5</th>
<th>Tumor 6</th>
<th>Tumor 7</th>
<th>Tumor 8</th>
<th>Tumor 9</th>
<th>Tumor 1</th>
<th>Tumor 2</th>
<th>Tumor 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>hpa-17</td>
<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>hpa-20</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>
<td>+</td>
<td>+</td>
<td>++</td>
</tr>
<tr>
<td>hpa-28</td>
<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>
mRNA expression of representative antigens in LNCaP and normal prostate, kidney, liver, stomach, lung and pancreas was also investigated by RNase protection. The results of these studies are provided in Table III.

Table III
Analysis of HPA clone mRNA expression by RNase protection in LNCaP and normal human tissues

<table>
<thead>
<tr>
<th>Clone</th>
<th>LNCaP</th>
<th>Prostate</th>
<th>Kidney</th>
<th>Liver</th>
<th>Stomach</th>
<th>Lung</th>
<th>Pancreas</th>
</tr>
</thead>
<tbody>
<tr>
<td>hpa-15</td>
<td>-</td>
<td>-</td>
<td>++</td>
<td>++</td>
<td>+</td>
<td>-</td>
<td>++</td>
</tr>
<tr>
<td>hpa-20</td>
<td>++++</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>NT</td>
<td>NT</td>
</tr>
<tr>
<td>hpa-25</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>++</td>
<td>++</td>
<td>NT</td>
</tr>
<tr>
<td>hpa-32</td>
<td>NT</td>
<td>++</td>
<td>+</td>
<td>+</td>
<td>NT</td>
<td>++</td>
<td>NT</td>
</tr>
<tr>
<td>hpa-35</td>
<td>+++</td>
<td>+++</td>
<td>NT</td>
<td>+</td>
<td>+</td>
<td>+++</td>
<td>+</td>
</tr>
<tr>
<td>hpa-36</td>
<td>+</td>
<td>+</td>
<td>NT</td>
<td>NT</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

Example 2
A. Isolation and Characterization of Rat Steroid Binding Protein

Immune sera was obtained from rats immunized with rat prostate extract to generate antibodies to self prostate antigens. Specifically, rats were prebled to obtain control sera prior to being immunized with a detergent extract of rat prostate (in PBS containing 0.1% Triton) in Freunds complete adjuvant. A boost of incomplete Freunds adjuvant was given 3 weeks after the initial immunization and sera was harvested at 6 weeks.

The sera thus obtained was subjected to ECL. Western blot analysis (Amersham International, Arlington Heights, Ill) using the manufacturer’s protocol and a rat prostate protein was identified, as shown in Fig. 1. After reduction, SDS-PAGE revealed a broad silver staining band migrating at 7 kD. Without reduction, a strong band was seen at 24 kD (Fig. 2). This protein was purified by ion exchange
chromatography and subjected to gel electrophoresis under reduced conditions. Three bands were seen, indicating the presence of three chains within the protein: a 6-8 kD chain (C1), a 8-10 kD chain (C2) and a 10-12 kD chain (C3). The protein was further purified by reverse phase HPLC on a Delta™ C18 300 Å 5 μm column, column size 3.9 x 300 mm (Waters-Millipore, Milford, MA). The sample containing 100 μg of protein was dissolved in 0.1% trifluoroacetic acid (TFA), pH 1.9 and polypeptides were eluted with a linear gradient of acetonitrile (0-60%) in 0.1% TFA pH 1.9 at a flow rate of 0.5 mL/min for 1 hour. The eluent was monitored at 214 nm. Two peaks were obtained, a C1-C3 dimer and a C2-C3 dimer. The amino terminus of the C2 chain was found to be blocked. The C1 and C3 chains were sequenced on a Perkin Elmer/Applied Biosystems Inc. Precise Model 494 protein sequencer and found to have the following amino terminal sequences (Seq. ID Nos. 44 and 45, respectively).

(a) Ser-Gln-Ile-Cys-Glu-Leu-Val-Ala-His-Glu-Thr-Ile-Ser-Phe-Leu; and

(b) Xaa-Xaa-Xaa-Xaa-Ser-Ile-Leu-asp-Glu-Val-Ile-Arg-Gly-Thr,

wherein Xaa may be any amino acid.

These sequences were compared to known sequences in the gene bank using the databases discussed in Example 1 and were found to be identical to rat steroid binding protein, also known as estramustine-binding protein (EMBP) (Forsgren, B. et al., Prog. Clin. Biol. Res. 75A:391-407, 1981; Forsgren, B. et al., Proc. Natl. Acad. Sci. USA 76:3149-53, 1979). This protein is a major secreted protein in rat seminal fluid and has been shown to bind steroid, cholesterol and proline rich proteins. EMBP has been shown to bind estramustine and estromustine, the active metabolites of estramustine phosphate. Estramustine phosphate has been found to be clinically useful in treating advanced prostate cancer in patients who do not respond to standard hormone ablation therapy (see, for example, Van Poppel, H. et al., Prog. Clin. Biol. Res. 370:323-41, 1991).

B. Isolation of putative human homologue to rat steroid binding protein

Purified rat steroid binding protein was obtained from freshly excised rat prostate and used to subcutaneously immunize a New Zealand white virgin female
rabbit (150 μg purified rat steroid binding protein in 1 ml of PBS and 1 ml of incomplete Freund’s adjuvant containing 100 μg of muramyl dipeptide (adjuvant peptide, Calbiochem, La Jolla, CA). Six weeks later the rabbit was boosted subcutaneously with the same protein dose in incomplete Freund’s adjuvant. Finally, the rabbit was boosted intravenously two weeks later with 100 μg protein in PBS and the sera harvested two weeks after the final immunization.

The resulting rabbit antisera was used to screen the LnCap.fgc cell line without success. The rabbit antisera was subsequently used to screen human seminal fluid anion exchange chromatography pools using the protocol detailed below in Example 3. This analysis indicated an approximately 18-22 kD cross-reactive protein. The seminal fluid fraction of interest (Fraction 1) was separated into individual components by SDS-PAGE under non-reducing conditions, blotted onto a PVDF membrane, excised and digested with CNBr in 70% formic acid. The resulting CNBr fragments were resolved on a tricine gel system, again electroblotted to PVDF and excised. The sequence for one peptide was determined as follows:

Val-Val-Lys-Thr-Leu-Ile-Ser-Ser-Ile-Pro-Leu-Gln-Gly-Ala-Phe-Asn-Tyr-Lys-Tyr-Thr-Ala (SEQ. ID No. 46).

This sequence was compared to known sequences in the gene bank using the databases identified above and was unexpectedly found to be identical to gross cystic disease fluid protein, a protein whose expression was previously found to correlate with the presence of metastatic breast cancer (Murphy, L.C. et al., J. Biol. Chem. 262:15236-15241, 1987). To the best of the inventors’ knowledge, this protein has not been previously identified in male tissues.

The ability of Fraction 1 as described above, to bind to steroid was investigated as follows. Purified rat steroid binding protein (RSBP) and fraction 1 were subjected to SDS-PAGE and transferred onto nitrocellulose filters. Specifically, 1.5 μg of RSBP/gel lane and 4 μg of fraction 1/gel lane were electrophoresed in parallel on a 4-20% gradient Laemmli gel (BioRad), then electrophoretically transferred to nitrocellulose. After protein transfer, the nitrocellulose was blocked for 1 hour at room temperature in 1% Tween 20 in PBS, rinsed three times for 10 min each in 10 ml
0.1% Tween 20 in PBS plus 0.5 M NaCl, then probed with either 1) 0.87 μM progesterone conjugated to horseradish peroxidase (HRP, Sigma) diluted in the rinse buffer; 2) 0.87 μM progesterone HRP with 200 μM estramustine; or 3) 0.87 μM progesterone HRP plus 400 μM unlabelled progesterone and 200 μM estramustine. Each reaction mixture was incubated for 1 hour at room temperature and washed three times for 10 min each with 0.1% Tween 20, PBS, and 0.5 M NaCl. The blots were then developed (ECL system, Amersham) to reveal progesterone HRP binding proteins that are also capable of binding estramustine.

With both rat steroid binding protein and Fraction 1, three bands were obtained that bound HRP-progesterone and that were competed out with unlabelled progesterone and estramustine (Fig. 3). These results indicate that the three bands isolated from human seminal fluid as described above bind hormone and correspond in number of polypeptides to the chains C1, C2 and C3 of rat steroid binding protein, although slightly bigger in size, either due to primary sequence or secondary post-translational modifications.

This putative homologue of rat steroid binding protein was also identified in a subsequent screen of human seminal fluid using the rabbit antisera detailed above. Specifically a hydrophobic 22kD/65kD aggregate protein was obtained which, following CNBr digestion of the 22kD band, provided a peptide having the following sequence:

Val-Val-Lys-Thr-Tyr-Leu-Ile-Ser-Ser-Ile-Pro-Leu-Gln-Ala-Phe-Asn-Tyr-Lys-Tyr-Thr-Ala (SEQ. ID No. 47).

This peptide was found to correspond to residues 67 through 87 of gross cystic disease fluid protein and was identified again utilizing human autoimmune prostatitis sera as discussed below in Example 4.
Example 3
Isolation and Characterization of Polypeptides Isolated from LnCaP.fgc
Using Rat Prostatitis Sera

A LnCaP.fgc cell pellet was homogenized (10 gm cell pellet in 10 ml) by resuspension in PBS, 1% NP-40 and 60 µg/ml phenylmethylsulfonyl fluoride (PMSF) (Sigma, St. Louis, MO) then 10 strokes in a Dounce homogenizer. This was followed by a 30 second probe sonication and another 10 strokes in the Dounce homogenizer. The resulting slurry was centrifuged at 10,000 x G, and the supernatant filtered with a 0.45 µM filter (Amicon, Beverly, MA) then applied to a BioRad (Hercules, CA) Macro-Prep Q-20 anion exchange resin. Proteins were eluted with a 70 minute 0 to 0.8 M NaCl gradient in 20 mM tris pH 7.5 at a flow rate of 8 ml/min. Fractions were cooled, concentrated with 10 kD MWCO centrifprep concentrators (Amicon) and stored at -20°C in the presence of 60 µg/ml PMSF. The ion exchange pools were then examined by electrophoresis on 4-20% tris glycine Ready-Gels (BioRad) and subsequent transfer to nitrocellulose filters. Ion exchange pools of interest were identified by ECL (Amersham International) Western analysis, using the rat sera described above in Example 3A. This analysis indicated an approximately 65 kD protein eluting at 0.08 to 0.13 M NaCl. The rat sera reactive ion exchange pool was subjected to HPLC and subsequent Western analysis to identify the protein fraction of interest. This protein was then digested for 24 hours at 25°C in 70% formic acid saturated with CNBr to cleave at methionine residues.

The resulting CNBr fragments were purified by microbore HPLC using a Vydac C18 column (Hesperia, CA), column size 1x150 mM in a Perkin Elmer/Applied Biosystems Inc. (Foster City, CA) Division Model 172 HPLC. Fractions were eluted from the column with a gradient of 0 to 60% of acetonitrile at a flow rate of 40 µl per minute. The eluent was monitored at 214 nm. The resulting fractions were loaded directly onto a Perkin Elmer/Applied Biosystems Inc. Procise 494 protein sequencer and sequenced using standard Edman chemistry from the amino terminal end. Two different peptides having the following sequences were obtained:
(a) Xaa-Ala-Lys-Lys-Phe-Leu-Asp-Ala-Glu-His-Lys-Leu-Asn-Phe-
Ala (SEQ. ID No. 48); and

(b) Xaa-Xaa-Xaa-Lys-Ile-Lys-Phe-Ile-Gln-Glu-Asn-Ile-Phe-
Gly,

wherein Xaa may be any amino acid (SEQ ID No. 49).

These sequences were compared to known sequences in the gene bank
using databases identified above, and identified as residues 286 through 300 and 228
through 242, respectively, of probable protein disulfide isomerase ER-60 precursor,
hereinafter referred to as ER-60 (Bado, R. J. et al., Endocrinology 123:1264-1273,
1988). This antigen is also known as phospholipase C-alpha (see PCT WO 95/08624).
Residues 285 and 227 of ER-60 are methionines, consistent with the above sequences
being cyanogen bromide fractions.

ER-60 is a resident endoplasmic protein with multiple biological
activities, including disulfide isomerase and restricted cysteine protease activity. In
particular, ER-60 has been shown to preferentially degrade calnexin, a protein involved
in presentation of antigens via the Class I major histocompatibility complex, or MHC,
pathway. ER-60 and a related family member, ER-72, have been shown to be over-
expressed in colon cancer, with truncated forms of ER-60 exhibiting increased
enzymatic activity (Egea, G. et al., J. Cell. Sci. (England) 105:819-30, 1993). However,
to the best of the inventors’ knowledge, this polypeptide has not been previously shown
to be present or overexpressed in human prostate. Recently, ER-60 gene expression has
been correlated with induction of contact inhibition of cell proliferation (Greene, J.J.
et al., Cell. Mol. Biol. 41:473-80, 1995). Thus, if ER-60 is also truncated and non-
functional in prostate cancer, as it is in colon cancer, the resultant loss of contact
inhibition would lead to neoplastic transformation and tumor progression.
Example 4
Isolation and Characterization of Polypeptides Isolated from LnCaP.fgc
Using Human Prostatitis Sera

The human prostatitis sera described above in Example 1 was used to screen the LnCaP.fgc cell line using the ion exchange techniques described above in Example 3. Reactive ion exchange pools were purified by reverse phase HPLC as described previously and the polypeptides shown in SEQ ID Nos. 50-51 were isolated utilizing cross-reactivity with said antisera as the selection criteria. Comparison of these sequences with known sequences in the gene bank using the databases described above revealed the homologies shown in Table II. However, none of these polypeptides have been previously associated with human prostate.

<table>
<thead>
<tr>
<th>SEQ ID No.</th>
<th>Database Search Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>53</td>
<td>glyceraldehyde-3-phosphate-dehydrogenase</td>
</tr>
<tr>
<td>54</td>
<td>alpha-human fructose biphosphate aldolase</td>
</tr>
<tr>
<td>55</td>
<td>calreticulin</td>
</tr>
<tr>
<td>56</td>
<td>calreticulin</td>
</tr>
<tr>
<td>57</td>
<td>malate dehydrogenase</td>
</tr>
<tr>
<td>58</td>
<td>cystic disease fluid protein</td>
</tr>
<tr>
<td>59</td>
<td>cystic disease fluid protein</td>
</tr>
</tbody>
</table>
Example 5
Isolation and Characterization of Polypeptides from Human Seminal Fluid

Polypeptides from human seminal fluid were purified to homogeneity by anion exchange chromatography. Specifically, seminal fluid samples were diluted 1 to 10 with 0.1 mM Bis-Tris propane buffer pH 7 prior to loading on the column. The polypeptides were fractionated into pools utilizing gel profusion chromatography on a Poros (Perceptive Biosystems) 146 II Q/M anion exchange column 4.6 mm x 100 mm equilibrated in 0.01 mM Bis-Tris propane buffer pH 7.5. Proteins were eluted with a linear 0-0.5 M NaCl gradient in the above buffer. The column eluent was monitored at a wavelength of 220 nm. Individual fractions were further purified by reverse phase HPLC on a Vydac (Hesperia, CA) C18 column.

The resulting fractions were sequenced as described above in Example 3. A peptide having the following N-terminal sequence was obtained:

(c) Met-Asp-Ile-Pro-Gln-Thr-Lys-Gln-Asp-Leu-Glu-Leu-Pro-Lys-Leu

(SEQ ID NO:57).

Comparison of this sequence with those of known sequences in the gene bank as described above revealed 100% identity with human placental protein 14 (PP14).

Example 6
Synthesis of Polypeptides

Polypeptides may be synthesized on an Applied Biosystems 430A 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, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture:

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 or other types of mass spectrometry and by amino acid analysis.

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

(1) GENERAL INFORMATION:

   (i) APPLICANT: Corixa Corporation

   (ii) TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
        AND IMMUNODIAGNOSIS OF PROSTATE CANCER

   (iii) NUMBER OF SEQUENCES: 57

   (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: 14-MAR-1997
        (C) CLASSIFICATION:

   (viii) ATTORNEY/AGENT INFORMATION:
          (A) NAME: Maki, David J.
          (B) REGISTRATION NUMBER: 31,392
          (C) REFERENCE/DOCKET NUMBER: 210121.424PC

   (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: 89 amino acids
       (B) TYPE: amino acid
       (C) STRANDEDNESS:
       (D) TOPOLOGY: linear

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

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

(2) INFORMATION FOR SEQ ID NO:2:  

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 89 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:  
Ala Arg Gly Arg Phe Gly Arg Leu Gly Val Gly Gly Glu Pro His Pro  
1 5 10 15  
Arg Arg Asn Pro Ala Leu Pro Thr Glu Leu Ala Glu Leu Thr Pro Gln  
20 25 30  
Val Arg Arg Ala Ala Xaa Lys Thr Gln Arg Ser Gln Val Lys Pro Arg  
35 40 45  
His Arg Arg Gly Trp Pro Pro Thr Val Pro Leu Ala Gly Arg Leu Glu  
50 55 60  
Glu Leu Lys Thr Pro Arg Ser Pro Arg Pro Pro Gln Gly Leu Asp  
65 70 75 80  
Pro Ser Pro Cys Ser Leu Pro Ser Pro  
85  

(2) INFORMATION FOR SEQ ID NO:3:  

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 858 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS:  
(D) TOPOLOGY: linear  

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:  
Gln Glu Ser Glu Pro Phe Ser His Ile Asp Pro Glu Glu Ser Glu Glu  
1 5 10 15  
Thr Arg Leu Leu Asn Ile Leu Gly Leu Ile Phe Lys Gly Pro Ala Ala  
20 25 30
Lys Glu Ser Pro Ala Met Leu Pro Thr Phe Leu Leu Met Asn Leu Leu
370 375 380
Ser Leu Ala Gly Asp Val Ala Leu Gln Gln Leu Val His Leu Glu Gln
385 390 395 400
Ala Val Ser Gly Leu Cys Arg Arg Arg Val Leu Arg Glu Glu Gln
405 410 415
Glu His Lys Thr Lys Asp Pro Lys Glu Lys Asn Thr Ser Ser Glu Thr
420 425 430
Thr Met Glu Glu Glu Leu Gly Leu Val Gly Ala Thr Ala Asp Asp Thr
435 440 445
Glu Ala Glu Leu Ile Arg Gly Ile Cys Glu Met Glu Leu Leu Asp Gly
450 455 460
Lys Glu Thr Leu Ala Ala Phe Val Pro Leu Leu Leu Lys Val Cys Asn
465 470 475 480
Asn Pro Gly Leu Tyr Ser Asn Pro Asp Leu Ser Ala Ala Ala Ser Leu
485 490 495
Ala Leu Gly Lys Phe Cys Met Ile Ser Ala Thr Phe Cys Asp Ser Gln
500 505 510
Leu Arg Leu Leu Phe Thr Met Leu Glu Lys Ser Pro Leu Pro Ile Val
515 520 525
Arg Ser Asn Leu Met Val Ala Thr Gly Asp Leu Ala Ile Arg Phe Pro
530 535 540
Asn Leu Val Asp Pro Trp Thr Pro His Leu Tyr Ala Arg Leu Arg Asp
545 550 555 560
Pro Ala Gln Gln Val Arg Lys Thr Ala Gly Leu Val Met Thr His Leu
565 570 575
Ile Leu Lys Asp Met Val Lys Val Lys Gly Gln Val Ser Glu Met Ala
580 585 590
Val Leu Leu Ile Asp Pro Glu Pro Glu Ile Ala Ala Leu Ala Lys Asn
595 600 605
Phe Phe Asn Glu Leu Ser His Lys Gly Asn Ala Ile Tyr Asn Leu Leu
610 615 620
Pro Asp Ile Ile Ser Arg Leu Ser Asp Pro Glu Leu Gly Val Glu Glu
625 630 635 640
Glu Pro Phe His Thr Ile Met Lys Gln Leu Leu Ser Tyr Ile Thr Lys
645 650 655
Asp Lys Glu Thr Glu Ser Leu Val Glu Lys Leu Cys Gln Arg Phe Arg
660 665 670
Thr Ser Arg Thr Glu Arg Gln Gln Arg Asp Leu Ala Tyr Cys Val Ser
675 680 685
Gln Leu Pro Leu Thr Glu Arg Gly Leu Arg Lys Met Leu Asp Asn Phe
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 127 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS: linear
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

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

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

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

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 751 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

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

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

<table>
<thead>
<tr>
<th></th>
<th>420</th>
<th>425</th>
<th>430</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys</td>
<td>Phe</td>
<td>Phe</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>Asp</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Pro</td>
<td>Arg</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>435</td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Ser</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Asn</td>
<td>Ser</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Glu</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Arg</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>450</td>
</tr>
<tr>
<td>Ile</td>
<td>Asp</td>
<td>Ser</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Ala</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Leu</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Leu</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Gly</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>465</td>
</tr>
<tr>
<td>Glu</td>
<td>Lys</td>
<td>Leu</td>
<td>Tyr</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Leu</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Glu</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Lys</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Phe</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>485</td>
</tr>
<tr>
<td>Leu</td>
<td>Pro</td>
<td>Asp</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Ser</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Arg</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Asn</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Glu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>500</td>
</tr>
<tr>
<td>Lys</td>
<td>Val</td>
<td>Glu</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>Ala</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Thr</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Arg</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Asp</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>515</td>
</tr>
<tr>
<td>Ser</td>
<td>Val</td>
<td>Pro</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>Ile</td>
<td>Ser</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>Asp</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Phe</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Val</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>530</td>
</tr>
<tr>
<td>Gly</td>
<td>Thr</td>
<td>Asp</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Arg</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Leu</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Asn</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>545</td>
</tr>
<tr>
<td>Leu</td>
<td>Glu</td>
<td>Ser</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Glu</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Ala</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Gly</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Lys</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>565</td>
</tr>
<tr>
<td>Gln</td>
<td>Thr</td>
<td>Asn</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Gly</td>
<td>Ile</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Asn</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Leu</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Pro</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>580</td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td>Cys</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Arg</td>
<td>Leu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Glu</td>
<td>Gln</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>Gln</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Gln</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>595</td>
</tr>
<tr>
<td>Lys</td>
<td>Gly</td>
<td>Ile</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ser</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Ser</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Lys</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Tyr</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>610</td>
</tr>
<tr>
<td>Ile</td>
<td>Thr</td>
<td>Glu</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Tyr</td>
<td>Ala</td>
</tr>
<tr>
<td></td>
<td>Gln</td>
<td>Glu</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Pro</td>
<td>Arg</td>
<td>Met</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Phe</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>625</td>
</tr>
<tr>
<td>Ala</td>
<td>Xaa</td>
<td>Val</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Ser</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Gly</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Met</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Pro</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>645</td>
</tr>
<tr>
<td>Ser</td>
<td>Asp</td>
<td>Pro</td>
<td>Lys</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Gln</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Gly</td>
<td>Val</td>
</tr>
<tr>
<td></td>
<td>Val</td>
<td>Glu</td>
<td>Ser</td>
</tr>
<tr>
<td></td>
<td>Xaa</td>
<td>Val</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>660</td>
</tr>
<tr>
<td>Ala</td>
<td>Asn</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Ser</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Arg</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Cys</td>
<td>Ser</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td>Leu</td>
<td>Ser</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>675</td>
</tr>
<tr>
<td>Gly</td>
<td>Ala</td>
<td>Asn</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Tyr</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Ser</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Asp</td>
<td>Thr</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>Thr</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>690</td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td>Leu</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Asp</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Thr</td>
<td>Asp</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td>Ala</td>
<td>Lys</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Glu</td>
<td>Glu</td>
<td>Pro</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>705</td>
</tr>
<tr>
<td>Thr</td>
<td>Ser</td>
<td>Leu</td>
<td>Phe</td>
</tr>
<tr>
<td></td>
<td>Xaa</td>
<td>Arg</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Glu</td>
<td>Thr</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>Gly</td>
<td>Glu</td>
</tr>
<tr>
<td></td>
<td>Ser</td>
<td>Gly</td>
<td>Leu</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>725</td>
</tr>
<tr>
<td>Val</td>
<td>Leu</td>
<td>Glu</td>
<td>Asp</td>
</tr>
<tr>
<td></td>
<td>Lys</td>
<td>Glu</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>His</td>
<td>Gln</td>
<td>Ile</td>
</tr>
<tr>
<td></td>
<td>Phe</td>
<td>Glu</td>
<td>Gly</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>740</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>745</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>750</td>
</tr>
</tbody>
</table>
(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 6 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

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

Ala Arg Gly Ser Thr Gln
1  5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 16 amino acids
   (B) TYPE: amino acid
   (C) STRANDEDNESS:
   (D) TOPOLOGY: linear

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

Ala Arg Gly Ser Ser Gln Val Arg Val Lys Ser Trp Arg Gly Asp Met
1  5  10  15

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 271 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

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

CCGCACGAGC CTCTGCATG CTTCTGGCA TGATGGCAGG AGGAAAGCCA GAAATGTGG 60
GAAGCAATTT AGACACACTG ATGAGCATAG GGCTGGATGA GAAGTTTCCA CAGGACTACA 120
GGCTGCCCCCA GCAGGTGTGC CATGCCATTG CCAAACATTCG GGACAGGAGA AAGCCTTCTC 180
TGGGCAAACG TCACCCCCC TGCCGGCTGC CTCAGGAACA CGGTTGTCTT GAGCGACTGC 240
GGGAGACAGT CACAAAAGGC TTTGTCACCA C 271

(2) INFORMATION FOR SEQ ID NO:10:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 403 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

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

GGGTGGAATAA CCTGAGGTAG GGAGTTCGAG ACCAGCCTGA CCAACATGGA GAAACCCCAT  
       60
CTCTACTAAA AATAAAAAAT TAGCCGGCGT ATGGGGTTCG GCTCTGAAATC CCAAGCTACTC  
       120
AAGAGGCTGAA GGCAGGAAAG TGGCGTCAAC ACCAGGGCGG AGGGTGTAGT GAGCCGGAAT  
       180
CACACCATTG CACTCCAGCT TGGGAACCAA TATGCGAATCT CACATCTCAAA TTAAAAAAA  
       240
TAAGGCCTACA CCCTTCCTTA AAATGCAGGG CTTCTCTCTA AATTAGCCTTA ACCTGAACCTG  
       300
GTGAGGCTGC TTCAACTTTC GAATATAATG TTGCAATCTT CCGTGTCTTAC TAATGAATTA  
       360
ATGTTTTTAT ATACTTTTAA AAAAAAAA AAAAAACTC GAG  
       403

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 2276 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

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

GGAGGTTTGG GGCAGCTGAG GCTGGAGAGG AGGCCCAACCC GCGGAGGAAC CGCCGCTTTCG  
       60
CAACGGAGCT GCGGAGGCTC ACTCCTGCAG TGACGCGCGG GGGGTGAAAA ACGGAGCGGA  
       120
GCCAGGTGAA ACCAGGCCAC GCCGCTGGCT GCCGCACGAA AGTTCTCTCTA GCCGGGAGGT  
       180
TGAGGAGCT GAAAACCGCG CGGAGCCCTC GCCGCACGGA GCCGGGCTTG GACCCCAAGCC  
       240
CTTGCAGGCT CCGCTTCCTT GGCACCCAAAG TGCGATCTCG GCTGCAAGAG GGGCCCGGGG  
       300
CGACACTGAT TTCCACCTC GTGTCAGGCT GTCGTGTGCTA GGCTCGACGCC TTAATGAGAG  
       360
GTGATTCCCTA AGCTGTCGGG AACGTGAGGT TGCCAAAGGG GGGGCAGGAAG ATGGGAGCCA  
       420
GTATACCAAG CCAGAGCACA ACTTGGAGGT TAAACCCTCA GCCCATTCAG CAAATGTTTT  
       480
CATCCATTTT GGACCAAAA GATGGGATTC GATTTCATT TTAAAAAGAT AATGTATAGAT  
       540
ATCTGATCTC ACTACAAATA TTTAGCTGAG AAGATTATCT AGCTTGTCCT TTTGTTGGAC  
       600
TGTCACCTCT TTCTGAAGGT TTCTTGACCC ACAAATTTCT TAATGAAATT CAAGTATCAA  
       660
AGGATACCAG GTTTAGAATG GTATAATGAT GTATTTTGTG TGAGGACTGC AAATTTTTATA 720
GAGACCACAG TTGATTTCGA GTGATATTCT GCAAATCAAA TGATTTGATA AACCTAATTT 780
TGAGGCTATT TATATTATTA AGCGACATCA AAAAGATGGGA GAAAAAAATG GCGATGCAA 840
AACTTTCTGG ATGGAGCTAG AAGATGATGG AAAAGTGCGG TTCTATTATTG AACNAGTACA 900
AAATGTGCTG CAGTCACCTGA AAACAAGCAG CAAAGATGGG TCTGGACACCA ATACAGAATA 960
CATCCAAGCA ATGATTCCTAG TGAATGAAGC AACTTAAATT AACAAGTCAA CATCAATAAA 1020
GGATCTCTAG CCTGGAACTC AGAAGGAACA GGAACAAAAA TCAAATGCAT TCCCTCTCAG 1080
ATCATGTGAA AACCTCTTTTC CAAAGAAGCT TACATTTCCTA ACAAGAGGAA AATAGGAAAT 1140
CTCTCTCTCT GAAAGAAAGG TTGTATGGCT TAGAGAAAAA GACTCATCTT CGAATTATATC 1200
TTACAAAGCT CATGACCTGCT CTGGATGGCTG TCTGATGAAG ATGCCACTGA AATGGAAGGG 1260
AGAAAACCCCT CTGCAAGCTGC CAACTCATAA GTCACTTCAAA AGACGACATG CAAAGACAAA 1320
CTCTCATTTG CAGAAGCTCC AACGAGTTTA AAAAAACCTT TGTGAGAGGA GTCTAGGAAA 1380
CGTGGAGGAA GTTTTTCTTT ACCGGGCTGA GACAGAGTGT AACCTTTTAT TTAAGATAAA 1440
CTTTTCTTTT AACTACATAG TTTGAGCTGG CAGGAATTAC CCAAGGCCAA AAGAGGTGGT 1500
TTCTGATGTG GATATAGCAG ATGGAGTGGA ATCCAGTCCC AATTCTCTTT CTATGAAAT 1560
TGACAGTGAG AAGCTCCCCA AATTTAACGA CAGGAAGACT GTGGGACCTG GAGCATATAA 1620
TCTAACCACAC TTTGCCAGCA TGTCTACATGA TTTCTGTGAC TGCTCTGGAG GCTGCCAATA 1680
CATAACAAAA TGTGCAATGTC TCTCAACTGAC AGCAAGAAT GCCAAAACCTT CCCCCCTGGTC 1740
AAGTGACAAA ATAAAACACTG CATATAATAA TAAAGACTA CAGAGCAGA TTTCTACTGG 1800
CATTATGAAA GACAGGCTCT TTGGCAATAG TAATCGCACA TTTGCTCAAAG ACCGAGTGGT 1860
CCACATGATT CCTCAAGTGA GGTTACAGGT GTTCAAAAAT GAGCCAGAGG GATGGGGTGT 1920
AGCGTCTGTA GATGACATTG ACAGAGGCCG ATTTGTTTGC ATTTATTCCG AGAGATTACT 1980
AAGCGAGCTT AACACTTGGAA AACTCTATTG TATTGTAAGA AAGGGGAGAG ATGAGAATAC 2040
TATGAAATAT ATATTTCGAA AAAAGAGGAA ATAGAAGTTT GCATGTTCAG ATGTTGAAGT 2100
TGAAGTCTCT CCATTAGGAT TGGAAACAAC TCTCTGAACCT GCTAAGACTG AGAAATGCTC 2160
ACCACAGTCC AGTAAATAATC CCAAGGAGCT TACTAGGAGA ACGAAAATAT GATATAATTT 2220
AAGAATTACG TACATTTCAAT ATTAGAGGAT TCCTGAATTC AAGACAGCCA TTTTTC 2276

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3114 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CAGGAGTCGG AACCCTTCAG TCATATAGAC CCAGAGGAGT CAGAGGAGAC CAGGCTCTTG
AATATCTTAG GACTTATCTT CAAAGGCCCA GCAGCTTCCA CACAAAGAAAA GAATCCCCCGG
GAGTCTCAG GGAAACATGGT CACAGGACAG ACTGCTGTGA AAAATAAACCC CAATATGTCG
GATCTCAGGG AATCCAGGGG AAATGATGAA CTAGTGAAAG AGGAGATGCT GGTACAGTAT
CTGAGGATG CTCTAGACCTT CTCCCCGAAG ATTCAAGGAG CCATTGGCAT CACTAGCAAG
ATGATGTATG AAAAAACACAC TCACATGGTG CAGGAGGTTG TTAGATNCTT TGTTGATGTGC
TTCCATTGTG GGCTACCCCA GCCCCTGTTTT GGCTGTGCCC GTATGCTGCC TCTCATCTGG
TTGTTAAGGAGC CTGCGTGTCGG GGAACCCGTCG CTATAATCCCT ACCCAGCACTC TCACTCTCCA
CCCAAAGGAGG ACTCTGGCCAG AGCCCAAAGGCC CAGGCTTGGAA TTCAGAAATCT CTCTCTGTGG
CTAGGTTGGAT GCCTGCTTGG GACCTATCTG TCTCTTGAGG AATTTCTCTG TGAATTGTTGG
CAGAGGAGGC AGCTGAAACCC ACTCTCGGCT CGAGGGCGGC ACCGAGGAGG AGGCAGAAGG
GTCGGCTGCT GCCTCTGAGA GGCGCTGTCC TCTGCTGATC TGCTTGGCAT GATGGCCAGA
AGAAGCCAGC AAATGTGGAG GACCAAATTC GCACACTCATC TGAGGATAGG GCTGGATGAG
AAATCTCCCAC AGGACATACAG CTGGGCCCCGC CAGGCTGTCG ATGCCATTTGC CAAACATCTG
GACAGGGAGAA GCCTTCTCTCT GGGCAAAAGGT CACCCCCCCT TCCGCTGCAG TCAGGACAG
AGGGTGGTTGG AGGGACTCGGC GAGGACAGTC AGAAAGGCCT TTGTCACCCCC AGACCCACTC
TGATGATGATG AACAAAGGGGC GAGCAGTGGC CACCTTACCAG AACTGAGAAGA GGGCCCGAAA
GTGATGATGATG CCAGATATTT GAGAGGCTGTG GCAAACAAGG CCCCCTGAGAA GCTAGAAGAG
AGAGAAACCA CTGGGAGAGA CCGGAGGAGC TCCCCGCGAA TGCTCTTCGAG TTCTCTCTCC
AGGATGCTGG TGGGAGCTTG GCTCTGCTGCA AGCTGTGGCCA CTTGGAGGAGG
GCCGACTGCTG GCAGGGCCGC GTGCTCGCGG AAGAACAGGA GCACCAAGACC
AAAGACCAAC AGGGAGAGAAG TACGAGCTCT GAGACCAAAA TGGAGGAGGA GCTTGAGGCTG
GTGGGGGCAA CAGGAGATGA CACAGAGGCA GAACATATCC GTGCCATCTG CGAGATGAAA
CTGCTGGAAG GCACACTGAG ACTGGCTGCC TTGTGCTCAG TCTGCTTAA GCATCTCTAC
AACCCGGGCC TCTATAGCAA CCCCAGACCTC TCTGAGCTCT GCTCTCAGTG CCGTGCCAGA
TCTCGATGAG CTGATGGCCAC TTTGCTGAGCC TCCAGCCTCC GTCTTGGCTT CACCATGGTC
GAAAGATGCTC CACTCCACCT TGTCCTGGCT TTGCTCATGGG GAATCTGGCC
AACCCCGGGCT CCAATCTGGT GGAACCCGTGG ACTCCCCATC TGATGCTCCG CCCCCGAGC
CCTGCTCAGC AAGTGCAGAA AACCACGGGG CTGGTGATGA CCCAACCTGAT CCTCAAGGAC 1740
ATGGTGAGGG TGAAGGGGCA GTGCTGAGTG ATGGGCAATGGC TGGCTCATAGA CCCCGACGCTT 1800
CAGATCTGCT CCCTGCGCAAA GAACTCCTCT AATGACCTCT CCCACAAAGG CAAGCCAATCC 1860
TATATACTCC TTTCAAGATAT CATCGAGCCC CTGTCAAGCA CCGAGCTGGG GGGTGGAGGA 1920
GAGCCCTTTCC ACACACCTCAT GAAACAGCTC CTGTCCTACA TACCAAGGA CAAGAGAGACA 1980
GAGAGCCTGG TGGAAGAAGCT GTGTCAGCGG TTCCGACATC CCGAAGCTGA GGGGCAGCAAG 2040
CGAGACCTGG CCTACTGTTT GTGCACAGCTG CCCCTCAGAG AGCGAGCCCT CGGTAAGAGTG 2100
CTTGACAATT TTGACTGTTT TGAGACAAAA CTGTCAAGATG ACGCCATCTT CAGTCCTTTT 2160
TTGCAGCTTG TGGGCAAGCT GCCAGCTGGG CCAAGCCTGG AGGGCAAGGC TATAATAGAT 2220
GAAATTGAGC AGAAGCTCTTG GGCCCTGCAT ACCAGAGTT TGGATGGAAT CAAGAGCTTT 2280
GAGATAGGCC AAGCCAGAGTC CCAGAGAGCG CCCATCAAGCG AGAAAACCACG CACTGATTTCT 2340
AGGTACCAGC CTCTGGAATC TACAGCAATCT GTGAGCAACC AGAGCCCGCC 2400
CGTACTCCCT GTCGCAATCC AAACACACCA CAGCGAGCTT CCAAAAGAAG ACCAAAGATTG 2460
GTTCATTCAA GTGATGATGC CAGTGAGGAA GACATTGACG CAGAGATGAC AGAAGAGCAG 2520
ACACCAAGCA AAAAACCACTC CATTCACTAG GCACTCGGCA GCAGCCGACAG ATTCCTAGGAA 2580
GTGCTGTGCT GTGCTCTCTG TGGAGGGATGC CCGCTAGGGT GACCTGGAAT TCGAATTCTG 2640
TTCCCTCTGT AAAATTTTGA TCTGTCTCTT TTGTTTAATA AAAAAAAGGC CCGGAGACTG 2700
TGCTCGACGC CTGTAAACCC AGCAATTTGC GATACCAAGG CGGATGGATA ACCTGAGGTA 2760
GGGACTGCAG GACCAGCGTG ACCAATCGAA AGAACCCCCA TCTTCTACAA AAATAAAAAA 2820
TTAGGCCGCG GTATTGCGCT GCACCGTCAA TCCAGCCTAC TCAAGAGCCT GAGGGAGGAG 2880
AATCGCGCTGA ACCAGAGGCG GGAGTGGTGA GTGAGGGCAA ATCGACACAT TCGACTCCAG 2940
CTTGGGGCAAC AATAGCGAAC CCTGATCTCA AAATAAAAAA AAATGCGATA CAGCGCTTCTT 3000
AAAATGCAAG GCTTTCTCTT AAATAGCCTT AACTGACTCG CTTGAGCTG TGCTCAGTTTT 3060
GGAATATATG TTTGCCAATC TCCTTGTTTT CTAATGGAATA AAAGTTTTTATA TATA 3114

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1797 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:
CGGCAGGAG TCGACTGGTG GCAAGTAAAA CAGATGGAAA AATTAGTACAG TATGAATGTG 60
AGGGGATAC TTGCGAGGA GAGAAAATAG ATGCTCTACA GTGAGGATAT TCTATATTAC 120
TAACAGCCA GCTGGAAATCT CAGCGGAATCT ACTGGGAAAA CAAGATAGTT CGGATAGGA 180
AGGACACAGC AGAGGAAAAT ACAAACATGA AGAACCAGTT TAAAGRAACA ATGGAGAAGT 240
GTGATAATCT AGAGCCAAAA CTAATAATGCT TCTAAAGAAA AAGGCGGCTG GCTGAAGAAG 300
AGTGCACCTCA GTAAACACAAG AAAGTTGCGCA AACTCCACAA CGAGTCCAAA GAGGACAGG 360
AAATGAACAA TGGTTTGCGA GCCCAACCAAG TCCCTCTGCCT GAACAGGCTA AAAGAGGAGG 420
AGAGGGTGCCT GAAGAGAGACC TTGGAACCAA AAGATCTGCA GATCACCAGG ATCCAGGAGC 480
AGCTGCGTGA GTGATGTCTC TACCTGGAGA CACACGAGAA GATGCAACAT CTGGCCTGCC 540
AGACCGGAGA GGAATCCCGA GAGGGACAGA TCAATACCGC CAGCCCTCGT GCCCTGAGCC 600
CTGGCCCTTC GGGGGCGAGT GGGAAGTCGG CCTCCAGGAA GGGCGCAGAG AAGGAGGCA 660
AGTGACCTTC AGAGCAACAG ACACTCCCTGA GACTGGTCTC CCTGACACTG TGAGAGTGTG 720
CTGGAGACCTT CATGAAATTG TGAGGTTGGG CCGTAAAAG TACAAGTGAG GATCAAGCCA 780
CAAGTGTGTTG GCGTTTTCAT TTGGTATGGT GTGATGTANT GAATGTAAGG GTTGCTGACT 840
GGAGAGCTGA TAGAACGCGT CCAGCTTCGCA AAGAGTTTCA AANAGTCTAC AATCCTCACAT 900
TCTTAATGACC ATTTGTGCTT CCGTTCTGCTG AAAGGCCCAGA AACTCTGCTTG CCACTTTC 960
ATTGATTTTA TGGGATGGGC GTATTTGACA TTCGTTTCCTG GGTTAGGTTT AAGATGTATAA 1020
GTTATTTTCTT GGAATCTCAA AGGCTAAGGTT ATCTAGACCT AAAGCACCAC AACTCTCTGA 1080
GGGCATAACA GCTGCCTTAAA AGAGAGGTTT CCAATGGGCTA TTAAGGAGTT ATGAAAATCT 1140
CCTAGCAATAT GGTATCATATC ATTATCATCT CCCCCCTCCT CTGGGAGAGT GAAGAATGTC 1200
TTGAATGTATA TCTGAAAAAG GGCCTGCTAG TAAACCAGCC CCTGGCTTCTT TACCGAGCT 1260
CATCTCTTCT TTGCTCTGGGG CCGAGCAGGA AAAAACAACA ACCCGGGGCC CATTTGGTAG 1320
ACTCAGTGTA GGGAAAAATGG TGCCAGCTCC ACTGTTTAATT TTTGTGACT TCGTAGGCTA 1380
TTATGAAACCG CATATTAAAGA GGAGGCTTAAA TGCGCTGTCC CAAAATCCAA TCTCAGAGTG 1440
GGTATCCTAG CATCTAGCAA NACCTGAGTG GGAGATTTCT CATCCGGTTG AAAATGTAGA 1500
GTGAGGCCTGC TGCAGAGCTT ATTTGTATTG TCTTGGGGTT TAGATTTTCT TAAATGTCTA 1560
CAAAATATTG GCCCTGCGAT TACAGCTGCA GCTAAGGAGA GCTGGGCCAN ATTTTCAATT 1620
ACGCTTTCAG GAATTAAACCA AAGCCTGTGT TCAAATCCTTA AAATTAGAAT TTCAACAGAN 1680
CCCGCTTTAG ACACTGCTAT TAAACCTGGTG GTGCCCCAAC AGANGGGCCT GCTTATCTTCT 1740
CTGGAAACCAT AAATGTCAAAT TAATTTTATAA CCTGCANTAA TTGAAGCAGCT TAAATAA 1797

(2) INFORMATION FOR SEQ ID NO:14:
(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 720 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:
TAATCACCAT CTGTTTTTGT GGATGTGCT GCAAGATTTTC CCAAAAAACT TNACGGTTAA 60
TGTTGAAAA TGAATGTACT CAGACATTNT TAATTTTTAC TTAGGGCAGA CCAAACCTTT 120
GAGTCCTCCT TGGACTTATA TATACAGATA TCCTAAGAGT GGGAAATGAA AGCATAACCT 180
AATTTNCTTT CCTATAGAGA TTCTATTTTA TTFAAAATNT ATTTNTACAC TAGTTAGAAT 240
CCTGCCTGTT TGGCCAAAGTA CTGTCCTGGC ATGCTCGAGC TTGCAAGAGC TGGGTTGGAT 300
CATACGATAC TAATGAGAGG AATTAGAGTT AGTTTACAAA GACTCGCTAC TCTCTATTT 360
TCTGTCATCC CTCTATCACA GTGCCCCAC CACCACTTGG GAAACAGAT TTTTCAGTAC 420
AGGTTGGATA AATGCCTCAG AAGGCTGGTC CCAGAGGAAAT GACAGAAATAG CCAAAGTGGTT 480
CCAAACTACT TGGAGGGTTA CAAAAAATAT GTCCCAAGAA AAAAAAATA CTTACCAAGA 540
TACGTANAGA AAAAAAATA TTGTAAAA GTACGCAAGA GTCATGTTTG AATTCCACA 600
AATCAGTCAA GAGAGAGAGT GAAAAATCAGCGTCAGAA GAACCACCTTA ATACACCCAT 660
ACTACCTTGA ACAATGAAAT TGAATTTAAA TAGCCAAACAT TTGAAAAAAA AAAAAAAA 720

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 1996 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:
CAGAAGTGCA GCGGTGCGCC CGGCTGTTTG CCGGCGGCGG GCGGGCTGCC GAGATGGAG 60
GTAACCTCAG ATCTTTTGCA AGATGGGTG GCTTCACAG CTACCCCTGG GACGGGAAA 120
TCTAAGCTGG AAACATTGCC CAAAGAAGAC CTCACTCAAGT TTGCCAAGA AAGAGATGATG 180
CTAATAGAGA AAGCATTCTC AAGGCTGACA GAATTGGAGA AAGAAATTGA AGAAACTCAGA 240
TCAAAAATCTC TAACTCAGAG AACTGGTGAT ATTATTAAGG CATTAAGCTGA AGCTCTGGAT 300
GCTCTTCTTC TGGAAAAAGC AGAGACTGAG CAACAGTGGC TTTCCTGGAA AAAGGAAAAT 360
ATAAAAATGA AGCAAGAGGT TGAGGTCTCT GTAACAAAGA TGGGAGATGC ACATAAGGAG 420
TTGGAACAAAT CACATATAAA CTATGTGAAA GAAATTGAAA ATTTGAAAAA TGAGTTGATG 480
GCAGTACGT TCAAACATAG TGAAGACAA TCTAATCCTAC AAAAGCGCTT GGAAGCAAA 540
TGATATCCCA ATTAGAACTT TCAAGACACG TTAATTTCGA GAACACTCTC GAAAGATAAG 600
TTAAAAACTT ACAAGAAGAG ATGAGAAAA TTAGGCCAGG CTTTGAGGAG CAATAATTAT 660
ATCTGCAAACA GCAAATAGAG GCCTACCAGTG ATGAAAAGAA GGAACAGTT ACTCAACTCC 720
AAAAATATCAT TGAGGCTAAT TCTACGACATT ACCAAAAATA TATTAATAGT TGCGAGGAAG 780
AGCTTTTACA GTTGGAAGCT ATACACCAAG AAGAGGTGAA AGAGTTGATG TGCCAGATTTG 840
AAGCATCACG TAAAGACATG GAAGAGAGAA TAAATAGGT GAAACGCTA AAGAGAATCT 900
TAAATAAACA ATGTGAGGCA AGTGAACAGA AACATCCAGA GAAATAGGAA TGTTAGTTTAG 960
AAAATTTAAG GAAAGCCACC TCAAATGCAA ACCAAAGACA TCAGATATGT TCTATTTCTC 1020
TGCAAGAAAA TACATTTGTA GAAACAGTG AATAGAAAAA AGTCAAACAC TTGAGATA 1080
CCTTAAAAGA ACTTGGAATCT CAACACAGTA TCTAAAAGA TGAGGTAACT TATATGAATA 1140
ATCTTTACTT AAAACTCTAA ATGGAGTGTCTT AACATATAAAA GGATGAGTATT TTTCTGAAC 1200
GGGAGACTT AGAGTTAAA GATTAATGAA TTTACTAGCG TAAAGAAAGAA CAGGGCTGTG 1260
TAGATTTAAA ATTAATATCT GAGCTAGCG GTTTAAATGA ACAGTTTGGC TATACCTAG 1320
AACAGCATAA CAGAGAGAGT CAGAGTCCTTA AGGAACACAC TCAAAAAGAAA ATATCAGAAC 1380
TAAATGAGAC ATTTTTGTCA GAATCAGAAA AAGAAAATT AACATTAATG TTGGAAATAC 1440
AGGTTCTCAA GGAACAGCTGT GAAAACCTAC AGGAAGAAAA GCAAAGAGCA ATTTTAATT 1500
ATGAGCATTT ATCCAGAGATT ATGGGAAATTT TACCAACAGA ACTGGGGGAA TCTGCGGAA 1560
AAATAACTCA AGAGTGCCCA TCAATGAGAC ACCAAGAACG ATCTGATGTT CATGAACTGC 1620
AGCAGAAGCT CAGACACGTC ATTTTACGAA AAGATGCCCT TCTCGAAACT GTGAAATCGC 1680
TCCAGGGAGA AATGAAAGG TTAGATATCTC ACCAAAGAATT GTTACCAGAA CTGAAAAATA 1740
CCATAAAAGA CCTTCAAGAA AAGATGAGG TATACCTACTCT TAGCATTACG CAAGAGATA 1800
CCATGTTAAA AGAATTGAGG GGAAGATAAA ATTCCTCTAC TGAGGAAAAA GATGATTTTA 1860
TAAATTAACCT GAACAAATCC CATGGAAGAA TGGATAATTT CCATAAGAAA TGTTAGAGCS 1920
AAGAAAGATT GATCGTTGAC CTTGGGAAGA AAGTAGACCA AACTATCCAG TACAAACAGT 1980
AACTAGAACA AAGGT

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3642 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

GTCTGTCTGA AGCTCAGTCA GATTCCTCA TTGATACCTT TCCTGAGTGT AGTACGGAAG  
GCTTCTCCAG TGACAGTGAT GCTGTACTCT TTACTGTGGA TTGGATTCTT CTGCCGTGAGT  
TAGATGATGG AATGGATTCC AATCAAAATT CTCCCCATTAG AACCTTTGGA CTCAATCTTT  
CTTCGGATTC TACAGCACTA GGCGCTTAGC TTCTGACAC TGAACAGAGC AAAACAGAAG  
AGGAACGGGA AAGTCGTAAGC CTCTTTCTTG GCAAGTTAAGG GCGAAGAGCA AGGAAGAGAG  
ATTATTTGGA GAAAGGCGAGA AATTAAATGA AGCTGGCTTT AAAAAGGCAA GAAGAGAGAC  
ACTATGAAGGC TGCTCTGTAT TTTATATAGGA AGGGAGTTGA TTTACTCTTA GAAAGCTGTC  
AAGGAGAGTC AAGCCCTACC CGTCGAGAAG CTTGTAAGAG AGAGAACAGCC GAGTACCTCA  
TGGGGCGGAGA AAGTATCTCT AGCTCTTTAG GAAACCTCTA GCTTGATGG GCTATCTCAGC  
CTCCAGGATC ACTAAGTCCA AGGGCCCTTT GGAACCTAAG GAGCCCTGGC GAGGAGCTGA  
AGGCCTTCCAG AGTCCTTGGG GTGATGAGCA AGTGTCTACT TGTAATGGAC ACAAGGACAG  
AACACACCTTT CATTTTAANA GGTCTAAAGA AAGGCTAGTG AGTACGACGG AAGCAAAGAGA  
CCATCCNCCC CGCGTGTGTG CCCANCATGG TGGTGCTCCA TAAGTACATC ATCTCTGAG  
AGTCANTATT TCTTGCTGCT CAGCATGCGG AANCTGGCAA ACTGTTGGCA TATATCAGTA  
AATTTCCTAA CAGAACTCCTT GAAARAAAGCT TTGACATCAA GGAAATGAAA AAACCTACAC  
TTGGAACAAGT TCACCTGCTG CAGCCCAACTT CTAGTCTCTCA AGCACGACGT AGCTTGGAT  
CCAGAGGAGA TGTAGGGTGG AGCATTGCTTA AAGGCTGCTCC TTGGAGAGCT AGCTTACTC  
CAAGTTTCTCA AGATGACGCCAACAGGAAG ATGATGAGCA ATAGAAGCTCT CCAAAGGTGCG  
CAGATTCTGG TTCAAGTCTCA GAAGAAAGCAT GTACTACTAG TTATTTAAAC TTAGAAAGG  
AATATGAGCA AGAGGAAGATT GACCGAGGTT CTTTGAAATG GAGACCGCTCA ATGAAAGACTG  
AAGGGAATGG TGTTGATACA AAAGCTATTA AAAGCTTCCC AGCACAACCTT GCTGCTGACA  
GTGACAGGCC CAGCACAAGC CTGAGACCTCC AGCAGCTGAA GTGTCCTCCC AACAGTGACC  
CAGAACAGCT TACTTCTCTCA AGAAGATCG AGTCCCTCAG TTAGAAGAAA AAATAGCCCCA  
TGGAATTCTT TAGATAGAC AGTAAAGATA GCCGAAGTGA ACTCTGGGGA CTGAGCTTTG  
GAGAAAAATT GTATGACTCA AAATCGAAGC CTTTGAAACC ATTTCTTACT CTCCAGGATG  
GAGACAGTCC TCTAGGAGT TTTAATACTA GTGAAAGCCA GGTAGAGTTT AAAGCTCAGG  
ACACCATTAG CAGGGGCTCA GATGACCTAG TGCCAGTTAT TTCAATTAAAG GATGCCTGCTT  

60 120 180 240 300 360 420 480 540 600 660 720 780 840 900 960 1020 1080 1140 1200 1260 1320 1380 1440 1500 1560 1620
TTGATGATGT CAGTGGTACT GATGAAAGGAA GACCTGGATCT TCTTGTAAT TTACCTGGTG 1680
AATTTGAGTC AACAAGAGAA GCTCGACGCAA TGGGACCTAC TAAGTTTACA CAACAATCTA 1740
TAGGGATAAT AGAAATATAA CTCTTGGGAAG CCCTCTGATG TTATGCTCCT AAGGCTTAGTA 1800
CTGAACAAATG CCAAGCAGAT GAGGAAAGGA GCATAGAGGA ACTGAGTGTAC CCCTCTGGGC 1860
CCCAAATCTTA TAGTATAACA GAGAAACACT AGTCGACGGA GGAATCCAGG AGTGTATTTG 1920
TAGCCANCTGT TGATCATAGT AGTTGAGGAG ATATGCTTTT GCTACCCAGC TGAGATCTA 1980
AGTTTCAACT ACTTTGAGTG GTGAGTCAN CGTAAAATGC AAACACAACA GAAGAAAGCT 2040
TATCCGTTAT TTGTAGTCCA CTCTCAGGTT GCTATGAAATA TATTTGGAGGC ACAGACATT 2100
TAAAAACAGA AGAAGATATTG CTGTTTACAG ACAGACCTGA TGATTTGCTG AAAGAGGAC 2160
CAACTTCTTT ATCCCANAGA GACCTGAGA CTAAAGGTGA AAGTGGTTTA GTGCTAGAAG 2220
GAGACAAAGGA AATACATCAG ATTTTTGAGG GACCTTGTATA AAAATTAAGC ACTACNCTCCC 2280
AGTTTTACA TCCAGAGGG CTGACATTCA AGNTGGCGAG CTGAAATGGT GGTAGCCCTT 2340
NGATGCTTTA ACATAGAGAG GAAATTTGTT CCCTCCGATTG AACCCAAAAA ANATMTATT 2400
GATGATAGA GGACATCTTC AGNTAACGTAT TTTTACGAGG TGGAGTGGAG TGGAAGATCT 2460
CTGTTGACAGC GATGCCATAG AGGAAAGTGA CTGTATACCAA GAGGTTTGGAG CAATCTTCTA 2520
AGAAACTGAA GCCCTGTGTT GGTGGAGTTT GGGTCTGTGC CTCTCTGAA CTTNCTACGG 2580
CAAGACTCTTG ATTTGATGCC ATCCAGCCAG AAATAATACT CACACTACTT TGACATGCC 2640
AGAATGTGTC TCTGAGAGGG TCTGCTCAGT CATCAGACAG CTCTTGAGAT TCAATCTCCT 2700
GGAACGACTT GTGCTGAGG TGCTCTGGTG TGAAGATATC AAATCTCCTC CATTCTTTAC 2760
CCCTGTGGAT TGGCCGAGAC TGAATGAGAT AGCTAATATC AGGTTTATCT TCACACATTG 2820
TGATTTTCTC TGTTGACAGGCC ATCTCCAGCA GTGAAACGCC TCTGACTCTG AGTTACTAT 2880
GGAGACCAAA AGCATTTGGA TAAAGGCAGG TATAGGAAAT AGGGGGGAAA TGCTGACAAAA 2940
AGAAACATTTG GTTTCAATTT ACAAGATATT AGCTAATTTG GCCAGGAGCT GTTATATACA 3000
TATATACACA ACCAAGGTTG GATCTGAAAT TAAATCCACAT TTGGTGTTGC AGATAGAGTTG 3060
TAAAGCCAAC TGAAGAGATT CCTCAGAAGA GTCACCTCTGA TAGGAAGCTA GAGTGTAGA 3120
ATGAAAGTTT ACTTGACAGA AGGACCTTTA CATGCGAGCT AACAGTGCTT TTTGCTGACC 3180
AGGATTGGTT TATATATATA AATTAATATT TGCTTAAAAA TACACTAAAAA GTATAGAAAC 3240
AATGTTCATCA ATGAAAACTTA AAAGCAGAGAA AAAAGAAATT AACACATATT TCTGAGCGAA 3300
AACCTGTACC CTGTGAGTTG ATATATGTAG TTGAATGTGG TCCAGATTTA TTTCACACAA 3360
AAGACACTCC ATGTTGTCAG TTTTGTACTC TTTGTTGATA CTGCTTATTT AGAGAAAGGT 3420
TCATATAAAC ACTCAGCTTG TGCTCTCAAT AACCATTTTAC TTCGCCCATC TTCCTATTTT 3480
CTGCACCCCTC TGCTTGCTCC CTCAATTCTC GTCTTCTCA GA CACACCTGCA 3540
ACAAAAAGGG GAGGGGAGTG CCATTCCCC TTTGGAAGAAG CACGATATAC 3600
AAATAAACAT GGTGAAACTC TNAANAAAAA AAAAAAAA AA 3642
(2) INFORMATION FOR SEQ ID NO:17:
   (i) SEQUENCE CHARACTERISTICS:
       (A) LENGTH: 1397 base pairs
       (B) TYPE: nucleic acid
       (C) STRANDEDNESS: single
       (D) TOPOLOGY: linear

   (xii) SEQUENCE DESCRIPTION: SEQ ID NO:17:
       GTTCACTCTCA ATAGAAAGTG ACGTGTTGCTA CCTAGTGTAT CTGGAAAGAG CTGAAGTCTG 60
       CAAATCTGAA GATGGCGCCA GCCCTCCACTG TATGGACTCTCG ACTGAACTCTC CCAAGTGAC 120
       GGTGTGTCTG GAGCCGATGG ACGAGTGTCTG GAATGGGATC TTCCTACAGTT TATGTAACCA 180
       CATCTTCCCAC AGCCAGTGTC TACAGCTGCT TGGGAACTCC ACCTGGCACCTT CTTTGCCTGA 240
       CTGCTCAAGGC CCCCAGGGCT TAGAAGGAAA TAAGTTGTTTT GTTGTTGGTG TCGGAGAAA 300
       TCTTGGATT TTGTATATAT TGGCGCAGAC TGGATGCTGA CCCTGAAGTC GCTGACATGC 360
       TTATAGGACAC TTGAGGAAA CGCAGCAGAC GTATGGCACTG CAGCTTACCA ACCATCGGAT 420
       CTGGGACTAT GCTGGAGATA ACTATGTCCA TGAGCTGTGGT GCAAGTTAAA CAGATGGAAC 480
       AATAGTACAG TAGAATGGTG AGGGGATAC TTGGCCAGGA GAGAAATAAG ATGCTTACCA 540
       GTTGGAGTTA TCTATTTTAT TAACTGCAGCC CTTGGAATCT CAGCGGATCT ACTGGGAAAA 600
       GAAGGATAGTT CGGATAGAGA AGGGACAGAC AGAGGGAATT AAACAACTGAT AGACCCAGTC 660
       TAAAGAAGACA ATGGAGAGAT GTGATAACTT AAGCAGCAAAA CTAAGGATAC CTTCTAAAAA 720
       AAAGCAGTCT GTGAAAGAGA AGTGCACTCA GCTAAACACA AAAGTGCCCA AACTCACCA 780
       CGAGCTCAA GAGGACAGG AAATGAAACAA GTTTTGGCAA GCCAACCAAG TCCTCTTGCAG 840
       GAACAAAGCTA AAGAGGGAGG AGAGGTGTCTA GAAGGAGACC TGTTGACAAA AGACTCTGCA 900
       GATCCAGGAG ATCCAGGAGC AGCTGCCTCTG CGTCTATGTTA TACCTGGAGA CACAGCAGA 960
       AGATCACAAC TCGGCTGCTG GAGAGCCAGG AGGAATCCCA GGAGGGACAG ATCAACATCG 1020
       CCATGGCCCT GGGCTGAGGC CTGGCTCTCT CGGGGGCGAG TGGGAAGTTG CCCCCCCAGGA 1080
       AGGGCCGAG CAGAGGGGAG ATGTAGAGCTT GAGGAACTAC GACATCCTG AGACTGTTCT 1140
       CCGGAGACTG TGTAAGTGCT GTGGGGACCT TCAGCTAAAT TGAGGGTGCG GCCCTAAATAA 1200
       GTACAAGTGA GGATCAAGCC ACAGTGTGTT GCCCTCTTCCA TTTGCTAGTG TGGTAGTGAG 1260
TGAAATGTAA AAGTGGCTGAC TGGAGAGCTG ATAGAAAGGC GCTGCCGTTCC AAAAGGTTCTT

AAGAGTTTCG TAACTCTCAA TTCTAAATGAC CANTTGGCTT TAGAAGCCCCC

ACACTCTGCT GTGCATT

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 800 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

CGTAAAATGGA GCANACTTAA AATAAGACCT CTGTTGGAAAT TTAGTTCTCT CGTGAAGAGGT 60

AGAGGGATAG GTTAGTAAAG TGTATTGTAA AAACACAGGT TTATAGTTTTTT GCCTTTTATAA 120

TTAGCCACAG GTTTTTCAAT CATCCAGCTTT CATGATAGGT TTTTAGGGCC GTTTAGGCC 180

TCATCCCTCT TGCACTTAAAT GTCTTACATG TTACTTGTTA GCACATCACA TGTATCAGT 240

ATCACCATCT GNTTPTTTGG GATGGTCTGC AGCATTTCCC AAAAATTCTT ACCTGGTATG 300

TTGCAAAATG AATGTACTCA GACATCTTAA ATTTTTACTTT AGGGCAGACC AACTCTTTGA 360

GTCTCTCTTG GACTTATAA TACAGATAC TTAAGAGTGG GAATGTAAG CATAACCTAA 420

TTTCTTTTCC TATAGAGATT CTAATTCTAT TAAAATCTAT TTATCCTCTA GTTGAAATCC 480

TGCTGTTTTG GCAAACTGTT TCTTCTGCTT GTCTGACCTT GCAGAAAGCTG GGATGGTAA 540

TAGCATACTA ATGAAGAGAA TTAGAAGTAG TTTCAAAACGC TGCTCACCTC ATCCATTCTC 600

TGTGATCCCC TCTATCCGAT GGCCGCCACCA CCACCTGAGA AAACAGATTT TCTAGTACAG 660

GTGGGATTAA TGCTCTGAAA GGCTGGGCC AGAGGAATGA GCAATAAGGC AAGTGGTTCC 720

AAACTACCTG GAGGTTTACA AAAATATGT CCCAGAAAAAAAATATCT TACCAAGATA 780

CGTAAAAAAA AAAAAAAA

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1810 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGTAAAATGGA GCANACTTAA AATAAGACCT CTGTTGGAAAT TTAGTTCTCT CGTGAAGAGGT 60

AGAGGGATAG GTTAGTAAAG TGTATTGTAA AAACACAGGT TTATAGTTTTTT GCCTTTTATAA 120

TTAGCCACAG GTTTTTCAAT CATCCAGCTTT CATGATAGGT TTTTAGGGCC GTTTAGGCC 180

TCATCCCTCT TGCACTTAAAT GTCTTACATG TTACTTGTTA GCACATCACA TGTATCAGT 240

ATCACCATCT GNTTPTTTGG GATGGTCTGC AGCATTTCCC AAAAATTCTT ACCTGGTATG 300

TTGCAAAATG AATGTACTCA GACATCTTAA ATTTTTACTTT AGGGCAGACC AACTCTTTGA 360

GTCTCTCTTG GACTTATAA TACAGATAC TTAAGAGTGG GAATGTAAG CATAACCTAA 420

TTTCTTTTCC TATAGAGATT CTAATTCTAT TAAAATCTAT TTATCCTCTA GTTGAAATCC 480

TGCTGTTTTG GCAAACTGTT TCTTCTGCTT GTCTGACCTT GCAGAAAGCTG GGATGGTAA 540

TAGCATACTA ATGAAGAGAA TTAGAAGTAG TTTCAAAACGC TGCTCACCTC ATCCATTCTC 600

TGTGATCCCC TCTATCCGAT GGCCGCCACCA CCACCTGAGA AAACAGATTT TCTAGTACAG 660

GTGGGATTAA TGCTCTGAAA GGCTGGGCC AGAGGAATGA GCAATAAGGC AAGTGGTTCC 720

AAACTACCTG GAGGTTTACA AAAATATGT CCCAGAAAAAAAATATCT TACCAAGATA 780

CGTAAAAAAA AAAAAAAA
GCAGCTCCCA GTGCGCTGT TT AAAAAATGGA GGGGGGATAT GTGATCCCG AGCCAAAGGC 60
GCCGCGCCAG ACTTATCGGT TCATTCAACA ACCAGTATTG AGTGCCCTGCT GTGTTCGCA 120
GGCCCTTGGCA TAGGGCCTTG TCATAGCGGT GCATAGCGTA TGAAAAAGAT CTGTTCTGCG 180
TGGCCATTCG TAATATAAAA ATCTGGAAAAT CGAAATGCTCA CAAAATCTCTA AACTTTTTGA 240
GCTGCTGACAT TAGGCTCCAA ATGGAAAATT TCATAATCTGA CTTATATTGG GTGCGANTCA 300
AAACACAGGT GCACAACCAC CAGTTTCATGC AACATCCCCCA ATGGGAAAA AGACCGGGCC 360
AGCTCTCTTC TGCTGCAGTT TTTCTGCTCA CACCTGGATT TCCCCATGCA TCCCCACAAA 420
AAGTAATTAA ATGGCATCGG TGCCAGCTTG AGACCGCAAC AACAGGTTTC CCACAATGCC 480
CCACATGCGG CCCAGACCTG TGTTGCAATC TCATTCCATT TTTGTCTTGA TCTCTGCTG 540
TGTGCTTAAA ATATATTTGG GAATATGTCA AAAACACCTC AAAGATACCC TGTGAATATC 600
AGTGAATAAGA AAAAGAGGA GCTTTATTGT TTATCTATAG CACAGAAGAT CAAGTTGTGG 660
GAGAAACTCGT ACAGTGTGTG AAGTGAGAAA CATCTTACAG AAGAGTATGG TGTTGGAATG 720
ACCACCATAT TGACCTGAAA GAAACAGAGA GATAAACCTG TGAGTTTTTA TGCTGAAGATT 780
GATGAGCGCA TATATTGAAA AAATAGAAAA ACACTTCATA AGCTAAAAAA TGAAGATCTT 840
GATGCGTATG TGAAGAGCTG CATCAGCGTG CTGGCGAGTG AACACATGCC ACTTAATGGT 900
ATGCTGATACA TGAAAACAAGC AAAGATATAT CACAATGAAC TAAAATTTGA GGGGAACGCT 960
GAATATCCAA CAGGCTCGGT GCCAGAAATTT AAGAAAGAC ATGGCATTAA ATTTTTAAAG 1020
ACTTGTGCGA ATAAAGCATC TGCTGGTCAT GAAGCAACAG AGAAGTTTAC TGGCAATTTT 1080
AGTAATGATG ATGAAACAGA TGTTACTATT GAAGAGTTCA ATATGTCGTA TGAGAAAAAA 1140
ATAATGCTTG ACCTCTCTTC ATATACAAAA AAATACACCT CAGACAGCTG CAGTAAGCTG 1200
GAAGAAGAGG ATAATTTTTNA TTGTTTTTAC AGTAAATAATG AGGCCCTCAG TGTTCTTCTCA 1260
TGTGCTTAAA GTGAGTAAAC AAAATGTTGT CTGAATCAAG ATGATCATGA TGATAATGAT 1320
AATGAGATG ATGTTAACAC TGCAAGAAAA GTGCCCTTATG ACGACATGCT AAAATGTTGT 1380
GATGGGCTTA TTAAAGGACT AGAGCGACAT GCATCTATAT CAGAGCAAGA AATCATCTCA 1440
GGTTATAAAA TCAGAGAGAG ACCTCTCAAGA CAAAAGACAT CATTAATGAG CAGATGACT 1500
CGGGAAGAAA CATTTAAAAA AGCCCTTCCG AGGAATGCTT CTCCCTCTCT ACAGAGCAGA 1560
CTTCCTGTTT CCGCAACTGT TTCTGACTGT TCTTCTTACC TAAAAATAAA ATAAAAATAACA 1620
GTGTCATGTA ACCTTTTATG CAAACACAGA TCATACTTGG AAACCTGAAAG CCACTACTGTA 1680
TTGGTTATGG TTGCTTACGA GCTGATACAG GTATCTCGGT GACACTACTG TGCTGGGTTA 1740
CTAAACACTGA ATACACTATT TTTTCCCTTG TAAAAATA AAAANAAAAA NAAANAAAAA 1800
AAAAANANANA 1810
(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 70 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

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

(2) INFORMATION FOR SEQ ID NO:21:

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

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

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

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 214 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

cggcagcaga aggtgcttgg aaagcactat gtgtgtgttg gcaacagtg
agtatagggc aatggagac tcattaccaa ccagctgca ggcccagcag gatgtgtca
acatanttg tcattcaag acccggcagca acctgagaa caagtcgggc cttatcacac
tggctaatga ctgtgaagtg ctgaccacac tcac

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 375 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:
	
tatggacaca tttgagccag ccaaggagga ggatgattac gacgtgatgc aggaccgccga

gtcccttcag agtgtcctag agaacctccc aggtgtggtat ccaacaatg aagccattcg
aaatgntatg ggctccctgg ctcggccagc cccaaaggg cgaagaaggg acaagaaagga

ggaagacaag aagtgagact ggagggaag ggtagctgag tctgcttagg ggaactgccatg

ggaagcagcgg aatattaggt tagatgttgtg tttactgtaa ccaattacgc cttaataaag

ctgggcaact tttttaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa aaaaaaaaaaaa

aaaaaaaaaac tgcag

375

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 304 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CGGCACGAGA AAGCAGTAGT GTGTGTGCGG ACAACAGTGA GTATATGCGG AATGGAGACT  60
TCCTTACCCAC CAGGCTGAG GCCGAGCAGG ATGCTGTCAA CATAGTTTG CATTCACAAGA  120
CCCGCAGCAA CCGTGGAGAC AGCGTGGGGCC TTATCACACT GGCTAATGAC TGGTAAGTGC  180
TGACCAACT CACCCCCAGAC ACTGGCCGTA TCCTGTCCAA GCTACATACT GTCCAACCCA  240
AGGGCAAGAT CACCTCTGAC GCGGCCATCC GCGTTGCCCA TCTGGCTCTG AAGCACCAGAC  300
AAGG  304

(2) INFORMATION FOR SEQ ID NO:25:

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

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

Arg Cys Gly Gly Gly Gly  20

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 78 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

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

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

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 384 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

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

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

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 68 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

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

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

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 97 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

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

(ii) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 116 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

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

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

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: linear

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

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

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 768 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

TACGAGAGG AGGAGGAGGA GGCCTCGGAG GAGGAGGGCT TGAGGTCGA TGGGAGGG
60
GAGGATAGG AGGCTCGGAG GCCCTCGGAG GGCCCTCGGC CCAGCGAGGG CAGGGCGCG
120
(2) INFORMATION FOR SEQUENCE ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:
   (A) LENGTH: 642 base pairs
   (B) TYPE: nucleic acid
   (C) STRANDEDNESS: single
   (D) TOPOLOGY: linear

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

TTAAATAAAA CACGCAGGT GCTAAAGAAGA GGCATTTTTT CTTAAAGTTAT TTTAATAGGT  60
GGTATAGCG TATTTTAAA TTTAGAGTTT GCTTTTACAG TTTAAACATT AATATGCTTT 120
CTCTGCTATG TCTGAAGATAAAGATTTT ATTATGAGCT TTTACAGGTA TTTTTAATA 180
GAGCAAGCCT CATGAATTTT AAAATATAGAT AACCCCCCCC AACAAATTTT AACCAATTATT 240
TGGCTTTTGG CATGAATTTT GTGTTTGCAT CACCCATGGA TTTTTTTTAT 300
TCTATATGAA TATTTTTTCA TGGTTTGTAG GGGAAATTTG AGCTAACAAT TTTCAATTCC 360
CCAGTCTGCA AAAGAAGCAG AATTCTATTG CTTTTGCTTG CTTATAGTCA TTTAATCATT 420
ACTTTTACAT ATATTGCGTG TACTTTCTGGT TTCTTTAAAA ATATAGTAAA GGATGTTTTA 480
TGAAGTCAAGA GTATACATAT AATTTTATTG TACCTAAAA TTTAGCAGTG CACCTTTGAAG 540
TTTTTTTCT AATTATAGAT GTAATAAGAA ATTTCTTTTG TAAATTGGGA AAAATCCAATA 600
AAAAGGATAT TCAATTAAAA AAAAATAAAA AAAAATAAAA AA 642
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

CGGCACGAGC TGCCAGAGCC AAGGCCAGG CTTTGATTCA GAATCTCTCT CTGCTGCTAG 60
TGGATGCCCTC GGTGGGACC ATTCAGTGTC TTGAGAAAAT TCTCTGTGAG TTTGTGCAGA 120
AGGATGAGTT GAAACCAGCA GTGACCCANC TGCTGTGGGA GCGGGCCACC GAGAAAGTGC 180
CCTGCTGTCC TCTGGAACGC TGTTCCTCTG TCATGCTTTCT TGACATGATG GCACGA 236

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 333 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

CCGGCGGTAT TGCGGTGGGC CTGTAATCCC AGCTAAACTCA AGAGGCTGAG GCAGGAGAAT 60
CGCCTGAACC CAGAGGGGGA GTTGTAGTG AGCCGAAATC ACACCATTG ACCTCAGCTT 120
GGGCAACAAT AGCGACCTCC CATCTCAAT 7AADADADAA AATGGCCATCA CGCTCTTAA 180
AATGCAAGGC TTTCCTTTAA ATTAGCCTAA CTGAACCTGGCG TTTGAGCTGCT TCAACTTTGG 240
AATATATGTT TGCCAACTCTC CTTTGTTTCT AATGAATAAA TGGTTCATA TACTTATAGA 300
AAAAAAAAAAA AAAAAAAAAAAAAA AACTCC GAG 333

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1272 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

GCAAGATGTT GTGGGAAGGC ACTATGTTGT GTGTGGACAA CAGTGAGTAT ATGCGGAATG 60
GAGACTTCTT ACCACCAGGG CTGCGGGGCCC AGCGAGTGC TGTAACATA GTTTGTCAAT 120
CAAGAGCCCG CAGCAACCCCT GAGAACAACG TGGGCCCTTTAT CACACTGGCT AATGACTGTTG 180
AAGTGCTGAC CACACTCACCC CAGACACTTG GGGTATCCTT GTCCAAAGCTA CATACTGTCC 240
AACCAAAAGG CAAGATCACCC TCTTGACCAAG GCATCCTGGGT GCACCATCTCG GCTCTGAAGC 300
ACCGACAAAGG CAAGATGCAAC AAGATGCCGGA TCATTGCCTT TGGGGAAGCC CCAGTGAGG 360
ACAATGAGAA GGATCCTGGTG AAACCTGGCTA AACGCCCTCAA GAAGAGAGAA GATAATGGTG 420
ACATTATCAA TTTTGGGGAA GAGGAGTGAG ACACAGAAAA GCTGACAGCA CTTTGAAACA 480
CGTTGAAATGG CAAAAAGTGAAC ACCGTTTCTC ATCTGTTGAC AGTGCCCTCTT GGGCCAGATT 540
TGCTCTGATGC TCTCATAGT CTCTGCCATT TGCCCTGGTG AGGTGCTGCG ATGCTGGTGTC 590
TTGTTGACAG TGACTTGGAA TTGGGACTAG ATCCCCAGTGC TGATCCTGAG CTGCGCCTTG 660
CCCTCCGTGT ATCTATGGAA GAGCACGCGGA AGCGGCGAGA GAGGAGGGCC CGGCCGCGCAG 720
CTGCACGCTCC TGCCGCTGAG GCGCGGATTT CTACACCTGGAG CACTGAAAGAC TCACGAGGATG 780
CCCTGCTGAAG GATGACATCT ACGCCAGCAAG AGTGTGGGGGC CACTGGGGCT CCTGACTCAA 840
GCAGATGAGC TGAGGAAGAC CAGATGCTTT ATGCCATGCA GATGCTCGTG CAGGGACGAG 900
AGTTGGCCCA GGGCGGATCA GCAGACATTG ATGCCAGCCTC AGCTATGGAC ACATCTGAGC 960
CAGCCAAAGA GGAGGATGAT TACGACTGTA TGCCAGACCGG CGAGCTCCCT CAGACTGTCC 1020
TAGAGAACTT CCCAGGCTGT GATCCCAACA AATGAAGCCAT TGCAATGCTC ATGGGCTCCC 1080
TGCCCTCCAG GCCACCAAGG AGCAACAGAA GGACAAAGAG GAGGAAGACA AGAAGTGAGA 1140
CTGGAGGGAA AGGGTAGCTG AGTCTAGTTA GGCGACTGGA TGGAAGCAG GGAATATAGG 1200
GTTAGATGTG TGGTATCTCTT AACATTACCA GCCTAAATAA AGCTGTTGCAA CCTTTAAAAA 1260
AAAAAAAAAA AA 1272

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 206 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:37:
CGGCACGAGA TGCTCACAGC TTCTCCCAGGA AGATTACAGA GGCCATGGGC ATCATACTGCA 60
AGATGATGTA TGAAACACACA ACTACACTGG TGCAAGGAGGT GAATGGAATTCT TTTTGAGATGG 120
TCTTCCAAATT TGGGTTACCC CAGGCCCCTGT TTGGGCGTGCG CGGTATGCTG CTCCTCATCT 180
GGTCTAAGGA GCCTGGGCTGC CGGGAA 206
(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 341 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:
TACTAAAAAT AAAAATTAG CCGGCCGTAT TGGCGTGCAG CTGTAATGCC AGCTACTCA 60
GAGGCTGAGG CAGGAAGATC GCCTGAACCC AGAGGGCGAG GTTGTATGTA GCCGAATCA 120
CACCATTGCA CTCCAGCTTG GGCACACATA GCAGACCTCC ATCTCAAAATT AAAAAAAA 180
TGCTCTACAG CTCTTTAAAA TGCAAGCTT TCTCTTAAA ATTAGCCTAAT GAACTGCAGT 240
GAGCTGCTTC AACCTTGGAA TATATTTTTT CCAATCTCCT TGTTTCTAA TGAATAATAG 300
TTTTTATATA CTTTTAANGA GAGAAAAAAA AANAAAACCTGA G 341

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 293 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:
CGGCACGAGC CCAGGCCTTG TTTGGGGTGC GCAGTATCTG TCCTCTCATA TGTTCTAAGG 60
AGCCTGCGTA CCGGGAGGCC GTGCTTAATG CCTACGCACA ACTCTACCTC AACCCAAAG 120
GGGACTCTTG CAGAGCCAG GCCAGGGCTT TGATTCAGAA TCTTCTCTTG GTGCCTAGTG 180
ATGCCTCGGT TCAGGACAAG ATGATTCTTG ATGAAATACT CTGTTAGTTT GTGCAGAAGG 240
ATGAGTTGAA ACCAGCAGTG ACCCATGCTGG TGTTGGGAAC GCACCCGGAG AAA 293

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 350 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:
CGGCACGAGC TACCACCAGG TCCGCGCTGA GAATTTGGGA TCCCTGCGCC GCGTTAACAA  60
TGAAGCAAAGT TGGAAACGCC CCGGCTTTCG TCAGCAAGCT GTGGGACGCTT GTGGAGAAAGA  120
CCCCACTCAAA CGAGTTTCTACT ACCTGGAGAC AGAAGTGGCCA AGATTTTTCTG GCTCTGGGATG  180
AGCAACAGATT TGCAAAAGAA ATTCTTCCCA AATATTTCAAA GCACAATAAT ATGGGCAAGCT  240
TTGTGAGGCAA ACTGAAATAGT TATGTTTCCG GTAAAGTAAAT ACATATCGGAC TCCTGGAAATGG  300
TTAAGCAAGAGAGGATGTC CCTGGAATAT TCTCAGCATCC TTACTTCCAA  350

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 377 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  

(xii) SEQUENCE DESCRIPTION: SEQ ID NO:41:
TCCTAAAGCT TTCTCTGCTC CAGTTATTTT TATTAATAAT TTTCACTTGG GCTTATTTTT  60
AAAACTGGGA ACATAAAAGTG CCTGATCTT GTAACAACCTC ATTTGTTTCT TTTGCTTCAG  120
AGAAGTTCAT TTATGGTCAA AGACGTTTTAT TCATGGTCAA CAGGAAGAGC AAGTGTCAGC  180
TGAACTGCTG CTGTCTGATA GGCTTCCAGC TCCATATATA TAGAAAGATC GGGGTTGGGA  240
TGSSATGGAG TGAGCCCAT CCACTTACTT GGCAGCTGTAT TAAATAAAAG TTTTCCCGTT  300
TGTGTTTTTT TGAACCATAC TGTATTCAGA AATAATACAA ATGAATGTGG NAAAAAAA  360
AAAAAAAAAA ACTGAG  377

(2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 374 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  

(xiii) SEQUENCE DESCRIPTION: SEQ ID NO:42:
CGGCACGAGG CGCCACTTGC GAGCGCTGCA AGGGCGGCTTT TGGCAGCGCCT GAGAAGATCG  60
TGAACAGTAA TGGGGAGCTG TACCAGGAGC AGTGGTTTCCT GTGCCGTTCAG TGCTTCCAGC  120
AGTTGCCAGA AGGACTCTTC TATGAGTTTG AAGGAAGAAA GTACTGTGCAA CATGACTTTTC  180
AGATGCTCTT TGCCCTTTGC TGTCATCAGT GTGGTGAATT CATCATTGCG CGAGTTATCA 240
AAGCCATGAA TAAACAGCTGG CATCCGGAGT GCTTCCGCTG TGACTCTCTG CAGGAAGTTC 300
TGGCAGATAT CGGGTTTGTC AAGAATGCTG GGAGACACCT GTGTCGCCCTG TGTCATAATC 360
GTGAGAAAGC CAGA 374

(2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 492 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

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

CTTTGCATT TACAGTAAAG ATCAAGTCC CTTTGAGTG CATTAGTAC GAAAATATGTG 60
ACCACCAATG ACAACCTTGG GAGTATTTT TAAATATAT GCTATGAA ATAGAGAACAC 120
AGAACAGGGT TTGACTATA GGTCTTTGG GCATAATTCT CATATACCTC TACACGAGAA 180
ATATGGAGGA GAAAAACAGG CATTTACATA TATTTCCGTT CACTTGTGAAG ATGCATGACC 240
TGAAGTCAGAC TGTTTGAGTG TGTATAGAC GTGGCAGCAT CCGAGCATCT CCGCAGCCCA 300
GAGGTTCAT TGGCTGCTTT GCTCAGTCTT CTTTAAAAAT ATGAATAGT GGACAGGCAC 360
GGTGCCTCAG ATCTGAAATC CCAAGCATTG GGGAGGTCGA GCCAGGTCGA TCACGAGGTGC 420
AGGAGATCAG AACCATTCTG GCTAGCAGTG AAAAAACATC TCTACTACAA AAAAAAABAA 480
AAAAAACGTG AG 492

(2) INFORMATION FOR SEQ ID NO:44:

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

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

Ser Gln Ile Cys Glu Leu Val Ala His Glu Thr Ile Ser Phe Leu 1 5 10 15

(2) INFORMATION FOR SEQ ID NO:45:

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

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

Xaa Xaa Xaa Xaa Xaa Ser Ile Leu Asp Glu Val Ile Arg Gly Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:46:

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

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

Val Val Lys Thr Tyr Leu Ile Ser Ser Ile Pro Gln Gly Ala Phe Asn
1 5 10 15
Tyr Lys Tyr Thr Ala
20

(2) INFORMATION FOR SEQ ID NO:47:

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

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

Val Val Lys Thr Tyr Leu Ile Ser Ser Ile Pro Leu Gln Ala Phe Asn
1 5 10 15
Tyr Lys Tyr Thr Ala
20

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:
Xaa Ala Lys Lys Phe Leu Asp Ala Glu His Lys Leu Asn Phe Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:
Xaa Xaa Xaa Lys Ile Lys Lys Phe Ile Gln Glu Asn Ile Phe Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:
Xaa Lys Val Lys Val Gly Val Asn Gly Phe Gly Arg Ile Gly Arg Leu
1 5 10 15
Val Thr

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:
Xaa Tyr Gln Tyr Pro Ala Leu Thr Xaa Gln Lys Lys Glu Leu
1 5 10 15
(2) INFORMATION FOR SEQ ID NO:52:

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

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:
Xaa Pro Ala Val Tyr Phe Lys Xaa Xaa Phe Leu Asp Xaa Asp
1 5 10

(2) INFORMATION FOR SEQ ID NO:53:

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

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:
Xaa Pro Ala Val Tyr Phe Lys Glu Gln Phe Leu Asp Gly Asp Gly
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:54:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:
Xaa Xaa Val Ala Val Leu Xaa Ala Ser Xaa Xaa Ile Gly Gln Pro Leu
1 5 10 15

Ser Leu

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

Val Val Lys Thr Tyr Leu Ile Ser Xaa Ile Pro Leu Gln Gly Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:56:

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

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

Xaa Xaa Lys Thr Tyr Leu Ile Ser Ser Ile Pro Leu Gln Gly Ala
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:57:

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

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

Met Asp Ile Pro Gin Thr Lys Gln Asp Leu Glu Leu Pro Lys Leu
1 5 10 15
CLAIMS

1. A polypeptide comprising an immunogenic portion of a prostate protein having a partial sequence selected from the group consisting of SEQ ID Nos. 2, 4, 5, 6, 7 and 8, or a variant of said protein that differs only in conservative substitutions and/or modifications.

2. A polypeptide comprising an immunogenic portion of a prostate protein or a variant of said protein that differs only in conservative substitutions and/or modifications wherein said protein comprises an amino acid sequence of a portion thereof encoded by a DNA sequence selected from the group consisting of the sequences recited in SEQ ID Nos. 11 and 13-19, the complements of said sequences, and DNA sequences that hybridize to a sequence recited in SEQ ID Nos. 11 and 13-19, or a complement thereof under moderately stringent conditions.

3. A DNA molecule comprising a nucleotide sequence encoding the polypeptide of claims 1 or 2.

4. An expression vector comprising the DNA molecule of claim 3.

5. A host cell transformed with the expression vector of claim 4.

6. The host cell of claim 5 wherein the host cell is selected from the group consisting of E. coli, yeast and mammalian cell lines.

7. A pharmaceutical composition comprising the polypeptide of claims 1 or 2 and a physiologically acceptable carrier.

8. A vaccine comprising the polypeptide of claims 1 or 2 and a non-specific immune response enhancer.
9. The vaccine of claim 8 wherein the non-specific immune response enhancer is an adjuvant.

10. A vaccine comprising a DNA molecule and a non-specific immune response enhancer, the DNA molecule comprising a nucleotide sequence encoding the polypeptide of claims 1 or 2.

11. The vaccine of claim 10 wherein the non-specific immune response enhancer is an adjuvant.

12. A pharmaceutical composition for the treatment of prostate cancer comprising a polypeptide and a physiologically acceptable carrier, the polypeptide comprising an immunogenic portion of a prostate protein having a partial sequence selected from the group consisting of SEQ ID Nos. 1, 3, 20, 21, 25-31 and 44-57.

13. A vaccine for the treatment of prostate cancer comprising a polypeptide and a non-specific immune response enhancer, the polypeptide comprising an immunogenic portion of a prostate protein having a partial sequence selected from the group consisting of SEQ ID Nos. 1, 3, 20, 21, 25-31 and 44-57.

14. The vaccine of claim 13 wherein the non-specific immune response enhancer is an adjuvant.

15. A pharmaceutical composition according to claim 7, for use in the manufacture of a medicament for inhibiting the development of prostate cancer.

16. A vaccine according to claim 8, for use in the manufacture of a medicament for inhibiting the development of prostate cancer.
17. A method for detecting prostate cancer in a patient, comprising:
   (a) contacting a biological sample obtained from a patient with a binding
   agent which is capable of binding to the polypeptide of claims 1 or 2; and
   (b) detecting in the sample a protein or polypeptide that binds to the
   binding agent, thereby detecting prostate cancer in the patient.

18. The method of claim 17 wherein the binding agent is a monoclonal
   antibody.

19. The method of claim 17 wherein the binding agent is a polyclonal
   antibody.

20. A method for monitoring the progression of prostate cancer in a
    patient, comprising:
    (a) contacting a biological sample obtained from a patient with a binding
    agent that is capable of binding to the polypeptide of claims 1 or 2;
    (b) determining in the sample an amount of a protein or polypeptide that
    binds to the binding agent;
    (c) repeating steps (a) and (b); and
    (d) comparing the amount of polypeptide detected in steps (b) and (c) to
    monitor the progression of prostate cancer in the patient.

21. A method for detecting prostate cancer in a patient, comprising:
    (a) contacting a biological sample obtained from a patient with a binding
    agent which is capable of binding to a polypeptide, the polypeptide comprising an
    immunogenic portion of a prostate protein having a partial sequence selected from the group
    consisting of SEQ ID Nos. 1, 3, 20, 21, 25-31 and 44-57; and
    (b) detecting in the sample a protein or polypeptide that binds to the
    binding agent, thereby detecting prostate cancer in the patient.
22. The method of claim 21 wherein the binding agent is a monoclonal antibody.

23. The method of claim 21 wherein the binding agent is a polyclonal antibody.

24. A method for monitoring the progression of prostate cancer in a patient, comprising:
   (a) contacting a biological sample obtained from a patient with a binding agent that is capable of binding to a polypeptide, the polypeptide comprising an immunogenic portion of a prostate protein having a partial sequence selected from the group consisting of: SEQ ID Nos. 1, 3, 20, 21, 25-31 and 44-57;
   (b) determining in the sample an amount of a protein or polypeptide that binds to the binding agent;
   (c) repeating steps (a) and (b); and
   (d) comparing the amount of polypeptide detected in steps (b) and (c) to monitor the progression of prostate cancer in the patient.

25. A monoclonal antibody that binds to the polypeptide of claims 1 or 2.

26. A monoclonal antibody according to claim 25, for use in the manufacture of a medicament for inhibiting the development of prostate cancer.

27. The monoclonal antibody of claim 26 wherein the monoclonal antibody is conjugated to a therapeutic agent.

28. A method for detecting prostate cancer in a patient, comprising:
   (a) contacting a biological sample from a patient with at least two oligonucleotide primers in a polymerase chain reaction, wherein at least one of the
oligonucleotide primers is specific for a DNA molecule selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43; and

(b) detecting in the sample a DNA sequence that amplifies in the presence of the oligonucleotide primer, thereby detecting prostate cancer.

29. The method of claim 28, wherein at least one of the oligonucleotide primers comprises at least about 10 contiguous nucleotides of a DNA molecule selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43.

30. A method for detecting prostate cancer in a patient, comprising:

(a) contacting a biological sample from the patient with at least one oligonucleotide probe specific for a DNA molecule selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43; and

(b) detecting in the sample a DNA sequence that hybridizes to the oligonucleotide probe, thereby detecting prostate cancer.

31. The method of claim 30 wherein the probe comprises at least about 15 contiguous nucleotides of a DNA molecule selected from the group consisting of SEQ ID Nos. 9-19, 22-24 and 32-43.
Rat Prostate Extracts

<table>
<thead>
<tr>
<th>Immune Sera</th>
<th>Control Sera</th>
</tr>
</thead>
<tbody>
<tr>
<td>TFA</td>
<td>TFA</td>
</tr>
<tr>
<td>TX-100</td>
<td>TX-100</td>
</tr>
</tbody>
</table>

205 — 121 — 86 — 50.7 — 33.6 — 27.8 — 19.4 — 7.4

Fig. 1
Rat Prostate Extract

Non-reduced SDS-PAGE

<table>
<thead>
<tr>
<th>Mol. Wt. (kD)</th>
<th>Immune</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19.4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Fig. 2