

(12) PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. AU 199940322 B2
(10) Patent No. 760305

(54) Title
FAP alpha-specific antibody with improved producibility

(51)⁷ International Patent Classification(s)
C07K 016/00

(21) Application No: **199940322**

(22) Application Date: **1999.04.22**

(87) WIPO No: **WO99/57151**

(30) Priority Data

(31) Number	(32) Date	(33) Country
98107925	1998.04.30	EP

(43) Publication Date : **1999.11.23**

(43) Publication Journal Date : **2000.02.03**

(44) Accepted Journal Date : **2003.05.15**

(71) Applicant(s)
Boehringer Ingelheim International GmbH

(72) Inventor(s)
John Edward Park; Pilar Garin-Chesa; Uwe Bamberger; Wolfgang J. Rettig; Olivier Leger; Jose Saldanha

(74) Agent/Attorney
DAVIES COLLISON CAVE, 1 Little Collins Street, MELBOURNE VIC 3000

(56) Related Art
WELT ET AL (1994) J. CLINICAL ONCOLOGY 12:1193-1203



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁶ : C07K 16/00	A2	(11) International Publication Number: WO 99/57151 (43) International Publication Date: 11 November 1999 (11.11.99)
(21) International Application Number: PCT/EP99/02711 (22) International Filing Date: 22 April 1999 (22.04.99) (30) Priority Data: 98107925.4 30 April 1998 (30.04.98) EP (71) Applicant (for all designated States except US): BOEHRINGER INGELHEIM INTERNATIONAL GMBH [DE/DE]; D-55216 Ingelheim (DE). (72) Inventors; and (75) Inventors/Applicants (for US only): PARK, John, Edward [US/DE]; Widdersteinstrasse 162, D-88400 Biberach (DE). GARIN-CHESA, Pilar [ES/DE]; Magnolienweg 7, D-88400 Biberach (DE). BAMBERGER, Uwe [DE/DE]; Grünweiler 90, D-88416 Ochsenhausen (DE). RETTIG, Wolfgang, J. [DE/DE]; Magnolienweg 7, D-88400 Biberach (DE). LEGER, Olivier [FR/FR]; Le Clos Saint André, 22, rue Marc Courriard, F-74100 Annemasse (FR). SALDANHA, Jose [GB/GB]; 22A Lincoln Way, Enfield, Middlesex EN1 1TE (GB). (74) Agents: LAUDIEN, Dieter et al.; Boehringer Ingelheim GmbH, A Patente, D-55216 Ingelheim (DE).		(81) Designated States: AU, BG, BR, CA, CN, CZ, EE, HU, ID, IL, JP, KR, LT, LV, MX, NO, NZ, PL, RO, SG, SI, SK, TR, UA, US, UZ, VN, YU, ZA, Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>Without international search report and to be republished upon receipt of that report.</i>
(54) Title: FAP α -SPECIFIC ANTIBODY WITH IMPROVED PRODUCIBILITY (57) Abstract Recombinant antibody proteins are provided that specifically bind fibroblast activation protein alpha (FAPα) and comprise framework modifications resulting in the improved producibility in host cells. The invention also relates to the use of said antibodies for diagnostic and therapeutic purposes and methods of producing said antibodies.		

FAP α -specific antibody with improved producibility

Field of the invention

5 The present invention relates to antibody proteins that specifically bind fibroblast activation protein alpha (FAP α). The invention also relates to the use of said antibodies for diagnostic and therapeutic purposes and methods of producing said antibodies.

Background of the invention

10 The invasive growth of epithelial cancers is associated with a number of characteristic cellular and molecular changes in the supporting stroma. A highly consistent molecular trait of the reactive stroma of many types of epithelial cancer is induction of the fibroblast activation protein alpha (from now on referred to as FAP), a cell surface molecule of reactive stromal fibroblasts originally identified with monoclonal antibody F19 (Garin-Chesa P., Old L. J. and Rettig W. J. (1990) Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. *Proc. Natl. Acad. Sci.* **87**: 7235). Since the FAP antigen is selectively expressed in the stroma of a range of epithelial carcinomas, independent of location and histological type, an FAP-targeting concept has been developed for imaging, diagnosis and treatment of epithelial cancers and certain other conditions. For this purpose a monoclonal antibody termed F19 that specifically binds to FAP was developed and described in US Patents 5,059,523 and 25 WO 93/05804, which are hereby incorporated by reference in their entirety.

30 One serious problem that arises when using non-human antibodies for applications *in vivo* in humans is that they quickly raise a human anti-non-human response that reduces the efficacy of the antibody in patients and impairs continued administration. Humanization of non-human antibodies is commonly achieved in one of two ways: (1) by constructing non-human/human chimeric antibodies, wherein the non-human variable regions are joined to human constant regions (Boulianne G. L., Hozumi N. and Shulman, M. J. (1984) Production of

functional chimeric mouse/human antibody *Nature* **312**: 643) or (2) by grafting the complementarity determining regions (CDRs) from the non-human variable regions to human variable regions and then joining these "reshaped human" variable regions to human constant regions (Riechmann L., Clark M., Waldmann H. and Winter G. (1988) Reshaping human antibodies for therapy. *Nature* **332**: 323). Chimeric antibodies, although significantly better than mouse antibodies, can still elicit an anti-mouse response in humans (LoBuglio A. F., Wheeler R. H., Trang J., Haynes A., Rogers K., Harvey E. B., Sun L., Ghrayeb J. and Khazaeli M. B. (1989) Mouse/human chimeric monoclonal antibody in man: Kinetics and immune response. *Proc. Natl. Acad. Sci.* **86**: 4220). CDR-grafted or reshaped human antibodies contain little or no protein sequences that can be identified as being derived from mouse antibodies. Although an antibody humanised by CDR-grafting may still be able to elicit some immune reactions, such as an anti-allotype or an anti-idiotypic response, as seen even with natural human antibodies, the CDR-grafted antibody will be significantly less immunogenic than a mouse antibody thus enabling a more prolonged treatment of patients.

Another serious limitation relating to the commercial use of antibodies for diagnosis, imaging and therapy is their producibility in large amounts. In many instances recombinant expression of native, chimeric and/or CDR-grafted antibodies in cell culture systems is poor. Factors contributing to poor producibility may include the choice of leader sequences and the choice of host cells for production as well as improper folding and reduced secretion. Improper folding can lead to poor assembly of heavy and light chains or a transport incompetent conformation that forbids secretion of one or both chains. It is generally accepted that the L-chain confers the ability of secretion of the assembled protein. In some instances multiple or even single substitutions can result in the increased producibility of antibodies.

Because of the clinical importance of specific immunological targeting *in vitro* and *in vivo* of specific disease-related antigens for diagnosis and therapy in humans, there is a growing need for antibodies that combine the features of antigen specificity, low immunogenicity and high producibility.

Therefore, the problem underlying the present invention was to provide antibody proteins that combine the properties of specific binding to FAP, low immunogenicity in humans, and high producibility in recombinant systems.

5

Disclosure of the invention

The technical problem is solved by the embodiments characterized in the claims.

The present invention provides new antibody proteins having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said new antibody proteins specifically binding to fibroblast activation protein (FAP), characterised in that they have framework modifications resulting in the improved producibility in host cells as compared to a chimeric antibody having the variable regions of F19 and foreign constant regions.

As used herein, an "antibody protein" is a protein with the antigen binding specificity of a monoclonal antibody.

"Complementarity determining regions of a monoclonal antibody" are understood to be those amino acid sequences involved in specific antigen binding according to Kabat (Kabat E. A., Wu T. T., Perry H. M., Gottesman K. S. and Foeller C. (1991) *Sequences of Proteins of Immunological Interest* (5th Ed.). NIH Publication No. 91-3242. U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD.) in connection with Chothia and Lesk (Chothia and Lesk (1987) *J. Mol. Biol.* 196:901-917).

As used herein, the term "framework modifications" refers to the exchange, deletion or addition of single or multiple amino acids in the variable regions surrounding the individual complementarity determining regions. Framework modifications may have an impact on the immunogenicity, producibility or binding specificity of an antibody protein.

"Fibroblast activation protein (FAP)", also designated fibroblast activation protein alpha (FAP α), is a membrane-bound glycoprotein belonging to the serine protease gene family (WO 97/34927). No shed or secreted form of FAP is known.

5 FAP can be characterized by its binding to the monoclonal antibody F19 (F19 is obtainable from the hybridoma cell line with the accession No. HB 8269 deposited at the ATCC).

The term "fibroblast activation protein specific binding" of an antibody protein is defined herein by its ability to specifically recognise and bind FAP-expressing human cells. The binding specificity of the proteins of the invention can be determined by standard methods for the evaluation of binding specificity such as described in an exemplary fashion in example 6, 8 and example 12.

15 The term "chimeric antibody" refers to an antibody protein having the light and heavy chain variable regions as described in figures 17 and 18 and foreign constant regions. "Foreign constant regions" as defined herein are constant regions which are different from the constant regions of F19. For comparing an antibody protein of the invention to a chimeric antibody it is to be understood that

20 such a chimeric antibody must contain the same constant regions as said antibody protein. For the purpose of demonstration and comparison alone the human constant heavy and light chains as described in Figures 19 to 22 are used in an exemplary fashion.

25 To provide the antibody proteins of the present invention, the nucleic acid sequences of the heavy and light chain genes of the murine antibody designated F19 were determined from RNA extracted from F19 hybridoma cells (ATCC Accession No. HB 8269).

30 In one embodiment the present invention relates to antibody proteins having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said new antibody proteins specifically binding to fibroblast activation protein (FAP), characterized in that they have framework modifications resulting in the improved producibility in host cells as compared to a chi-

meric antibody having the variable regions of F19 and foreign constant regions, wherein said antibody protein is derived from the murine antibody designated F19 (ATCC Accession No. HB 8269).

5 To generate humanised FAP-specific antibody proteins a chimeric antibody was constructed, having variable regions of the light and heavy chains of F19 and human light and heavy constant regions, respectively. The construction and production of chimeric mouse/human antibodies is well known (Boulianne et al. (1984), referenced above) and demonstrated in an exemplary fashion in examples
10 1 and 2.

The variable regions of the antibody proteins of the present invention are typically linked to at least a portion of the immunoglobulin constant region (F_C), typically that of a human immunoglobulin. Human constant region DNA sequences
15 can be isolated in accordance with well-known procedures from a variety of human cells, but preferably immortalized B cells (see Kabat et al., *supra*, and WO 87/02671). Hence the antibody proteins of the invention may contain all or only a portion of the constant region as long as they exhibit specific binding to the FAP antigen. The choice of the type and extent of the constant region depends on
20 whether effector functions like complement fixation or antibody dependent cellular toxicity are desired, and on the desired pharmacological properties of the antibody protein. The antibody protein of the invention will typically be a tetramer consisting of two light chain/heavy chain pairs, but may also be dimeric, i.e. consisting of a light chain/heavy chain pair, e.g. a Fab or Fv fragment.

25 Therefore, in a further embodiment the invention relates to antibody proteins according to the invention, characterised in that they have a variable light chain region and a variable heavy chain region, each joined to a human constant region. In particular, the variable region of the light chain was joined to a human kappa
30 constant region and the variable region of the heavy chain was joined to a human gamma-1 constant region. Other human constant regions for humanising light and heavy chains are also available to the expert.

Therefore, in one particular embodiment the antibody proteins of the invention contain a human kappa constant region.

Also, in another particular embodiment the antibody proteins of the invention
5 contain a human gamma-1 constant region.

One particular "chimeric F19 antibody" protein (cF19) consists of the light and heavy chain variable and constant regions described in Figures 17 to 22. cF19 demonstrates specific binding and high avidity to the FAP antigen. As demonstrated in example 2, the expression of cF19 in COS cells (cells derived from
10 the kidney of an African green monkey) is poor, ranging from about 10 to 60 ng/ml, which is at least 10 fold less than most antibodies.

In an attempt to increase expression levels of cF19, the leader sequence of the F19 V_L region was changed by substitution of proline to leucine at position -9.
15 This single change in amino acid in the leader sequence resulted in at least doubling the amount of chimeric antibody produced in COS cells. For the expression of this particular chimeric antibody in COS cells the following mutated leader sequence of the light chain: MDSQAQVLMLLLLWVSGTCG, and the following leader
20 sequence of the heavy chain: MGWSWVFLFLLSGTAGVLS were used.

According to the invention the term "improved producibility" in host cells refers to the substantial improvement of expression levels and/or purified antibody yields when compared with the expression levels and/or antibody yields of a chimeric antibody without framework modifications as defined above. Two particular
25 but not limiting examples for demonstrating improved producibility are exemplified for the COS cell expression system (in examples 2 and 5) and for the CHO cell expression system (in example 10 and 11).

30 While the mutation of the leader sequence only leads to a doubling of the expression yield of the chimeric F19 antibody, a substantial improvement as defined herein refers to an improvement in expression level and/or purification yield of at least a factor of 10.

In a preferred embodiment, the invention refers to antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 10.

In more preferred embodiment, the invention refers to antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 20.

In a most preferred embodiment, antibody proteins, characterised in that their expression levels in crude media samples as determined by ELISA and/or purified antibody yields exceed the expression levels and/or purification yields of the chimeric antibodies without framework modifications by at least a factor of 100.

Improved producibility of the recombinant antibody proteins of the invention can be demonstrated for eukaryotic cells in general as shown for COS and CHO (Chinese hamster ovary derived cells) eukaryotic cells (see examples 5 and 11). In a further embodiment, the present invention relates to recombinant antibody proteins characterised in that they display improved producibility in eukaryotic cells.

In a preferred embodiment the present invention relates to antibody proteins, wherein said eukaryotic cell is a Chinese hamster ovary cell (CHO cell).

It was unexpectedly found that certain framework modifications of the light chain variable regions determine the improved producibility of the antibody proteins of the invention. Three versions of reshaped light chain variable regions, designated version A, B, and C, as described in Figures 1 to 6, were prepared.

Light chain variable region versions A, B, and C demonstrate substantially improved producibility in CHO cells (see example 11). While light chain variable region versions A and C differ from light chain variable region version B by only two common amino acid residues they display an even further substantial improvement in producibility. There is at least another 10 fold difference in antibody secretion levels between the human reshaped F19 light chain version B and versions A or C. Reshaped human F19 light chain version A and B only differ in their amino acid sequences by two residues at positions 36 (Tyr to Phe mutation) and 87 (Tyr to Asp mutation) (nomenclature according to Kabat). This negative effect on the secretory capability of antibodies containing the light chain variable region version B could have been indirect if the Tyr to Asp and Tyr to Phe mutations, considered individually or together, merely caused improper folding of the protein. But this is unlikely to be the case since antigen binding assays show that immunoglobulins containing F19 light chain version B have similar avidities to those paired with F19 light chain version A or C, suggesting that they were not grossly misfolded.

Residue 87 in reshaped human F19 light chain version B seems particularly responsible for the reduction of secretion when compared to versions A and C.

In a preferred embodiment, the present invention relates to antibody proteins according to the invention, wherein the amino acid in Kabat position 87 of the light chain region is not asparagine.

In a more preferred embodiment, the invention relates to antibody proteins according to the invention, wherein the amino acid in Kabat position 87 of the light chain region is selected from aromatic or aliphatic amino acids.

In a most preferred embodiment, the present invention relates to antibody proteins according to the invention, wherein the aromatic amino acid in Kabat position 87 of the light chain region is a tyrosine or phenylalanine.

In a further embodiment, the present invention also pertains to antibody proteins according to the invention, wherein the amino acid in Kabat position 36 of the light chain region is selected from aromatic amino acids.

5 In a particular embodiment the invention relates to the specific antibody proteins that may be prepared from the individually disclosed reshaped variable regions of the light and heavy chains.

Especially light chain variable region versions A and C are particularly suitable to practice the invention because of their exceptionally high producibility, while retaining full FAP-binding specificity and achieving low immunogenicity. This holds especially true when compared to the chimeric antibody having the variable regions of F19 and the same constant regions but also when compared to light chain version B.

15 Therefore, in one embodiment the present invention relates to antibody proteins that contain the variable region of the light chain as set forth in SEQ ID NO: 2.

20 In a further embodiment the invention also relates to antibody proteins, characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1.

In one embodiment the present invention relates to antibody proteins that contain the variable region of the light chain as set forth in SEQ ID NO: 6.

25 In a further embodiment the invention also relates to antibody proteins characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 5.

30 The present invention also discloses several different variable regions of the heavy chain that work particularly well with the variable regions of the light chain versions A and C in terms of improved producibility.

In one embodiment the invention relates to antibody proteins containing a variable region of the heavy chain as set forth in any one of SEQ ID NOs: 8, 10, 12, 14.

5 In another embodiment the invention relates to antibody proteins characterised in that the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in any one of SEQ ID NOs: 7, 9, 11, 13.

In a very particular embodiment the invention relates to antibody proteins
10 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NO: 12. Most preferably, this antibody protein additionally contains the constant region of the light chain as set forth in SEQ ID NO: 20 and the constant region of the heavy chain as set forth in SEQ ID NO: 22.

15 Thus a further aspect of the present invention is an antibody protein containing an amino acid sequence as set forth in SEQ ID NO: 2. More preferably, such an antibody protein further contains an amino acid sequence as set forth in SEQ ID NO: 12. More preferably, said antibody protein further contains an amino
20 acid sequence as set forth in SEQ ID NO: 20 and an amino acid sequence as set forth in SEQ ID NO: 22. A further aspect of the invention is an antibody protein as described in this paragraph which is conjugated to a radioisotope, preferably ^{131}I , ^{125}I , ^{186}Re , ^{188}Re , or ^{90}Y . An additional aspect of the present invention is a DNA molecule coding for an antibody protein as described in this paragraph. A further
25 aspect of the invention is a host cell carrying such a DNA molecule. Accordingly, a further aspect of the invention is a method of producing an antibody protein as described in this paragraph, said method comprising the steps of cultivating such a host cell under conditions where said antibody protein is expressed by said host cell, and isolating said protein. A further aspect of the invention is a pharmaceutical
30 composition comprising an antibody protein of the present invention and a pharmaceutically acceptable carrier. In a further particular embodiment the invention relates to antibody proteins characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and

the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 11.

In a further particular embodiment the invention relates to antibody proteins
5 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NO: 8.

In a further particular embodiment the invention relates to antibody proteins characterised in that the variable region of the light chain is encoded by a nucleotide
10 sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 7.

Humanization of the variable region of a murine antibody may be achieved employing methods known in the art. EP 0239400 discloses grafting of the CDRs
15 of a murine variable region into the framework of a human variable region. WO 90/07861 discloses methods of reshaping a CDR-grafted variable region by introducing additional framework modifications. WO 92/11018 discloses methods of producing humanized Ig combining donor CDRs with an acceptor framework that has a high homology to the donor framework. WO 92/05274 discloses the preparation of framework mutated antibodies starting from a murine antibody. Further
20 prior art references related to humanization of murine monoclonal antibodies are EP 0368684; EP 0438310; WO 92/07075, or WO 92/22653. Thus, the expert can produce the antibodies of the present invention starting from the publicly available murine monoclonal antibody F19 and employing techniques known in the art, e.g.
25 from WO 92/05274. DNA molecules coding for the antibody proteins of the present invention may of course also be obtained by state-of-the-art synthetic procedures, e.g. by chemical synthesis of appropriate oligonucleotides and subsequent ligation and amplification procedures (see e.g. Frank et al. (1987) *Methods Enzymol.* 154: 221-249).

In a further aspect, the present invention relates to nucleic acid molecules containing the coding information for the antibody proteins according to the invention as disclosed above. Preferably, a nucleic acid molecule according to the pre-

sent invention is a nucleic acid molecule containing a nucleotide sequence selected from SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, or 15.

5 A further aspect of the present invention is a recombinant DNA vector containing the nucleotide sequence of any one of the above-mentioned nucleic acids, especially when said nucleotide sequence is operationally linked to an expression control sequence as in expression vectors. Preferred is a recombinant DNA vector, said vector being an expression vector.

10 A further aspect of the present invention is a host cell carrying a vector as described, especially an expression vector. Such a host cell can be a prokaryotic or eukaryotic cell. Preferably, such a host cell is an eukaryotic cell, a yeast cell, or a mammalian cell. More preferably, said host cell is a CHO (Chinese hamster ovary) cell or a COS cell.

15 Accordingly, a still further aspect of the present invention is a method of producing antibody proteins according to the invention. Such a method comprises the steps of:

- 20 (a) cultivating a host cell as described above under conditions where said antibody protein is expressed by said host cell, and
(b) isolating said antibody protein.

25 Mammalian host cells, preferably CHO or COS cells are preferred. Host cells for producing the antibody proteins of the invention may be transfected with a single vector containing the expression units for both, the light and the heavy chain (see e.g. WO 94/11523). In one particular embodiment the method of producing antibody proteins according to the invention pertains to host cells, wherein said host cells are cotransfected with two plasmids carrying the expression units for the light and heavy chains respectively (See e.g. EP 0481790).

30 The antibody proteins of the invention provide a highly specific tool for targeting therapeutic agents to the FAP antigen. Therefore, in a further aspect, the invention relates to antibody proteins according to the invention, wherein said antibody protein is conjugated to a therapeutic agent. Of the many therapeutic

agents known in the art, therapeutic agents selected from the group consisting of radioisotopes, toxins, toxoids, inflammatogenic agents, enzymes, antisense molecules, peptides, cytokines, and chemotherapeutic agents are preferred. Among the radioisotopes, gamma, beta and alpha-emitting radioisotopes may be used as a therapeutic agent. β -emitting radioisotopes are preferred as therapeutic radioisotopes. $^{186}\text{Rhenium}$, $^{188}\text{Rhenium}$, $^{131}\text{Iodine}$ and $^{90}\text{Yttrium}$ have been proven to be particularly useful β -emitting isotopes to achieve localized irradiation and destruction of malignant tumor cells. Therefore, radioisotopes selected from the group consisting of $^{186}\text{Rhenium}$, $^{188}\text{Rhenium}$, $^{131}\text{Iodine}$ and $^{90}\text{Yttrium}$ are particularly preferred as therapeutic agents conjugated to the antibody proteins of the invention. For example, for the radioiodination of an antibody of the invention, a method as disclosed in WO 93/05804 may be employed.

A further aspect of the present invention pertains to antibody proteins according to the invention, characterised in that they are labelled. Such an FAP-specific labelled antibody allows for the localisation and/or detection of the FAP antigen *in vitro* and/or *in vivo*. A label is defined as a marker that may be directly or indirectly detectable. An indirect marker is defined as a marker that cannot be detected by itself but needs a further directly detectable marker specific for the indirect marker. Preferred labels for practicing the invention are detectable markers. From the large variety of detectable markers, a detectable marker selected from the group consisting of enzymes, dyes, radioisotopes, digoxigenin, and biotin is most preferred.

A further aspect of the present invention relates to antibody proteins according to the invention, characterised in that they are conjugated to an imageable agent. A large variety of imageable agents, especially radioisotopes, are available from the state of the art. For practising the invention gamma-emitting isotopes are more preferred. Most preferred is $^{125}\text{Iodine}$.

One aspect of the present invention relates to pharmaceutical compositions containing an antibody protein according to the present invention as described above and a pharmaceutically acceptable carrier. Such pharmaceutical compositions are useful for treating tumors, wherein said tumors are associated with acti-

vated stromal fibroblasts. There are two possible effector principles for an anti-tumor stroma immunotherapy that may act synergistically: (a) An unmodified (unconjugated, 'naked') antibody according to the invention may induce immune destruction or inflammatory reactions in the tumor stroma while (b) an antibody conjugated to a therapeutic agent, such as for example, a radioisotope or other toxic substance, may achieve localized irradiation and destruction of the malignant tumor cells. Accordingly, a further aspect of the present invention is the use of an antibody protein as described for the manufacture of a pharmaceutical composition, especially for the treatment of tumors.

One further embodiment are pharmaceutical compositions containing an antibody protein according to the invention conjugated to a therapeutic agent as described above and a pharmaceutically acceptable carrier useful for treating tumors, wherein said tumors are associated with activated stromal fibroblasts. Another embodiment pertains to pharmaceutical compositions containing an antibody protein according to the present invention conjugated to an imageable agent as described above and a pharmaceutically acceptable carrier useful for imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient. A most preferred embodiment relates to the pharmaceutical compositions mentioned above, wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, invasive bladder cancers, pancreatic cancers and cancers metastatic of the brain.

In an animal or human body, it can prove advantageous to apply the pharmaceutical compositions as described above via an intravenous or other route, e.g. systemically, locally or topically to the tissue or organ of interest, depending on the type and origin of the disease or problem treated, e.g. a tumor. For example, a systemic mode of action is desired when different organs or organ systems are in need of treatment as in e.g. systemic autoimmune diseases, or allergies, or transplantations of foreign organs or tissues, or tumors that are diffuse or difficult to localise. A local mode of action would be considered when only local manifestations of neoplastic or immunologic action are expected, such as, for example local tumors.

The antibody proteins of the present invention may be applied by different routes of application known to the expert, notably intravenous injection or direct injection into target tissues. For systemic application, the intravenous, intravascular, intramuscular, intraarterial, intraperitoneal, oral, or intrathecal routes are preferred. A more local application can be effected subcutaneously, intracutaneously, intracardially, intralobally, intramedullary, intrapulmonarily or directly in or near the tissue to be treated (connective-, bone-, muscle-, nerve-, epithelial tissue). Depending on the desired duration and effectiveness of the treatment, pharmaceutical antibody compositions may be administered once or several times, also intermittently, for instance on a daily basis for several days, weeks or months and in different dosages.

For preparing suitable antibody preparations for the applications described above, the expert may use known injectable, physiologically acceptable sterile solutions. For preparing a ready-to-use solution for parenteral injection or infusion, aqueous isotonic solutions, such as e.g. saline or corresponding plasma protein solutions are readily available. The pharmaceutical compositions may be present as lyophilisates or dry preparations, which can be reconstituted with a known injectable solution directly before use under sterile conditions, e.g. as a kit of parts. The final preparation of the antibody compositions of the present invention are prepared for injection, infusion or perfusion by mixing purified antibodies according to the invention with a sterile physiologically acceptable solution, that may be supplemented with known carrier substances or/and additives (e.g. serum albumin, dextrose, sodium bisulfite, EDTA).

The amount of the antibody applied depends on the nature of the disease. In cancer patients, the applied dose of a 'naked' antibody may be between 0.1 and 100 mg/m², preferably between 5 and 50 mg/m² per application. For radiolabeled antibodies, e.g. with iodine-131, the maximally tolerated dose (MTD) has to be determined which must not be exceeded in therapeutic settings. Application of radiolabeled antibody to cancer patients may then be carried out by repeated (monthly or weekly) intravenous infusion of a dose which is below the MTD (See e.g. Welt et al. (1994) *J. Clin. Oncol.* 12: 1193-1203).

Furthermore, one aspect of the present invention relates to the use of the antibody proteins according to the invention for the treatment of cancer. In a preferred embodiment the present invention relates to the use of antibody proteins according to the invention conjugated to a therapeutic agent as described above for the treatment of cancer. In another preferred embodiment the present invention relates to the use of antibody proteins according to the invention conjugated to an imageable agent for imaging activated stromal fibroblasts. In a further preferred embodiment the present invention relates to the use of labelled antibody proteins according to the invention for detecting the presence of activated stromal fibroblasts in a sample.

One aspect of the invention relates to a method of treating tumors, wherein the tumor is associated with activated stromal fibroblasts capable of specifically forming a complex with antibody proteins according to the invention, present as naked/unmodified antibodies, modified antibody proteins, such as e.g. fusion proteins, or antibody proteins conjugated to a therapeutic agent, which comprises contacting the tumor with an effective amount of said antibodies. In a preferred embodiment the present invention relates to a method of treating tumors as mentioned above, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, invasive bladder cancers, pancreatic cancers and metastatic cancers of the brain. The method of treating tumors as described above may be effected *in vitro* or *in vivo*.

A further aspect of the invention relates to a method of detecting the presence of activated stromal fibroblasts in wound healing, inflammation or in tumors, characterised in that

- (a) a sample, possibly containing activated stromal fibroblasts, is contacted with an antibody protein according to the invention under conditions suitable for the formation of a complex between said antibody and antigen,

- (b) detecting the presence of said complex, thereby detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor.

5 In a preferred embodiment, the present invention relates to a method of detecting the presence of activated stromal fibroblasts in a tumor, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and
10 metastatic cancers of the brain. Most preferred antibody proteins of the invention are those which are characterised in that they are labelled as mentioned above.

A further aspect of the invention relates to a method of imaging the presence of activated stromal fibroblasts in a healing wound, inflamed tissue (rheumatoid
15 arthritis and cirrhosis are also positive) or a tumor, in a human patient, characterised in that

- (a) an antibody protein according to the present invention conjugated to an imageable agent is administered to a human patient under conditions suitable for the formation of an antibody-antigen complex,
20 (b) imaging any complex formed in this manner,
(c) thereby imaging the presence of activated stromal fibroblasts in a human patient.

In a preferred embodiment the present invention relates to a method of imaging
25 the presence of activated stromal fibroblasts as described above in tumors, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancer, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

30 In a further aspect the present invention relates to a method of detecting tumor-stroma, characterised in that

- (a) a suitable sample is contacted with an antibody protein according to the present invention, under conditions suitable for the formation of an antibody-antigen complex,
- (b) detecting the presence of any complex so formed,
- 5 (c) relating the presence of said complex to the presence of tumor-stroma.

Antibody proteins for practicing the invention are preferably labelled with a detectable marker.

10 In a further aspect the present invention relates to a method of imaging tumor-stroma in a human patient, which comprises

- (a) administering to the patient an antibody according to the invention conjugated to an imageable agent as described above under conditions suitable for the formation of an antibody-antigen complex,
- 15 (b) imaging any complex so formed, and thereby imaging the presence of tumor-stroma in a human patient.

Figure legends

20 **Fig. 1.** DNA sequence of F19 human reshaped light chain variable region version A (hF19L_A) SEQ ID NO:1.

25 **Fig. 2.** Amino acid sequence of F19 human reshaped light chain variable region version A (hF19L_A) SEQ ID NO: 2.

Fig. 3. DNA sequence of F19 human reshaped light chain variable region version B (hF19L_B) SEQ ID NO: 3. Nucleotides differing from version A are underlined and in bold type.

30 **Fig. 4.** Amino acid sequence of F19 human reshaped light chain variable region version B (hF19L_B) SEQ ID NO: 4. Amino acids differing from version A are underlined and in bold type.

5

Fig. 5. DNA sequence of F19 human reshaped light chain variable region version C (hF19L_C) SEQ ID NO:5. Nucleotides differing from version A are underlined and in bold type.

10 **Fig. 6.** Amino acid sequence of F19 human reshaped light chain variable region version C (hF19L_C) SEQ ID NO: 6. Amino acids differing from version A are underlined and in bold type.

Fig. 7. DNA sequence of F19 human reshaped variable region heavy chain version A (hF19H_A) SEQ ID NO: 7.

15

Fig. 8. Amino acid sequence of F19 human reshaped heavy chain variable region version A (hF19H_A) SEQ ID NO: 8

20 **Fig. 9.** DNA sequence of F19 human reshaped heavy chain variable region version B (hF19H_B) SEQ ID NO: 9. Nucleotides differing from version A are underlined and in bold type.

Fig. 10. Amino acid sequence of F19 human reshaped heavy chain variable region version B (hF19H_B) SEQ ID NO: 10. Amino acids differing from version A are underlined and in bold type.

25

Fig. 11. DNA sequence of F19 human reshaped heavy chain variable region version C (hF19H_C) SEQ ID NO: 11. Nucleotides differing from version A are underlined and in bold type.

30

Fig. 12. Amino acid sequence of F19 human reshaped heavy chain variable region version C (hF19H_C) SEQ ID NO: 12. Amino acids differing from version A are underlined and in bold type.

35

5 **Fig. 13.** DNA sequence of F19 human reshaped heavy chain variable region version D (hF19H_D) SEQ ID NO: 13. Nucleotides differing from version A are underlined and in bold type.

10 **Fig. 14.** Amino acid sequence of F19 human reshaped heavy chain variable region version D (hF19H_D) SEQ ID NO: 14. Amino acids differing from version A are underlined and in bold type.

15 **Fig. 15.** DNA sequence of F19 human reshaped heavy chain variable region version E (hF19H_E) SEQ ID NO: 15. Nucleotides differing from version A are underlined and in bold type.

20 **Fig. 16.** Amino acid sequence of F19 human reshaped heavy chain variable region version E (hF19H_E) SEQ ID NO: 16. Amino acids differing from version A are underlined and in bold type.

Fig. 17. Amino acid sequence of F19 chimeric light chain variable region (chF19LC) SEQ ID NO: 17.

25 **Fig. 18.** Amino acid sequence of F19 chimeric heavy chain variable region (chF19HC) SEQ ID NO: 18.

Fig. 19. DNA sequence of human kappa light constant chain SEQ ID NO: 19.

Fig. 20. Amino acid sequence of human light constant chain SEQ ID NO: 20.

30 **Fig. 21.** DNA sequence of human heavy constant chain SEQ ID NO: 21.

Fig. 22. Amino acid sequence of human heavy constant chain SEQ ID NO: 22.

5 **Fig. 23.** *Mammalian cell expression vectors used to produce chimeric and reshaped human antibodies with human kappa light chains and human gamma-1 heavy chains.*

A. Light chain expression vector: pKN100

B. Heavy chain expression vector: pG1D105

10

Fig 24. *DNA and amino acid sequences of mouse F19 light chain variable region as modified for use in the construction of chimeric F19 light chain. Restriction sites are indicated by bold letters. The Kozak sequence, CDRs 1 to 3 and the splice donor site are underlined.*

15

Fig 25. *DNA and amino acid sequences of mouse F19 heavy chain variable region as modified for use in the construction of chimeric F19 heavy chain. Restriction sites are indicated by bold letters. The Kozak sequence and the splice donor site are underlined.*

20

Fig. 26. *DNA sequence of F19 chimeric antibody cloned into pKN100 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDRs 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the mouse F19 light chain inside the pKN100 eukaryotic expression vector. This vector has a cDNA version of the human kappa constant region gene (allotype Km(3)) terminated by a strong artificial termination sequence. In addition, the Neo selection gene is also terminated by this artificial sequence and is also in the same orientation as the kappa light chain expression cassette.*

25

30 The essential components of the pKN100 eukaryotic expression vector are:

1 - 6 = EcoRI site

7 - 1571 = HCMV promoter/enhancer

583 - 587 = TATAA box

35 610 = Start of transcription

- 5 728 - 736 = Splice donor site
 731 = Beginning of intron
 1557 = End of intron
 1544 - 1558 = Splice acceptor site
 1590 - 1598 = Kozak sequence
 10 1599 - 1658 = peptide leader sequence
 1659 - 1997 = mouse F19 light chain
 1996 - 2004 = splice donor site
 2011 - 2657 = cDNA copy of human Kappa constant region (Km(3)) gene
 2664 - 2880 = Artificial spaC2 termination sequence
 15 2887 - 7845 = This is the pSV2neo vector DNA fragment comprising of
 the Amp-resistance gene (in the opposite orientation),
 the ColEI and SV40 origins of replication and the
 Neo-resistance gene (in the same orientation as the
 HCMVi-KCT cassette)
 20 7852 - 8068 = Artificial spaC2 termination signal

This sequence ends immediately upstream of the EcoRI site (position 1-6) at the beginning of the sequence. As a vector this DNA sequence would be circular.

25

Fig. 27. *DNA sequence of F19 chimeric antibody cloned into pg1d105 mammalian expression vector.* Restriction sites are indicated by bold letters and underlined. CDRs 1 to 3 and the splice donor site are underlined. This is the DNA sequence of the eukaryotic expression vector pG1D105 containing the mouse

30 F19 heavy chain variable region. This vector contains a cDNA version of the human gamma-1 constant region (allotype G1m (non-a, z, -x) also known as Gm1(17) allotype).

The essential components of the construct are:

- 35 1 - 2501 = pBR322 based sequence including Ampicillin resistance

- 5 gene and ColEI origin plus the SV40 origin and the
crippled SV40 early promoter
- 2502 - 3226 = dhfr gene
- 3233 - 4073 = SV40 poly A sequence etc.
- 4074 - 4079 = ligated BamHI and BglII site (BstYI)
- 10 4080 - 4302 = SPA site plus C2 termination signal
- 4303 - 5867 = HCMVi promoter
- 5879 - 5885 = unique HindIII restriction site for cloning of
immunoglobulin variable genes
- 5886 - 5894 = Kozak sequence
- 15 5895 - 5951 = signal peptide
- 5952 - 6323 = mouse F19 heavy chain
- 6323 - 6330 = splice donor site
- 6331 - 6336 = unique BamHI restriction site for cloning of
immunoglobulin variable genes
- 20 6337 - 7388 = cDNA copy of human gamma-1 constant regions preceded by a
62 bp intron
- 7389 - 7709 = Arnie termination sequence

25 The human gamma-1 constant region used in this construct has a G1m (non-a,
z, -x) also known as Gm1(17) allotype which is defined by a glutamic acid (E)
residue at position 356 (according to Eu numbering), a methionine (M) residue
at position 358 (according to Eu numbering) and a lysine (K) residue at position
214 (according to Eu numbering). These three residues are underlined in the
sequence above.

30

Fig. 28. *PCR-based method for the construction of human reshaped F19 light chain.*

This figure provides a schematic overview of the strategy of construction. The
dotted lines indicate a complementary sequence of at least 21 bases between
the primers.

35

5 **Fig. 29.** Nucleotide and deduced amino acid sequences of reshaped human F19 light chain variable regions version A, B and C. Nucleotide and deduced amino acid sequences are aligned and compared with that of version A, dashes indicate nucleotide identity, dots indicate amino acid identity with this sequence. Amino acids are numbered according to Kabat *et al.* (1991). The locations of CDRs are indicated in boxes.

Fig. 30. DNA sequence of F19 L_A (human reshaped light chain version A) cloned into pKN100 mammalian expression vector. Restriction sites are indicated by bold letters and underlined. CDRs 1 to 3 and the splice donor site are underlined.

15 This is the DNA sequence of the reshaped F19 light chain version A cloned into pKN100 eukaryotic expression vector. This vector has a cDNA version of the human kappa constant region gene (allotype Km(3)) terminated by a strong artificial termination sequence. In addition, the Neo selection gene is also terminated by this artificial sequence and is also in the same orientation as the

20 kappa light chain expression cassette.

The components of the vector are:

7 - 1571 = HCMVi promoter/enhancer

583 - 587 = TATAA box.

25 610 = Start of transcription.

728 - 736 = Splice donor site.

731 = Beginning of intron.

1557 = End of intron.

1544 - 1558 = Splice acceptor site.

30 1590 - 1598 = Kozak sequence

1599 - 1658 = peptide leader sequence

1659 - 1997 = reshaped F19 light chain version A

1996 - 2004 = splice donor site

2011 - 2657 = cDNA copy of human kappa constant region (Km(3)) gene.

35 2664 - 2880 = Artificial spaC2 termination sequence.

- 5 2887 - 7845 = This is the pSV2neo vector DNA fragment comprising of
 the Amp-resistance gene (in the opposite orientation),
 the ColEI and SV40 origins of replication and the
 Neo-resistance gene (in the same orientation as the
 HCMVi-KCT cassette).
- 10 7852 - 8068 = Artificial spaC2 termination signal.

This sequence ends immediately upstream of the EcoRI site (position 1-6) at the beginning of the sequence below. As a vector this DNA sequence would be circular.

15

Fig. 31. *PCR-based method for the construction of human reshaped F19 heavy chain.*
 This figure provides a schematic overview of the strategy of construction. The dotted lines indicate a complementary sequence of at least 21 bases between the primers.

20

Fig. 32. *Nucleotide and deduced amino acid sequences of reshaped human F19 heavy chain variable region versions a to e.* Nucleotide and deduced amino acid sequences are aligned and compared with that of version A, dashes indicate nucleotide identity, dots indicate amino acid identity with this sequence. Amino acids are numbered according to Kabat *et al.* (1991). The location of CDRs is indicated by boxes.

25

Fig. 33. *DNA sequence of F19Ha (human reshaped heavy chain version a) cloned into pg1d105 mammalian expression vector.* Restriction sites are indicated by bold letters and underlined. CDRs 1 to 3 and the splice donor site are underlined.
 This is the DNA sequence of the eukaryotic expression vector pG1D105 containing the reshaped version A of F19 heavy chain variable region. This vector contains a cDNA version of the human gamma-1 constant region (allotype G1m (non-a, z, -x) also known as Gm1(17) allotype).

30

35

The essential components of the construct are:

5

1 - 2501 = pBR322 based sequence including Ampicillin resistance gene
and ColEI origin plus the SV40 origin and the crippled SV40 early
promoter

2502 - 3226 = dhfr gene

10 3233 - 4073 = SV40 poly A sequence etc.

4080 - 4302 = SPA site plus C2 termination signal

4303 - 5867 = HCMVi promoter/enhancer

5879 - 5885 = unique HindIII restriction site for cloning of
immunoglobulin variable genes

15 5886 - 5894 = Kozak sequence

5895 - 5951 = signal peptide

5952 - 6323 = reshaped F19 heavy chain version A

6323 - 6330 = splice donor site

6331 - 6336 = unique BamHI restriction site for cloning of
20 immunoglobulin variable genes

6337 - 7388 = cDNA copy of human gamma-1 constant regions preceded by a
62 bp intron

7389 - 7709 = Arnie termination sequence

25

The human gamma-1 constant region used in this construct has a G1m (non-a,
z, -x) also known as Gm1(17) allotype which is defined by a glutamic acid (E)
residue at position 356 (according to Eu numbering), a methionine (M) residue
at position 358 (according to Eu numbering) and a lysine (K) residue at position
214 (according to Eu numbering). These three residues are underlined in the
30 sequence above.

30

Fig. 34. Heavy (panel A) and light (panel B) chains RNA splicing events taking place
during antibody F19 expression in mammalian cells – schematic overview.

35

A. Heavy chain RNA splicing

5

B. Kappa light chain RNA splicing

Fig. 35. *Concentration dependence of L_AH_C supernatant binding to CD8-FAP.*

10

Fig. 36. *Binding of biotinylated L_AH_C to human FAP.*

Fig. 37. *CD8-FAP carries the F19 epitope as detected with cF19.*

15

Examples

Example 1: Construction of mouse - human chimeric genes

20

The chimeric F19 (cF19) antibody was designed to have the mouse F19 V_L and V_H regions linked to human kappa and gamma-1 constant regions, respectively. PCR primers were used to modify the 5'- and 3'- sequences flanking the cDNA sequences coding for the mouse F19 V_L and V_H regions (Table 1). PCR primers specific for F19 light chain V-region were designed. These adapted mouse F19 variable regions were then subcloned into mammalian cell expression vectors already containing the human kappa (pKN100 vector) or gamma-1 (pG1D105 vector) constant regions (Figure 23). These vectors employ the human cytomegalovirus (HCMV) promoter/enhancer to efficiently transcribe the light and heavy chains. The vectors also contain the SV40 origin of replication to permit efficient DNA replication and subsequent protein expression in COS cells. The expression vectors were designed to have the variable regions inserted as HindIII-BamHI DNA fragments. PCR primers were designed to introduce these restriction sites at the 5'- (HindIII) and 3'- (BamHI) ends of the cDNAs coding for the V-regions. In addition the PCR primers were designed to introduce the Kozak sequence (GCCGCCACC) at the 5'-ends of both the light and heavy chain cDNAs to allow efficient

35

translation (Kozak M.: At least six nucleotides preceding the AUG initiator codon enhance translation in mammalian cells. (1987) *J. Mol. Biol.* **196**: 947), and to introduce splice donor sites at the 3'-ends of both the light and heavy chain cDNAs for the variable regions to be spliced to the constant regions. The PCR primers used in the construction of the chimeric F19 light and heavy chains are shown in Table 1. The DNA and amino acid sequences of the mouse F19 V_L and V_H regions as adapted for use in the construction of chimeric F19 light and heavy chains are shown in Figures 24 and 25. The DNA sequences of mouse F19 light and heavy chains cloned into the eukaryotic expression vectors pKN100 and pG1D105, respectively, are shown in Figures 26 and 27.

TABLE 1: PCR primers for the construction of chimeric F19 antibody.

A. Light chain variable region

1. Primer for the construction of the 5'-end (37mer)

5' CAGA **AAGCTT** GCCGCCACC ATG GAT TCA CAG GCC CAG 3'

HindIII Kozak sequence M D S Q A Q

2. Primer for the construction of the 3'-end (35mer)

5' CCGA **GGATCC** ACTCACG TTT CAG CTC CAG CTT GGT 3'

BamHI Splice donor site

B. Heavy chain variable region

1. Primer for the construction of the 5'-end (37mer)

5' CAGA **AAGCTT** GCCGCCACC ATG GGA TGG AGC TGG GTC 3'

HindIII Kozak sequence M G W S W V

2. Primer for the construction of the 3'-end (35mer)

5 5' CCGA **GGATCC** ACTCACC TGA GGA GAC GGT GAC TGA 3'

BamHI Splice donor site

10 **Example 2: Expression and binding activity of chimeric F19 antibody**

The two plasmid DNAs coding for the chimeric F19 light and heavy chains (see example 1) were co-transfected into COS cells to look for transient expression of chimeric F19 antibody as described below. After incubation for 72 hours, the medium was collected, centrifuged to remove cellular debris, and analysed by ELISA for the production of a human IgG1-like antibody. The COS cell supernatant containing the chimeric F19 antibody was analysed for its ability to bind to HT 1080 cells (see example 13) expressing the FAP antigen on their surface.

20 Transfection of COS cells using electroporation

The mammalian expression vectors pg1d105 and pKN100 containing the chimeric or reshaped human heavy and light chains versions, respectively, were tested in COS cells to look for transient expression of F19 antibodies. COS7 cells were passaged routinely in DMEM (Gibco BRL cat. #41966) containing penicillin (50 IU/ml), streptomycin (50 µg/ml), L-glutamine and 10% heat-inactivated gamma globulin-free foetal calf serum (FCS, Harlan Sera-Lab cat. # D0001). The DNA was introduced into the COS cells by electroporation using the Gene Pulsar apparatus (BioRad). DNA (10µg of each vector) was added to a 0.8 ml aliquot of 1×10^7 cells/ml in Phosphate-buffered saline (PBS, Ca^{2+} and Mg^{2+} free). A pulse was delivered at 1,900 volts, 25 µF capacitance. After a 10 min recovery period at ambient temperature the electroporated cells were added to 8 ml of DMEM containing 5% FCS. After incubation at 37°C for 72 hours, the medium was collected, centrifuged to remove cellular debris, and stored under sterile conditions at 4°C for short periods of time, or at -20°C for longer periods.

5 ELISA method for measuring assembled IgG1/kappa antibody concentrations in COS cell supernatants

Samples of antibodies produced in transfected COS cells were assayed by ELISA to determine how much chimeric or reshaped human antibody had been produced. For
10 the detection of antibody, plates were coated with goat anti-human IgG (Fc γ fragment specific) antibody (Jackson ImmunoResearch Laboratories Inc., #109-005-098). The samples from COS cells were serially diluted and added to each well. After incubation for 1 h at 37°C and washing, horseradish peroxidase conjugated goat anti-human kappa light chain (Sigma, A-7164) was added. After incubation for 30 minutes at 37°C and
15 washing, K-blue substrate (a mixture of 3,3',5,5' tetramethylbenzidine and hydrogen peroxide, Bionostics Limited, #KB175) was added for 30 minutes at room temperature. The reaction was stopped using Red Stop solution (Bionostics Limited, #RS20) and the optical density read on a microplate reader at 650 nm. Purified human IgG1/Kappa antibody (Sigma, I-3889) of known concentration was used as a standard.

20

The expression of chimeric F19 antibody in COS cells was poor (Table 2), between 10 and 60 ng/ml, which is at least 10 fold less than most antibodies.

In an attempt to increase expression levels of the chimeric F19 antibody, the
25 leader sequence of F19 V_L region was changed by substitution of leucine to proline at position -9. This single change in amino acid in the leader sequence resulted in at least doubling the amount of chimeric antibody produced in COS cells.

Cell binding results show that chimeric F19 binds specifically and with the ex-
30 pected avidity to the FAP target.

5 **TABLE 2: Chimeric F19 antibody concentrations in COS cell supernatants(These are the results of three independent transfections)**

Transfected Antibody components		Human $\gamma 1/K$
Heavy chain	Kappa light chain	[in $\mu\text{g/ml}$]
cF19	cF19 (F19 leader sequence)	0.060
cF19	cF19 (mutated leader sequence)	0.212
cF19	cF19 (F19 leader sequence)	0.056
cF19	cF19 (mutated leader sequence)	0.108
cF19	cF19 (F19 leader sequence)	0.011
cF19	cF19 (mutated leader sequence)	0.087

10 **Example 3: Construction of the reshaped human F19 light chain versions A to C ($L_A - L_B$)**

The construction of the first version of reshaped human F19 V_L region (L_A) was carried out using overlapping PCR fragments in a method similar to that described by Daugherty B. L., DeMartino J. A., Law M. F., Kawka D. W., Singer I. I. and Mark G. E. (1991) Polymerase chain reaction (PCR) facilitates the cloning, CDR-grafting, and rapid expression of a murine monoclonal antibody directed against the CD18 component of leukocyte integrins. *Nucl. Acids Res.* 19: 2471. Ten oligonucleotides were synthesised that consisted of five primer pairs, APCR1-vla1, vla2-vla3, vla4-vla5, vla6-vla7, and vla8-APCR4 (Table 3 and Figure 28). There was an overlapping sequence of at least 21 bases between adjacent pairs (Figure 28). APCR1 and APCR4 hybridised to the flanking pUC19 vector sequences. The mutagenic primers were designed such that their 5' end immediately followed the wobble position of a codon. This strategy was used to counteract the gratuitous addition of one nucleotide to the 3' end of the strand complementary to the mutagenic primer by the DNA polymerase during PCR (Sharrocks A. D. and Shaw P. E. (1992) Improved primer design for PCR-based, site-directed mutagenesis. *Nucl. Acids Res.* 20: 1147). The appropriate primer pairs (0.2 μM of each) were combined with 10ng of version "B" of reshaped human L25 V_L region cDNA, and 1 unit of AmpliTaq (Perkin Elmer Cetus) DNA polymerase in 50 μl of PCR buffer containing 10

5 mM Tris-HCl (pH 8.3), 50 mM KCl, 200 μ M dNTPs, and 1.5 mM $MgCl_2$. This was overlaid with mineral oil and PCR was performed for 25 cycles, each cycle consisting of a denaturation step at 94°C for 1 min, a primer annealing step at 55°C for 1 min, and an extension step at 72°C for 2 minutes. This was followed by a single cycle consisting of a further elongation step at 72°C for 10 minutes followed by cooling to 4°C. The ramp time
10 between the primer-annealing and extension steps was 2.5 minutes. The PCR products of the five reactions (A, B, C, D and E) were then purified by gel electrophoresis followed by DNA elution using Wizard PCR preps (Promega). PCR products A, B, C, D, and E were assembled by their complementarity to one another. In the second set of PCR reactions, PCR products B and C, and D and E, (50 ng of each) were added to 50
15 μ l PCR reactions (as described above) each containing 1 unit of AmpliTaq (Perkin Elmer Cetus) DNA polymerase. The reactions were cycled for 20 cycles as described above with the exception that the annealing temperature was raised to 60°C. In the third set of PCR reactions, PCR products F and G were PCR-amplified using 1 μ l of each prior PCR reaction and the appropriate pair of PCR primers (vla2-vla5 or vla6-APCR4).
20 The PCR reactions contained 1 unit of AmpliTaq DNA polymerase in 50 μ l PCR reaction (as described above) and were amplified for 25 cycles as in the first stage. In the fourth set of PCR reactions, the PCR product H was PCR-amplified using 1 μ l of each prior PCR reaction and the vla2-APCR4 pair of PCR primers. Finally, PCR products A and H were assembled by their own complementarity in a two step-PCR reaction similar to that
25 described above using RSP and UP as the terminal primers. The fully assembled fragment representing the entire reshaped human F19 V_L region including a leader sequence was digested with HindIII and BamHI and cloned into pUC19 for sequencing. A clone having the correct DNA sequence was designated reshF19La (Figure 29) and was then subcloned into the eukaryotic expression vector pKN100. The DNA sequence of
30 reshF19La cloned into pKN100 is shown in Figure 30.

The second version of reshaped human F19 V_L region (L_B) was constructed using the same scheme as that described for L_a but where vla4 and vla7 primers were substituted by vlb4 and vlb7 respectively (Table 3). The DNA sequence of L_B is shown in
35 Figure 29.

5 The third version of reshaped human F19 V_L region (L_C) was constructed using the QuikChangeTM site-directed mutagenesis kit from Stratagene. The QuikChange site-directed mutagenesis method was performed according to the manufacturer's instructions, using reshF19La in pKN100 vector as double stranded DNA template. The mutagenic oligonucleotide primers F19Lc-sense and F19Lc-antisense (Table 3) for use in
10 this protocol were designed according to the manufacturer's instructions. Briefly, both the mutagenic primers contained the desired point mutation (codon TTT at Kabat residue position 49 (Phe) changed to TAT coding for Tyr) and annealed to the same sequence on opposite strands of L_A in pKN100 vector. The point mutation was verified by
15 DNA sequencing the entire V_L region. The DNA sequence of L_C is shown in Figure 29. To eliminate the possibility that random mutations occurred in the pKN100 during the PCR reaction, the V_L region was cut out of the pKN100 vector as an HindIII/BamHI fragment and re-subcloned into an unmodified pKN100 vector cut with the same two restriction enzymes beforehand.

20
TABLE 3: PCR primers for the construction of reshaped human F19 light chain variable regions

25 1. Primers for the synthesis of version "A"

F19vla1 (36 mer):

5' GTCATCACAATGTCTCCGGAGGAACCTGGAACCCAG 3'

30 F19vla2 (29 mer):

5' CTCCGGAGACATTGTGATGACCCAATCTC 3'

F19vla3 (45 mer):

5' GAATATAAAAGGCTCTGACTGGACTTGCAGTTGATGGTGGCCCTC 3'

5 F19vla4 (72 mer):

5' CAGTCAGAGCCTTTTATATTCTAGAAATCAAAGAAGTACTTGGCCTGGTAT
CAGCAGAAACCAGGACAGCC 3'

F19vla5 (44 mer):

10 5' ACCCCAGATTCCCTAGTGCTAGCCCCAAAAGATGAGGAGTTTGGG 3'

F19vla6 (67 mer):

5' TAGCACTAGGGAATCTGGGGTACCTGATAGGTTCAAGTGGCAGTGGGTTTG
GGACAGACTTCACCCTC 3'

15

F19vla7 (53 mer):

5' GTCCCTTGTCCGAACGTGAGCGGATAGCTAAAATATTGCTGACAGTAA
TAAAC 3'

20 F19vla8 (33 mer):

5' GCTCACGTTCCGGACAAGGGACCAAGGTGGAAAT 3'

2. Primers for the synthesis of version "B"

25 F19vib4 (72 mer):

5' CAGTCAGAGCCTTTTATATTCTAGAAATCAAAGAAGTACTTGGCCTGG
TTCCAGCAGAAACCAGGACAGCC 3'

F19vib7 (57 mer):

30 5' GTCCCTTGTCCGAACGTGAGCGGATAGCTAAAATATTGCTGACAGTCATA
AACTGCC 3'

3. Primers for the synthesis of version "C"

35 F19Lc-sense (34 mer):

5 5' CCCAAACTCCTCATCTATTGGGCTAGCACTAGGG 3'

F19Lc-antisense (34 mer):

5' CCCTAGTGCTAGCCCAATAGATGAGGAGTTTGGG 3'

10 4. Primers hybridizing to the flanking PUC19 vector sequences

APCR1 (17 mer, sense primer): 5' TACGCAAACCGCCTCTC 3'

APCR4 (18 mer, anti-sense primer): 5' GAGTGCACCATATGCGGT 3'

RSP (-24) (16 mer, sense primer): 5' AACAGCTATGACCATG 3'

15 UP (-40) (17 mer, anti-sense primer): 5' GTTTTCCCAGTCACGAC 3'

Example 4: Construction of the reshaped human F19 heavy chain versions A to E ($H_A - H_E$)

20

Version "A" of reshaped human F19 V_H regions (H_A) was constructed using the same PCR methods as described for the construction of version "A" of reshaped human F19 V_L region (L_A) (Figure 31). The template DNA was version "A" of reshaped human 226 V_H (Léger O. J. P., Yednock T. A., Tanner L., Horner H. C., Hines D. K., Keen S., Saldanha J., Jones T., Fritz L. C. and Bendig M. M. (1997). Humanization of a mouse antibody against human alpha-4 integrin: a potential therapeutic for the treatment of multiple sclerosis. *Hum. Antibod.* 8: 3). Six PCR primers were designed and synthesized for the construction of version "A" of reshaped human F19 V_H region (Table 4). PCR products A, B, C, and D were obtained using APCR1-Vha1, Vha2-Vha3, Vha4-Vha5 and Vha6-APCR4 as PCR primer pairs, respectively. The PCR conditions were essentially as described for the construction of reshaped human F19 V_L region. A clone having the correct DNA sequence was designated reshF19Ha (Figure 32) and was then subcloned into the eukaryotic expression vector pG1D105. The DNA sequence of reshF19Ha cloned into pG1D105 is shown in Figure 33.

5 The third version of reshaped human F19 V_H region (H_C) was constructed using the same scheme as that described for H_A but where Vha4 primer was substituted by Vhc4 (Table 4). The DNA sequence of H_C is shown in Figure 32. The second (H_B) and fourth (H_D) version of reshaped human F19 V_H region were constructed based on the PCR-mutagenesis methods of Kamman et al. (Kamman M., Laufs J., Schell J. and Gronenborn B. (1989) Rapid insertional mutagenesis of DNA by polymerase chain re-
10 action (PCR). *Nucl. Acids Res.* 17: 5404). For H_B and H_D, a mutagenic primer F19VHbd6 (Tyr-91 to Phe-91, Table 4) was used paired with APCR4 in PCR reactions with H_A and H_C as the template DNA, respectively. The PCR products VHB and VHD were restriction enzyme digested with PstI and BamHI and subcloned into reshF19Ha
15 and reshF19Hc, respectively, previously digested with the same two restriction enzymes. The DNA sequences of H_B and H_D are shown in Figure 32.

Version "E" of reshaped human F19 V_H region (H_E) was constructed based on the PCR-mutagenesis methods of Kamman *et al.* (1989) already mentioned above:

20 For reshF19He mutagenic primer F19MscIHe (Table 5) was used paired with primer F19V_HHindIII (Table 5) in PCR reactions with H_C cloned in pg1d105 mammalian expression vector as the template DNA. The appropriate primer pairs (0.2 µM of each) were combined with 10 ng of cDNA of version "A" of reshaped human 226 V_H region in
25 100 µl of PCR buffer containing 10 mM KCl, 10 mM (NH₄)₂SO₄, 20 mM Tris-HCl (pH 8.8) 2 mM MgSO₄, 0.1% Triton X-100 and 200 µM dNTPs. Reaction mixtures were overlaid with mineral oil and kept at 94°C for 5 minutes. Then 1 unit of Deep Vent DNA polymerase (New England Biolabs) was added ("Hot Start" PCR; Chou Q., Russell M., Birch D., Raymond J. and Bloch W. (1992) Prevention of pre-PCR mis-priming and
30 primer dimerization improves low-copy-number amplifications. *Nucl. Acids Res.* 20: 1717) and PCR was performed for 25 cycles on a TRIO-Thermoblock Thermal Cycler (Biometra, Göttingen, Germany). Each cycle consisting of a denaturation step at 94°C for 1 min, a primer-annealing step at 70°C for 1 min, and an extension step at 72°C for 2 minutes. This was followed by a single cycle consisting of a further elongation step at

5 72°C for 10 minutes followed by cooling at 4°C. The PCR products were then extracted and purified from a TAE 1.4% standard agarose gel using a QIAquick™ gel extraction kit, following the protocol supplied by the manufacturer (QIAGEN Ltd., UK). The PCR product V_He was then restriction enzyme digested with MscI and HindIII and ligated into reshF19Hc cloned in pg1d105 previously digested with the same two restriction en-
10 zymes. The MscI restriction recognition site is unique to all the reshaped human F19 V_H region versions and is not present in the pg1d105 expression vector. The HindIII restriction recognition site is a unique site in pg1d105 for cloning of V_H immunoglobulin genes. Electroporation-competent XL-1 Blue E. coli cells were transformed with 1 µl of the ligated DNA and plated on agarose plates containing Ampicillin. Colonies were then
15 screened for the presence and correct size of inserts by direct PCR on colonies (Güssow D. and Clackson T. (1989) Direct clone characterization from plaques and colonies by the polymerase chain reaction. *Nucl. Acids Res.* 17: 4000) with primers HCMi and Hucy1 hybridising to the flanking pg1d105 vector sequences (Table 5). DNA from positive colonies was prepared using a Plasmid Midi kit, following the protocol supplied
20 by the manufacturer (QIAGEN Ltd., UK). DNA sequencing was performed by the dideoxy chain termination method (Sanger F., Nicklen S. and Coulson A. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* 74: 5463) directly from circular vector DNA using conventional heat denaturation (Andersen A., Pettersson A. and Kielsen T. (1992) A fast and simple technique for sequencing plasmid DNA with sequenase using heat denaturation. *Biotechniques* 13: 678) and Sequenase 2.0 (USB, Cleveland, OH). The DNA sequence of reshF19He is shown in Figure 32.

30 **TABLE 4: PCR primers for the construction of reshaped human F19 heavy chain variable regions versions A to D.**

1. Primers for the synthesis of version "A"

5 F19vha1 (47mer):

5' GTGTATTCAGTGAAGGTGTATCTACTAGTTTTACAGCTGACTTTCAC 3'

F19vha2 (53 mer):

5' TAGTAGATACACCTTCACTGAATACACCATACTGGGTTAGACAGG

10 CCCCTG 3'

F19vha3 (71 mer):

5' CCCTTGAACCTTCTGGTTGTAGTTAGGAATACCATTGTTAGGATTAATACC
TCCTATCCACTCCAGCCTTTG 3'

15

F19vha4 (71 mer):

5' TAACTACAACCAGAAGTTCAAGGGCCGGGCCACCTTGACCGTAGGCAA
GTCTGCCAGCACCGCCTACATGG 3'

20 F19vha5 (63 mer):

5' GCATGGCCCTCGTCGTAACCATAGGCGATTCTTCTTCTGGCGCAGTAGT
AGACTGCAGTGTCC 3'

F19vha6 (48 mer):

25 5' CTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAAGGAAC 3'

2. Primers for the synthesis of version "c"

F19vhc4 (71 mer):

30 5' TAACTACAACCAGAAGTTCAAGGGCCGGGTCACCATCACCGTAGACA
CCTCTGCCAGCACCGCCTACATGG 3'

3. Primers for the synthesis of version "B" and "D"

35 F19vhbd6 (27 mer):

5 5' GGACACTGCAGTCTACTTCTGCGCCAG 3'

4. Primers hybridizing to the flanking PUC19 vector sequences

APCR1 (17 mer, sense primer): 5' TACGCAAACCGCCTCTC 3'

10 APCR4 (18 mer, anti-sense primer): 5' GAGTGCACCATATGCGGT 3'

TABLE 5: PCR primer for the construction of reshaped human F19 heavy chain variable regions version E

15

1. Primer for the synthesis of version "E"

F19MscIHe (65 mer, anti-sense):

20 5' CCTTTGGCCAGGGGCCTGTCTAACCCAGTGTATGGTGTATTCAGTGAAGGTG
MscI
TATCCACTAGTTTCCACTAGTTT 3'

2. Primers hybridizing to the flanking pg1d105 mammalian expression vector sequences

25

HCMi (28 mer, sense): 5' GTCACCGTCCTTGACACGCGTCTCGGGA 3'

Hucy1 (17 mer, anti-sense): 5' TTGGAGGAGGGTGCCAG 3'

30 **Example 5: Reshaped human F19 antibody concentrations in COS cells supernatants**

COS cells were transfected with one pair of a series of reshaped human F19 antibody constructs and the human antibody concentration was measured using the

5 IgG1/Kappa ELISA as described in example 2.

TABLE 6: Reshaped human F19 antibody concentrations in COS cell supernatants

10

Transfected Antibody components		Human $\gamma 1/K$
Heavy chain	Kappa light chain	Concentration [$\mu\text{g/ml}$]
H _A	L _A	2.50
H _A	L _B	0.18
H _B	L _A	1.25
H _B	L _B	0.10
H _D	L _A	1.15
H _D	L _B	0.18
H _A	L _A	1.50
H _A	L _C	1.56
H _C	L _A	1.47
H _C	L _C	1.97
cF19	L _A	1.54
cF19	L _B	0.07
cF19	L _C	2.14

TABLE 7: Reshaped human F19 antibody concentrations in COS cell supernatants

15

Transfected Antibody components		Human $\gamma 1/K$
Heavy chain	Kappa light chain	concentration [$\mu\text{g/ml}$]
H _A	L _A	2.00
H _A	L _C	2.50
H _C	L _A	2.90
H _C	L _C	3.00
H _E	L _A	2.80
H _E	L _C	3.50

5 RNA splicing events required for the expression of immunoglobulin genes in mammalian cells

Both mammalian expression vectors pKN100 and pg1d105 have an intron between the variable and the constant regions which is removed during the process of
10 gene expression to give rise to an messenger RNA. The splicing event which consists of a DNA recombination between the heavy or light chain splice donor sites and the immunoglobulin splice acceptor site is described in Figure 34.

15 **Example 6: Flow cytometric analysis of the binding of cF19 and L_AH_C to FAP-expressing human cells**

The ability of L_AH_C to bind to both recombinant and endogenously expressed FAP on cell surface was tested.

20 The example was conducted to determine the binding of L_AH_C to cellular FAP. Both naturally FAP expressing MF-SH human tumour cells (Shirasuma, K, et al., Cancer 1985, 55:2521-32) and FAP-transfected human tumour cell lines were used as cellular targets. L_AH_C was studied in cytofluorometric assays evaluating direct binding to target
25 cells as well as by the inhibitory effect on the binding of either murine F19 or chimeric cF19 anti-FAP antibodies.

Antibodies and cell lines used were F19 (murine monoclonal anti-human FAP antibody, IgG1 subclass), mIgG (murine immunoglobulin, IgG class), cF19 (chimeric
30 monoclonal anti-human FAP antibody, IgG1 subclass), L_AH_C (reshaped monoclonal anti-human FAP antibody, IgG1 subclass), hIgG1 (human immunoglobulin, IgG1 subclass), MF-SH (human malignant fibrous histiocytoma cell line), HT-1080 (human fibrosarcoma cell line), HT-1080FAP clone 33 (HT-1080 cell line transfected with cDNA encoding human FAP). Antibodies were biotinylated as described in examples 8 and 12.

5 Direct binding of LAHC to FAP on the surface of human tumour cell lines

5x10⁵ cells of the tumour cell line under investigation were incubated with the indicated concentration of test or control antibody in a total volume of 0.2 ml phosphate-buffered saline (PBS) supplemented with 1% bovine serum albumin (BSA) for 30 min on
10 ice. Subsequently, cells were washed twice with 2 ml of PBS, resuspended in 0.2 ml of PBS supplemented with 1% BSA, a 1:20 dilution of mouse anti-human IgG FITC-labelled [Dianova] as secondary reagent was added) and incubated for another 30 min on ice.

15 Alternatively, 5x10⁵ cells of the tumour cell line under investigation were incubated with the indicated concentration of biotin-labelled cF19 in a total volume of 0.2 ml PBS supplemented with 1% BSA for 30 min on ice. Subsequently, cells were washed twice with 2 ml of PBS, resuspended in 0.2 ml of PBS supplemented with 1% BSA, and incubated for another 30 min on ice with 1:40 dilution of streptavidin-FITC (Dianova) as
20 secondary reagent.

Cells were again washed twice with 2 ml of PBS, resuspended in a total volume of 0.5 ml of PBS supplemented with 1% paraformaldehyde (PFA) and kept on ice. Single cell fluorescence was determined cytofluorometrically by analysing the cellular green
25 fluorescence at 488nm in an EPICS XL (Coulter) fluorescence-activated cell analyzer.

Competition of LAHC for binding of biotinylated cF19 to cell-surface FAP on FAP-expressing human cell lines

30 5x10⁵ cells of the tumour cell line under investigation were incubated with the indicated amounts of unlabelled test or control antibody added together with 1 µg/ml biotin-labelled cF19 antibody. Subsequently, cells were washed twice with 2 ml of PBS, resuspended in 0.2 ml of PBS supplemented with 1% BSA, 1:40 diluted streptavidin-FITC (Dianova) as secondary reagent and incubated for another 30 min on ice.

5 Cells were then washed twice with 2 ml of PBS, resuspended in a total volume of 0.5 ml PBS supplemented with 1% PFA and kept on ice. Single cell fluorescence was determined cytofluorometrically by analysing the cellular green fluorescence at 488nm in an EPICS XL (Coulter) fluorescence-activated cell analyzer.

10 Both, cF19 and LAHC bind in a concentration dependent manner specifically to FAP-transfected HT-1080FAP clone33 human tumour cells (Table 8). No binding to FAP-negative HT-1080 cells was detectable (Table 9). Both cF19 and LAHC bound in a concentration dependent manner to human MF-SH cells endogenously expressing FAP (Table 10).

15 Biotinylated cF19 bound to human HT-1080FAP clone 33 (Table 11) in a concentration dependent manner. No binding was detectable to FAP-negative HT-1080 cells (Table 12).

20 Binding of biotinylated cF19 to HT-1080FAP clone 33 cells was inhibited by both unlabelled cF19 and unlabelled LAHC (Table 13).

Chimeric anti-human FAP monoclonal antibody cF19 as well as reshaped human anti-human FAP monoclonal antibody LAHC (example 10) were shown to bind directly
25 to FAP expressed on human cell lines either endogenously expressing this protein or transfected with cDNA encoding for it. This binding was shown to be concentration dependent. Binding of biotinylated cF19 could be inhibited by both unlabelled cF19 and unlabelled LAHC.

30 Using cytofluorometric technology, direct binding as well as inhibition of specifically binding reagents showed specificity of chimeric cF19 and reshaped LAHC human monoclonal antibodies to cell surface expressed FAP.

5 **Table 8: Binding of anti-FAP antibodies to HT-1080FAP clone 33 cells**

Concentration of antibody	Mean fluorescence intensity		
[ng/ml]	hlgG1	cF19	LAHC
500.0	0.12	6.65	2.76
100.0	0.12	1.63	0.66
20.0	0.12	0.43	0.22
4.0	0.12	0.17	0.15
0.8	0.12	0.14	0.13

Table 9: Binding of anti-FAP antibodies to non-transfected HT-1080 cells

10

Concentration of antibody	Mean fluorescence intensity		
[ng/ml]	hlgG1	cF19	LAHC
500.0	0.11	0.11	0.12
100.0	0.11	0.11	0.11
20.0	0.11	0.11	0.12
4.0	0.11	0.11	0.12
0.8	0.11	0.11	0.11

Table 10: Binding of anti-FAP antibodies to MF-SH cells

Concentration of antibody	Mean fluorescence intensity		
[ng/ml]	hlgG1	cF19	LAHC
4000	0.6	3.6	2.8
2000	n.d.	3.3	2.5
1000	n.d.	2.4	1.9
500	n.d.	1.8	1.3

15

n.d.: not done

5 **Table 11: Binding of biotinylated cF19 antibody to HT-1080FAP clone 33 cells**

Concentration of antibody [ng/ml]	Mean fluorescence intensity	
	Biotinylated hlgG1	Biotinylated cF19
5,000.0	0.2	36.5
1,000.0	0.2	18.1
200.0	0.2	4.5
40.0	0.2	1.3
8.0	0.2	0.5
1.6	0.3	0.3

Table 12: Binding of biotinylated cF19 antibody to non-transfected HT-1080 cells

10

Concentration of antibody [ng/ml]	Mean fluorescence intensity	
	Biotinylated hlgG1	Biotinylated cF19
5,000.0	0.1	0.1
1,000.0	0.1	0.1
200.0	0.1	0.1
40.0	0.1	0.1
8.0	0.1	0.1
1.6	0.1	0.1

5 **Table 13: Competition of anti-FAP antibodies with the binding of biotinylated cF19 to HT-1080FAP clone 33 cells**

Competitor antibody	Concentration of competitor antibody [µg/ml]	Mean fluorescence concentration
No	0.00	11.2
hIgG1	1.00	9.0
hIgG1	3.16	11.3
hIgG1	10.00	9.8
hIgG1	31.66	10.3
cF19	1.00	7.5
cF19	3.16	4.8
cF19	10.00	1.3
cF19	31.66	1.2
LAHC	1.00	8.0
LAHC	3.16	5.5
LAHC	10.00	2.9
LAHC	31.66	1.7

Biotinylated cF19 was used at a concentration of 1 µg/ml in all tests shown in the Table 13.

10

Example 7: In vitro immune effector functions of monoclonal antibody LAHC

This experiment was conducted to determine the potential of the monoclonal antibody (mAb) LAHC with specificity for fibroblast activation antigen (FAP) to lyse FAP-expressing targets in the presence of human complement or human mononuclear leukocytes, respectively.

In particular, the ability of LAHC to mediate cytotoxic effects against HT-1080FAP clone 33 cells, which expressed human FAP on the surface, was studied. Cytotoxicity was determined in vitro using the following approach: ⁵¹Cr-labelled target cells were incubated in the presence of LAHC with human serum as source of complement or human MNC (peripheral blood mononuclear cells) as effectors. Release of ⁵¹Cr was measured as measure of target-cell lysis.

5 Antibodies and cell lines used were LAHC (reshaped human anti-human FAP IgG1 antibody), hIgG1 (human IgG1 isotype control), 3S193 (murine monoclonal anti-Lewis^Y IgG3 antibody), mIgG (murine IgG control), HT-1080 (human fibrosarcoma), HT-1080FAP clone 33, (HT1080 transfected with cDNA encoding human FAP), MCF-7
10 (human breast adenocarcinoma cell line).

Complement-mediated lysis of target cells by LAHC

15 Tumour cells were radiolabelled by incubation in RPMI1640 medium with 100 μ Ci ⁵¹Cr (NEN) at 37° C for one hour. Subsequently, cells were washed twice in ⁵¹Cr -free medium and resuspended at a concentration of 2×10^5 cells per ml.

Human serum as source of complement was freshly prepared from blood of different volunteers. Blood was taken by puncturing the arm vein, remained at room temperature for one hour to allow clotting to occur, and was kept at 4° C over night. Serum
20 was separated by centrifugation and taken off from the sediment.

The antibody under study was diluted from the stock solution to the appropriate concentration in RPMI1640 cell culture medium.

25 1×10^5 radiolabelled tumor cells of the indicated cell line were incubated for 2 h at 37° C in an incubator (95% air and 5% CO₂) in the presence of different concentrations of test or control antibody and 25% (v/v) human serum as the source of human complement. Incubations were performed in U-shaped 96-well plates in a total volume of
30 200 μ l RPMI1640 and done in triplicate. After the incubation period, plates were centrifuged, 100 μ l supernatant was removed and radioactivity was counted in a gamma-counter. The total amount of incorporated radioactivity was determined by measuring 10^4 target cells. Spontaneous release was defined as activity released from the target cells in the absence of both antibody and complement during the described incubation
35 period.

Specific lysis was calculated as follows:

$$\text{specific lysis (in \%)} = \frac{[\text{activity sample}] - [\text{activity spontaneous release}]}{[\text{maximum activity}] - [\text{activity spontaneous release}]} \times 100$$

Antibody-dependent cellular cytotoxicity (ADCC) of LAHC

Tumour cells were radiolabelled by incubation in RPMI1640 medium with 100 μCi ^{51}Cr at 37°C for one hour. Subsequently, cells were washed twice in ^{51}Cr -free medium and resuspended at a concentration of 2×10^5 cells per ml.

MNC (peripheral blood mononuclear cells) were prepared from peripheral blood taken by puncturing the arm vein of healthy human volunteers. Clotting was prevented by the addition of 20% citrate buffer. MNC from 4 ml of this blood preparation were purified by centrifugation (30 min at 400 G and room temperature) on 3 ml of lymphocyte preparation medium (Boehringer Mannheim, Germany). MNC (peripheral blood mononuclear cells) were taken off from the gradient, washed three times and diluted with RPMI1640 to the appropriate concentration. Lymphocyte activated killer (LAK) cells were derived from MNC (peripheral blood mononuclear cells) by incubation for 5 days at 37°C in an 95% air and 5% CO_2 incubator at an initial density of 1.3×10^6 cells per ml in the presence of 100 U recombinant human Interleukin-2 (IL-2). The antibody under study was diluted from the stock solution to the appropriate concentration in RPMI1640 cell culture medium.

1×10^4 radiolabelled tumour cells of the indicated cell line were incubated for 5 h at 37°C and 5% CO_2 in the presence of different concentrations of test or control antibody and MNC. MNC were added in amounts to reach the indicated effector:target cell ratio. Incubation was performed in U-shaped 96-well plates in a total volume of 200 μl RPMI1640 and done in duplicate.

After the incubation period, plates were centrifuged, 100 μ l of the supernatant were taken off and radioactivity was determined in a gamma-counter. The total amount of incorporated radioactivity was determined by measuring 10^4 target cells. Spontaneous release was defined as activity released from the target cells in the absence of both antibody and effector cells during the described incubation period.

Specific lysis was calculated as follows:

$$\text{specific lysis (in \%)} = \frac{[\text{activity sample}] - [\text{activity spontaneous release}]}{[\text{maximum activity}] - [\text{activity spontaneous release}]} \times 100$$

Antibody-mediated complement lysis of tumor cells

No L_AH_C -specific complement-mediated lysis (above that seen with an isotype control) was observed in HT-1080FAP clone 33 cells treated with L_AH_C at concentrations up to 50 μ g/ml (Table 14, Table 15a)

Lytic activity of human serum used as source of complement was shown by lysis of MCF-7 human breast carcinoma cells in the presence of 12.5 μ g/ml 3S193, a murine monoclonal anti-Lewis^Y antibody with known complement activating ability (Table 15b)

Antibody-mediated cellular lysis of tumor cells

In the presence of L_AH_C at concentrations up to 10 μ g/ml, no ADCC (antibody-dependent cellular toxicity) mediated by human MNC (Table 16) or human LAK cells (lymphokine activated killer cell, Table 17) of L_AH_C on HT-1080FAP clone 33 as measured by lysis was detectable above that seen with an isotype control at an effector:target ratio of 50:1.

5 In appropriate in vitro assays with either human complement or with human MNC - as effector mechanisms, human anti-FAP monoclonal antibody L_AH_C revealed no - detectable cytotoxic effects above isotype controls on FAP-expressing tumor cell line HT-1080FAP clone 33.

10

Table 14: Specific complement lysis (in %) of HT-1080FAP clone 33 tumor cell targets mediated by L_AH_C

Source of human serum:		HT-1080 clone 33:	
Concentration of antibody		hlgG1 isotype control	L _A H _C
A	50 µg/ml	5	4
A	10 µg/ml	5	3
B	50 µg/ml	7	5
B	10 µg/ml	6	5
	0 µg/ml	0	0

15 Incubation: 2 hours at 37°C, 25% serum from human volunteers A or B, respectively, as source of complement.

5 **Table 15a: Specific complement lysis (in %) of HT-1080FAP clone 33 tumor cell targets mediated by human anti-FAP monoclonal antibody LAHC**

Source of human serum:		HT1080clone 33:	
concentration of antibody		hlgG1	LAHC
A	10.00 µg/ml	2	1
A	2.50 µg/ml	2	2
A	0.60 µg/ml	1	1
A	0.15 µg/ml	1	2
A	0.00 µg/ml	2	2
B	10.00 µg/ml	2	2
B	2.50 µg/ml	2	2
B	0.60 µg/ml	2	2
B	0.15 µg/ml	2	2
B	0.00 µg/ml	2	2
C	10.00 µg/ml	2	2
C	2.50 µg/ml	1	1
C	0.60 µg/ml	1	1
C	0.15 µg/ml	2	1
C	0.00 µg/ml	3	3

10 Incubation: 2 hours at 37°C, 25% serum from human volunteers A, B or C, respectively, as source of complement.

5 **Table 15b: Specific complement lysis (in %) of MCF-7 tumour cell targets mediated by murine anti-Lewis^y monoclonal antibody 3S193**

Source of human serum:		MCF-7:	
concentration of antibody		mlgG	3S193
A	10.00 µg/ml	0	21
A	2.50 µg/ml	1	21
A	0.60 µg/ml	0	21
A	0.15 µg/ml	1	18
A	0.00 µg/ml	0	0
B	10.00 µg/ml	1	13
B	2.50 µg/ml	0	17
B	0.60 µg/ml	1	18
B	0.15 µg/ml	1	15
B	0.00 µg/ml	0	0
C	10.00 µg/ml	1	22
C	2.50 µg/ml	0	23
C	0.60 µg/ml	1	26
C	0.15 µg/ml	1	20
C	0.00 µg/ml	1	1

Incubation: 2 hours at 37° C, 25% serum from human volunteers A, B or C, as source of complement.

10

Table 16: ADCC (antibody-dependant cellular cytotoxicity) (specific lysis in %) of HT-1080FAP clone 33 target cells by human MNC (peripheral blood mononuclear cells) mediated by LAHC.

15

HT-1080FAP clone 33:

Concentration of antibody: [in µg/ml]	HT-1080FAP clone 33:	
	hlgG1	LAHC
10.000	2	2
2.500	2	2
0.625	2	2
0.156	3	3
0.000	3	3

Incubation: 5 hours at 37°C, 10⁴ target cells and an effector:target cell ration of 50:1.

5 **Table 17: ADCC (antibody-dependent cellular cytotoxicity, specific lysis in %) of HT-1080FAP clone 33 target cells by LAK cells (lymphokine activated killer cells) mediated by L_AH_C.**

Concentration of antibody:	HT-1080FAP clone 33:	
[in µg/ml]	hlgG1	L _A H _C
10.000	12	14
2.500	14	17
0.625	14	21
0.156	15	21
0.000	14	14

Incubation: 5 hours at 37°C, 10⁴ target cells and an effector:target cell ration of 50:1.

10

Example 8: Immunohistochemical analysis of monoclonal antibody L_AH_C binding to normal and neoplastic human tissues

15 This experiment was performed to determine the binding characteristics of the humanized mAb L_AH_C to normal and neoplastic human tissues.

The following antibodies were used: L_AH_C, cF19, and the negative control hlgG1 were directly biotinylated according to methods of the state of the art and used at concentrations of 2.5 to 0.25 mg/ml in 2% BSA/PBS (bovine serum albumin in phosphate-buffered saline). Murine mAb F19 was used as tissue culture supernatant of the F19 hybridoma, at dilutions of 1:5 to 1:10 in 2% BSA/PBS.

25 The following reagents were used for immunochemical assays: Streptavidin peroxidase complex (Vector Labs, Burlingame, CA, USA), Avidin-biotin peroxidase complex (Vector Labs.), Biotinylated horse anti-mouse (Vector Labs.), DAB (diaminobenzidine, Sigma Chemical Co. St. Louis, MO, USA), Harris' hematoxylin.

30 Fresh frozen tissue samples examined included the following: Normal colon, breast, lung, stomach, pancreas, skin, larynx, urinary bladder, smooth and skeletal

5 muscle. Among the tumors tested were carcinomas from breast, colon, lung, esophagus, uterus, ovary, pancreas, stomach, and head and neck.

An indirect immunoperoxidase method was carried out according to state of the art methods (Garin-Chesa P, Old LJ, Rettig WJ: Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. *Proc Natl Acad Sci USA* (1990) 87:7235-7239) on five micrometer thickness fresh frozen sections. DAB was used as a substrate for the final reaction product. The sections were counter-stained with Harris' hematoxylin and examined for antigen expression.

15 L_AH_C expression in normal human tissues

The normal tissues tested were negative for L_AH_C expression, except for the normal pancreas in which a subset of positive endocrine cells in the islets of Langerhans (A cells) were identified with L_AH_C, cF19 and F19. (Table 18). No immunoreactivity was observed with the hIgG1 (human immunoglobulin IgG1 subclass) used as a negative control.

L_AH_C expression in tumors

In the tumor samples, L_AH_C, cF19 and F19 showed an indistinguishable pattern of expression in the tumor stromal fibroblasts. A strong and homogeneous expression was found in the majority of the cases examined, especially in the cancer samples derived from breast, colon, lung, pancreas and in the squamous cell carcinomas (SQCC) of the head and neck tested (Table 19). No immunoreactivity was observed with the hIgG1 used as negative control.

30 L_AH_C, cF19 and F19 showed immunoreactivity with the tumor stromal fibroblasts in the epithelial cancer samples tested. No L_AH_C or F19 immunoreactivity was seen with either the fibrocytes of the normal organ mesenchyme or the parenchymal cells of normal adult organs. Anti-FAP immunoreactivity was only observed in a subset of endocrine cells in the pancreatic islets, presumably glucagon-producing A cells, and in

- 5 four of nine uterine samples tested, representing subsets of stromal fibroblasts in these tissues.

Immunohistochemical analysis of L_AH_C in normal human tissues and FAP-expressing human carcinomas showed indistinguishable patterns of binding for L_AH_C,
 10 cF19 and murine mAb F19.

Table 18: Immunoreactivity of mAbs L_AH_C, cF19 and F19 with normal human tissues

Tissue type	No.	BIBH 1	cF19	F19
Breast	4			
-Epithelial cells ducts / acini		-	-	-
-Myoepithelial cells		-	-	-
-Connective tissue		-	-	-
-Blood vessels		-	-	-
Colon	6			
-Crypts of Lieberkühn		-	-	-
-Connective tissue		-	-	-
-Lymphoid tissue		-	-	-
-Smooth muscle		-	-	-
-Blood vessels		-	-	-
-Myenteric plexus		-	-	-
Lung	4			
Bronchus:				
-Bronchial epithelium		-	-	-
-Hyaline cartilage		-	-	-
-Connective tissue		-	-	-
-Mucous glands		-	-	-
Alveolus:				
-Pneumocytes (type I/II)		-	-	-
-Alveolar phagocytes		-	-	-
-Alveolar capillaries		-	-	-

5

Stomach	3			
-Surface epithelium		-	-	-
-Gastric glands:		-	-	-
-Chief cells		-	-	-
-Parietal (oxyntic) cells		-	-	-
-Mucous cells		-	-	-
-Neuroendocrine cells		-	-	-
-Connective tissue		-	-	-
-Blood vessels				
-Smooth muscle				
Esophagus	1			
-Surface epithelium		-	-	-
-Connective tissue		-	-	-
Small intestine	1			
-Epithelium of villi&crypts		-	-	-
-Connective tissue		-	-	-
-Smooth muscle		-	-	-
-Blood vessels		-	-	-
-Lymphoid tissue		-	-	-
Urinary bladder	2			
-Urothelium		-	-	-
-Connective tissue		-	-	-
-Smooth muscle		-	-	-
-Blood vessels		-	-	-
Pancreas	3			
-Duct epithelium		-	-	-
-Acinar epithelium		-	-	-
-Islets of Langerhans:		-	-	-
-B-cells		-	-	-
-A-cells		++	++	++
-D-cells		-	-	-
-Connective tissue		-	-	-
-Blood vessels		-	-	-
-Nerves		-	-	-
Larynx	1			
-Squamous epithelium		-	-	-
-Mucous glands		-	-	-
-Connective tissue		-	-	-
-Hyaline cartilage		-	-	-
-Blood vessels		-	-	-
-Skeletal muscle		-	-	-
Lymph node	1			

-Lymphoid cells		-	-	-
-Lymph sinuses		-	-	-
-Connective tissue		-	-	-
-Blood vessels		-	-	-
Spleen	1			
-Red & white pulp		-	-	-
-Sinuses		-	-	-
-Connective tissue		-	-	-
Liver	1			
-Hepatocytes		-	-	-
-Bile ducts		-	-	-
-Portal triad		-	-	-
Thyroid gland	2			
-Follicular epithelium		-	-	-
-Parafollicular cells		-	-	-
-Connective tissue		-	-	-
Prostate gland	1			
-Glandular epithelium		-	-	-
-Stroma		-	-	-
Testicle	1			
-Seminiferous tubules		-	-	-
-Stroma		-	-	-
Ovary	3			
-Follicles		-	-	-
-Stroma		-	-	-
Uterine cervix	1			
-Epithelium		-	-	-
-Stroma		-	-	-
Uterus	9			
-Endometrium:				
glands		-	-	-
stroma		+#	+#	+#
blood vessels		-	-	-
-Myometrium		-	-	-
Cerebral cortex	1			
-Neurons		-	-	-
-Neuroglial cells		-	-	-
-Blood vessels		-	-	-
Cerebellum	1			
-Molecular layer		-	-	-
-Granular cell layer		-	-	-
-Purkinje cells		-	-	-
-Blood vessels		-	-	-
Skin	3			
-Squamous epithelium		-	-	-
-Melanocytes		-	-	-

-Skin appendages		-	-	-
-Connective tissue		-	-	-
-Blood vessels		-	-	-

Acetone-fixed frozen sections were tested by the avidin-biotin complex immunoperoxidase procedure.

No, Number of tissue samples derived from different individuals tested.

* Identification of A-cells, based on morphology and location within the islets.

Positive immunoreactivity in the stroma in 4 out of 9 samples tested. The positive samples represent early (x2) and intermediate (x2) phase proliferative endometrium.

Example 9: Species specificity of LAHC binding in tissue sections

This experiment was conducted to assess the reactivity of LAHC with tissues from mouse, rat, rabbit and cynomolgus monkeys by immunohistochemical methods.

Also used in these tests were cF19 and human IgG1 (hlgG1) as negative controls. The reagents used for immunohistochemistry were Streptavidin peroxidase complex (Vector Labs., Burlingame, CA, USA), DAB (Sigma Chemical Co., St. Louis, MO, USA) and Harris' hematoxylin.

The following fresh frozen tissue samples from mouse, rat, rabbit and cynomolgus were tested: Brain, liver, lung, kidney, stomach, pancreas, intestine, thymus, skin, muscle, heart, spleen, ovary, uterus and testes. As positive control, sections from normal human pancreas and a breast carcinoma sample were included in every assay.

Immunohistochemistry

An indirect immunoperoxidase method was carried out as described in the state of the art (Garin-Chesa P, Old LJ, Rettig WJ: Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers. *Proc Natl Acad Sci USA* (1990) 87:7235-7239) on five micrometer thickness fresh frozen sections. The antibodies LAHC, cF19 and hlgG1 (at 1 µg/ml) were biotinylated according to the state of the art and were detected with streptavidin peroxidase complex. DAB was used as a

5 substrate for the final reaction product. The sections were counterstained with Harris' hematoxylin and examined for antigen expression.

The normal tissues tested did not react with either LAHC or cF19 in the experiments (Table 1).

10

The normal human pancreas used as positive control showed LAHC and cF19 binding in a subset of endocrine cells in the islets of Langerhans as previously described for F19. In addition, binding of LAHC and cF19 was seen in the tumor stromal fibroblasts in the breast carcinoma sample.

15

Immunohistochemical analysis of normal tissues from mouse, rat, rabbit and cynomolgus failed to detect any binding of either LAHC or cF19, in the experiments performed.

20

5 **Table 20: Binding of LAHC to tissue sections of non-human species, as determined by immunohistochemistry.**

Organ / Tissue type	Mouse	Rat	Rabbit	Cynomolgus
Brain –Cerebral cortex	-	-	-	
-Cerebellum	-	-	-	-
Liver –Hepatocytes	-	-	-	-
-Portal triad	-	-	-	-
Lung –Bronchi	-	-	-	-
-Alveoli	-	-	-	-
Kidney –Glomeruli	-	-	-	-
-Tubular epithelium	-	-	-	-
Stomach –Glandular epithelium				
-Smooth muscle	-	-	-	-
Pancreas –Exocrine acini	-	-	-	-
-Endocrine islets	-	-	-	-
Intestine –Glandular epithelium	-	-	-	-
-Smooth muscle	-	-	-	-
Thymus –Lymphocytes	-	-	-	-
Skin –Keratinocytes	-	-	-	-
-Sweat glands	-	-	-	-
-Hair follicles	-	-	-	-
Skeletal muscle	-	-	-	-
Heart	-	-	-	-
Spleen –Lymphocytes	-	-	-	-
Ovary –Follicular epithelium	-	-	-	-
-Stroma	-	-	-	-
Uterus –Myometrium	-	-	-	-
-Cervix uteri	-	-	-	-
Testis –Tubular epithelium	nt	nt	nt	-
Connective tissue	-	-	-	-

nt, not tested

5 **Example 10: Construction of cell lines producing chimeric and reshaped anti-FAP monoclonal antibodies**

The objective of this experiment was to demonstrate stable cell lines according to the invention expressing L_AH_C, L_AH_A, L_BH_B, L_BH_D, and cF19 in CHO DG44 cells.

10 Stable cell lines transfected with humanized or chimeric F19 antibodies were produced and their identity was confirmed by PCR amplification of heavy and light variable regions using genomic DNA derived from each transfectant as template.

CHO DG44 cells maintained under serum-free conditions in SFM-II medium.

15 Lipofectin and SFM-II serum-free medium were obtained from Gibco/BRL. Geneticin and all restriction enzymes were obtained from Boehringer Mannheim. Pfu polymerase was obtained from Stratagene.

DNA for transfections was purified from E. coli cells using QiaFilter Maxi Cartridges (Qiagen) as directed by the manufacturer. All DNA preparations were examined by
20 restriction enzyme digestion. Sequences of L_AH_C variable regions in their respective vectors were confirmed using an ABI PRISM 310 Sequencer (Perkin-Elmer).

Further information regarding the vectors and DNA sequences employed is
25 available in the prior examples.

Transfection of CHO DG44 cells

Cells in logarithmic growth were plated into 6 well plates containing 1 ml fresh
30 SFM-II medium. Plasmids encoding heavy and light chains of humanized or chimeric F19 versions were cotransfected into CHO DG44 cells using liposomal transfection. Liposomes were prepared using 6 µl Lipofectin reagent and 0.5 µg of each vector (one for the desired heavy chain and one for the light) as described for LipofectAMINE transfections except that SFM-II medium was used to dilute all reagents. Twenty-four
35 hours later, cells were diluted 1:10 into SFM-II medium containing 300 µg/ml Geneticin.

5 After the initial phase of cell killing was over (10-14 days), the concentration of Geneticin was reduced to 200 mg/ml and methotrexate was added to a final concentration of 5 nM. Methotrexate concentrations were increased after 10-14 days to a final concentration of 20 nM.

10 PCR Amplification of transfectant DNA

10⁷ CHO DG44 cells were centrifuged in an Eppendorf microcentrifuge briefly at full speed, washed once with PBS, and pelleted once again. Genomic DNA was prepared by ethanol precipitation after SDS lysis and Proteinase K treatment of the cell
15 pellets.

A mixture containing one of the following primer pairs, dNTPs, buffer, and Pfu polymerase was used to amplify either the heavy or light chain variable region using genomic DNA as template. The resulting PCR products were digested with the appropriate restriction enzyme and analyzed by agarose gel electrophoresis to confirm their
20 identity.

Light chain primer set:

25 5'- GAG ACA TTG TGA CCC AAT CTC C - 3' PKN 1690

5'- GAC AGT CAT AAA CTG CCA CAT CTT C - 3' PKN.1930.R

Heavy chain primer set:

30 5'- TTG ACA CGC GTC TCG GGA AGC TT - 3' PG 5863

5'- GGC GCA GAG GAT CCA CTC ACC T - 3' PG 6332.R

5 The undigested heavy chain PCR product has a predicted size of 469 bp while the light chain PCR product has a predicted size of 286 bp. Verification of identity was determined by restriction enzyme digest with BstEII (heavy chain) or NlaIV (light chain).

CHO cell lines were transfected with LAHC, LAHA, LBHB, LBHD, as well as cF19. Geneticin-resistant cells were obtained and these cells were further selected for resistance to methotrexate. PCR amplification, followed by restriction enzyme cleavage of the light and heavy chain DNA produced the expected bands and confirmed the identity of LAHC, LBHB, LAHA and LBHD transfectants..

15 The cells described were maintained under serum-free conditions at all times and were not treated with animal-derived products such as trypsin.

Producer cell lines transfected with expressing monoclonal LAHC, LAHA, LBHB, LBHD and cF19 antibodies were produced. Their identities were confirmed using PCR amplification and restriction enzyme cleavage of the resulting PCR products of both their heavy and light chain variable regions.

Example 11: Expression of antibody proteins in Chinese hamster ovary DG44 cells and their purification

25 The objective of this experiment was to express and purify LAHC, LAHA, LBHB, and LBHD mAbs to enable their characterization. Other goals included the establishment of a quantitative ELISA to permit measurement of antibody concentrations in both crude media samples as well as purified Ig samples and determination of relative expression levels of various humanized F19 constructs using this assay.

Serum-free CHO DG44 cells and USP-grade methotrexate were obtained from the Biotechnical Production Unit of the Dr. Karl Thomae GmbH, Biberach, Germany; both products are also commercially available. Cells were maintained under serum-free con-

5 ditions at all times. SFM-II serum-free medium was obtained from Gibco/BRL. Protein A agarose was from Pierce Chemical (Indianapolis, IN, USA). Human IgG1 standards (Cat. No. I 3889), p-Nitrophenyl phosphate tablets (N 2640), bovine serum albumin (BSA) (A 7906), and goat anti-human kappa chain specific alkaline phosphatase-conjugated antibody (A 3813) were obtained from Sigma Chemical (St. Louis, MO, USA).

10 Goat anti-human gamma-chain specific alkaline phosphatase-conjugated antibody was obtained from Jackson ImmunoResearch Laboratories (through Stratech Scientific). Tris-buffered saline (TBS) consisted of 150 mM NaCl, 50 mM Tris, pH 7.5.

Cell culture conditions for antibody expression

15 Cells were cultured and maintained in T-175 flasks in SFM-II serum-free medium without agitation. The medium contained 200 µg/ml Geneticin and 20 nM methotrexate without antibiotics. Cells were passaged by dilution, were not adherent, and grew in small clusters. When the cells reached stationary phase, the medium was collected and
20 centrifuged to remove cells and frozen at -20°C until needed.

Purification of LAHC

All purification steps were carried out at 4°C. A C10/10 column (Pharmacia Fine
25 Chemicals) was packed with Protein A agarose (3 ml bed volume). The column was washed with TBS and preeluted once with 0.1 M Na citrate, pH 3.0 to insure that no loosely bound material remained on the column. The column was then immediately re-equilibrated with TBS and stored at 4°C. Spent culture supernatants were thawed and centrifuged at 10,000 x g for 30 minutes prior to Protein A chromatography to remove
30 debris and diluted with an equal volume of TBS. This material was loaded onto the Protein A column at 0.5 ml/min using a P-1 peristaltic pump (Pharmacia) and washed with TBS until the absorbance at 280 nm was undetectable. Elution of the antibody was initiated with 0.1 M Na citrate pH 3.0 at approximately 0.2 ml/min. The elution was monitored at 280 nm and one ml fractions of the eluted material were collected into
35 tubes containing sufficient Tris base pH 9 to neutralize the citrate buffer. Protein-con-

5 taining fractions were pooled and concentrated using an Amicon filtration apparatus with a YM-30 filter and dialyzed against PBS. The column was immediately regenerated with TBS. Protein dye-binding assays were performed with the BioRad (Hercules, California) protein determination kit, according to the manufacturer's instructions, using bovine serum albumin as a standard.

10 Human IgG (gamma immunoglobulin) ELISA

ELISA plates were coated overnight with 100 μ l goat anti-human gamma-chain specific alkaline phosphatase-conjugated antibody at 0.4 mg/ml in coating buffer at 4°C. 15 Coating antibody was removed and plates were blocked with 2% BSA in PBS for 2 hours. All subsequent steps were performed at 37°C. Blocking buffer was replaced with antibody samples or human IgG1 standard diluted in dilution buffer, serially diluted in a 200 μ l volume, and incubated for one hour. Negative controls included dilution buffer and/or culture medium of nontransfected cells. Wells were washed and 100 μ l goat anti- 20 human kappa chain specific alkaline phosphatase-conjugated antibody diluted 1:5000 was added and incubated for one hour. Wells were washed and 100 μ l reaction buffer was added and incubated for 30 minutes. The reaction was stopped by addition of 1 M NaOH and absorbance read at 405 nm in an ELISA plate reader. Results were analyzed by four-parameter iterative curve fitting.

25 Amino acid analysis was performed according to methods available in the state of the art.

Monoclonal antibody LAHC was produced and purified to homogeneity using Protein A affinity chromatography. ELISA assays using human IgG1 as standard indicated 30 LAHC recoveries exceeding 70%. The purity of the material was estimated to be >90% by SDS-polyacrylamide gel electrophoresis. Representative expression data and typical purification yields are shown in Table 21.

Table 21: Expression data and purification yields FAP antibody proteins in CHO cells

Antibody	Expression levels in crude media samples (ELISA)	Purified antibody yields	Yield improvement [purified antibody]
LAHC	7 - 10 mg/l	~ 5 - 7 mg/l	500 - 700
LAHA	5 - 7 mg/l	~ 3 - 4 mg/l	300 - 400
LBHB	0.5 - 1 mg/l	~ 0.2 - 0.5 mg/l	20 - 50
LBHD	0.8 - 1.5 mg/l	~ 0.3 - 0.8 mg/l	30 - 60
Chimeric F19	~ 0.02 mg/l	< 0.01 mg/l	1

Representative expression data for each of the anti-FAP antibodies produced in this study are shown. Recoveries after Protein A agarose affinity chromatography were based on protein dye-binding measurements of the purified Ig using BSA as a standard.

Example 12: Binding of monoclonal antibody LAHC to isolated recombinant human FAP

The objective of this study was to characterize binding of LAHC to isolated recombinant human FAP.

CD8-FAP ELISA

ELISA plates were coated overnight with 100 µl of mouse anti-rat antibody (Sigma Chemical R0761) at 1:2000 in coating buffer at 4°C. Coating antibody was removed and plates were blocked with 2% BSA in PBS for one hour. All subsequent steps were performed at room temperature. Blocking buffer was replaced with 100 µl of 1 µg/ml rat anti-CD8 antibody (Pharmingen 01041D) and incubated for one hour. Plates were washed and 100 µL CD8-FAP culture supernatant (see Example 14) (1:2 in PBS) was added and allowed to bind for one hour. Plates were washed and antibody samples were added (two-fold serial dilutions) in a 100 µl volume and incubated for one hour. Negative controls included human IgG and/or culture medium of nontransfected cells. Wells were washed and 100 µl of horseradish peroxidase (HRP) conjugated mouse anti-human IgG1 antibody (Zymed 05-3320) diluted 1:500 in dilution buffer were added

5 and incubated for one hour. Wells were washed and 100 μ l HRP substrate, (azino-bis (3-ethylbenzthiazoline 6-sulfonic) acid, Sigma Chemical A9941), were added and incubated for 60 minutes. The reaction was stopped by addition of 1 M NaOH and absorbance read at 405/490 nm in an ELISA plate reader. Results were analyzed by four-parameter curve iterative curve fitting.

10 Alternatively, plates were coated directly with cF19. FAP (recombinant human FAP, see example 13) was allowed to bind to these plates as above and biotinylated LAHC (~1 μ g/ml) was then added. Antibody binding was detected with HRP-streptavidin conjugate as above.

15 Solubilization of membrane-bound human FAP

FAP-expressing 293FAP 1/2 cells or control 293 cells were washed with PBS and lysed with 1% Triton X-114 in Tris-buffered saline. Nuclei and debris were removed by
20 centrifugation at 10,000 x g. The supernatant was phase-partitioned (Estreicher A, Wohlend A, Belin D, Scheuning WD Vasalli JD. Characterization of the cellular binding site for the urokinase-type plasminogen activator. (1989) *J Biol Chem* **264**:1180-1189) to enrich membrane proteins. The detergent phase was collected and diluted in buffer containing 1% Empigen BB (Calbiochem) to prevent reaggregation of the Triton X-114.
25 This material was subjected to Concanavalin A agarose chromatography (Rettig WJ, Garin-Chesa P, Healey JH, Su SL, Ozer HL, Schwab, M, Albino AP, Old LJ. Regulation and heteromeric structure of the fibroblast activation protein in normal and transformed cells of mesenchymal and neuroectodermal origin. (1993) *Cancer Res* **53**:3327-3335).

30 Biotinylation of LAHC

LAHC (1-2 mg) was dialyzed against 50 mM bicarbonate buffer and biotinylated with a ten-fold molar excess of sulfosuccinimidyl-6-biotinamido hexanoate (NHS-LC

5 biotin, Pierce Chemical, Rockford, Illinois, USA) for 2 hours at room temperature. Unreacted product was removed by repeated microdialysis in a microconcentrator.

Transient transfections

10 COS-7 cells (American Type Tissue Culture Collection, reference number CRL 1651) were cotransfected by electroporation with the heavy and light chain vectors encoding LAHC.

Anti-CD8 monoclonal antibody was immobilized onto microtiter plates. CD8-FAP
15 from medium of insect cells infected with CD8-FAP baculovirus was allowed to bind to these plates. Spent medium from COS-7 cell cultures transiently transfected with two separate vectors encoding LAHC was serially diluted and added to the wells containing the immobilized CD8-FAP. LAHC bound to isolated immobilized CD8-FAP protein (Figure 35). Culture supernatants from mock-transfected COS-7 cells failed to demonstrate
20 binding.

Recombinant membrane-bound FAP from detergent extracts of 293FAP 1/2 cells or control extracts was serially diluted and immobilized via chimeric F19 monoclonal antibody bound to microtiter plates. Biotinylated LAHC bound recombinant human FAP
25 immobilized with cF19 (Figure 36) in a concentration-dependent manner.

LAHC recognized isolated immobilized recombinant human FAP carrying the epitope for murine F19. LAHC bound to both CD8-FAP produced in insect cells, as well as FAP protein produced in 293FAP 1/2 cells.

30 Culture supernatants from COS7 cells transfected with either heavy and light chain vectors encoding LAHC or without DNA (Control) were collected three days posttransfection. CD8-FAP was immobilized via an anti-CD8 antibody as described in the text. Serial dilutions of the COS7 supernatants were allowed to bind to the immobilized CD8-FAP and subsequently detected with an HRP-conjugated anti-human IgG1 antibody.
35

5 Detergent extracts of FAP-expressing 293FAP 1/2 cells or control 293 cells were serially diluted and added to cF19-coated microtiter plates. Biotinylated LAHC was added and binding of biotinylated LAHC was detected with HRP-conjugated strept-
10 avidin.

**Example 13: Characterization of HT-1080 fibrosarcoma cells and 293 human em-
bryonic kidney cells transfected with cDNA for human FAP**

15 Fibroblast activation protein (FAP) is a cell-surface, membrane-bound protein which carries the F19 epitope and is expressed on tumor stromal fibroblasts. Cell lines expressing recombinant FAP protein and matched controls lacking FAP were generated for the characterization of anti-FAP monoclonal antibodies.

20 Cells used were HT-1080 cells (reference number CCL 121) and 293 human embryonic kidney cells (reference number CRL 1573) were obtained from the American Type Culture Collection (Maryland, USA). Transfectam was obtained from Promega (Madison, WI). Geneticin and all restriction enzymes were obtained from Boehringer Mannheim. DNA for transfections was purified from E. coli cells using QiaFilter Maxi
25 Cartridges (Qiagen) as directed by the manufacturer. All DNA preparations were examined by restriction enzyme digestion. Vector sequences were confirmed using an ABI PRISM 310 Sequencer.

30 Further information regarding the vectors and DNA sequences employed has been described in Scanlan MJ, Raj BK, Calvo B, Garin-Chesa P, Sanz-Moncasi MP, Healey JH, Old LJ, Rettig WJ. Molecular cloning of fibroblast activation protein alpha, a member of the serine protease family selectively expressed in stromal fibroblasts of epithelial cancers. (1992) *Proc Natl Acad Sci USA* 89:10832-10836. The FAP cDNA sequence has been deposited in Genbank (accession number HS09287).

5 Cell culture and immunoassays

HT-1080 cells were transfected with 1 mg DNA using Transfectam according to the manufacturer's instructions. Human embryonic kidney 293 cells were transfected by calcium phosphate transfection (Brann MR; Buckley NJ; Jones SVP; Bonner TI. Expression of cloned muscarinic receptor in A9 L cells. (1987) *Mol Pharmacol* 32: 450-455) with 10 mg DNA. Twenty-four hours later, cells were diluted 1:10 into fresh medium containing 200 mg/ml Geneticin. Colonies were picked and examined by immunofluorescence for FAP expression as described in Rettig WJ; Garin-Chesa P; Beresford HR; Oettgen HF; Melamed MR; Old LJ. Cell-surface glycoproteins of human sarcomas: differential expression in normal and malignant tissues and cultured cells. (1988) *Proc Natl Acad Sci USA* 85: 3110-3114.

Immunoprecipitations with cF19 were performed with metabolically labelled cells as described in Rettig WJ, Garin-Chesa P, Healey JH, Su SL, Ozer HL, Schwab, M, Albino AP, Old LJ. Regulation and heteromeric structure of the fibroblast activation protein in normal and transformed cells of mesenchymal and neuroectodermal origin. (1993) *Cancer Res* 53:3327-3335.

HT-1080 and 293 cells were tested for FAP antigen expression in immunofluorescence assays with anti-FAP antibodies and were found to be antigen-negative. Transfection of these cells with FAP.38 vector resulted in the generation of Geneticin-resistant colonies. Isolated colonies were picked and analyzed by immunofluorescence for FAP expression. Two cell clones were identified, designated HT-1080FAP clone 33 and 293FAP I/2, which express cell surface-bound FAP protein, as recognized by cF19 antibody. Staining of nonpermeabilized HT-1080FAP clone 33 cells and 293FAP I/2 with cF19 antibody confirmed the cell surface localization of the FAP protein.

Immunoprecipitation of radiolabelled FAP protein with cF19 from extracts of ³⁵S-methionine labelled HT-1080FAP clone 33 cells or 293FAP I/2 cells resulted in the ap-

pearance of a 93 kilodalton band after autoradiography. This band is not detectable in immunoprecipitates of parental HT-1080 or 293 cell extracts.

Two stably transfected cell lines, HT-1080FAP clone 33 and 293FAP 1/2, express FAP on the cell surface as determined in immunological assays with anti-FAP mAbs.

Neither parental HT-1080 cells nor parental 293 cells express detectable levels of FAP.

Example 14: Generation and characterization of CD8-FAP fusion protein

A soluble form of human FAP (fibroblast activation protein) in the form of a CD8-FAP fusion protein was produced in insect cells for the characterization of LAHC containing the binding site for anti-FAP mAbs. Murine CD8 was chosen to permit secretion of the protein and to provide an additional epitope tag.

The cDNA encoding the extracellular domain of CD8, consisting of the first 189 amino acids of murine CD8 α (Genbank M12825), was linked to that of the extracellular domain of FAP (amino acids 27 to 760), essentially as described by Lane, et al. (Lane P, Brocker T, Hubele S, Padovan E, Lazavecchia A, McConnell. Soluble CD40 ligand can replace the normal T cell-derived CD40 ligand signal to B cells in T cell-dependent activation. (1993) *J Exp Med* 177:1209-1213) using standard PCR protocols. The authenticity of all clones was verified by DNA sequencing. The resulting DNA was inserted into the pVL1393 vector (Invitrogen) and transfection of Sf9 cells (Invitrogen) with this vector and amplification of the resulting recombinant baculovirus were performed as described (Baculovirus Expression Vectors. A Laboratory Manual. O'Reilly DR, Miller LK, Luckow VA, (Eds.), Oxford University Press: New York, 1994). The spent medium of High Five™ cells (Invitrogen) infected with recombinant CD8-FAP baculovirus for four days was collected and cleared by ultracentrifugation.

The CD8-FAP ELISA (enzyme-linked immunosorbent assay) has been described above (Example 12).

5

Insect cell cultures infected with CD8-FAP virus secreted a fusion protein into the medium which carries the F19 epitope and is recognized by an anti-FAP antibody (Figure 37). Neither the cell culture medium alone nor medium from insect cells infected with CD8-CD40L fusion protein bound anti-FAP antibody.

10

Soluble CD8-FAP protein carrying the F19 epitope was secreted into the medium of infected insect cell cultures. Culture supernatant from cells infected with a control construct did not contain antigen bearing the F19 epitope.

15

A soluble form of FAP, CD8-FAP, was produced in insect cells and CD8-FAP was shown to carry the epitope recognized by cF19.

20
25

Supernatants from insect cells infected with recombinant baculovirus encoding either CD8-FAP or CD8-CD40L fusion protein were collected four days postinfection. Cell culture medium without cells was used as an additional control (medium). Serial dilutions of these materials were added to anti-CD8 antibody-coated microtiter plates and allowed to bind. cF19 (1 mg/ml) was subsequently added and allowed to bind. Bound cF19 was detected with horseradish peroxidase-conjugated anti-human antibody.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

30

The reference to any prior art in this specification is not, and should not be taken as, an acknowledgment or any form of suggestion that that prior art forms part of the common general knowledge in Australia.



SEQUENCE LISTING

<110> Boehringer Ingelheim International GmbH

<120> FAP alpha-specific antibody with improved producibility

<130> 12-196-PCT

<140> PCT/EP99/02711

<141> 1999-04-22

<150> EP 98107925.4

<151> 1998-04-30

<160> 101

<170> PatentIn Ver. 2.1

<210> 1

<211> 339

<212> DNA

<213> Homo sapiens

<400> 1

```

gacattgtga tgaccaatc tccagactct ttggctgtgt ctctagggga gagggccacc 60
atcaactgca agtccagtca gagcctttta tattctagaa atcaaaagaa ctacttggcc 120
tggtatcagc agaaaccagg acagccaccc aaactcctca tcttttgggc tagcactagg 180
gaatctgggg tacctgatag gttcagtggc agtgggtttg ggacagactt caccctcacc 240
attagcagcc tgcaggctga agatgtggca gtttattact gtcagcaata ttttagctat 300
ccgctcacgt tcggacaagg gaccaagggtg gaaataaaa 339

```

<210> 2

<211> 113

<212> PRT

<213> Homo sapiens

<400> 2

```

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

```



<210> 3
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 3
 gacattgtga tgacccaatc tccagactct ttggctgtgt ctctagggga gagggccacc 60
 atcaactgca agtccagtca gagcctttta tattctagaa atcaaaagaa ctacttggcc 120
 tggttccagc agaaaccagg acagccaccc aaactcctca tcttttgggc tagcactagg 180
 gaatctgggg tacctgatag gttcagtggc agtggggttg ggacagactt caccctcacc 240
 attagcagcc tgcaggctga agatgtggca gtttatgact gtcaacaata ttttagctat 300
 ccgctcacgt tcggacaagg gaccaagggtg gaaataaaaa 339

<210> 4
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 4
 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15
 Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30
 Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
 35 40 45
 Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60
 Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80
 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
 85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110
 Lys

<210> 5
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 5
 gacattgtga tgacccaatc tccagactct ttggctgtgt ctctagggga gagggccacc 60
 atcaactgca agtccagtca gagcctttta tattctagaa atcaaaagaa ctacttggcc 120
 tggatcagc agaaaccagg acagccaccc aaactcctca tctattgggc tagcactagg 180
 gaatctgggg tacctgatag gttcagtggc agtggggttg ggacagactt caccctcacc 240
 attagcagcc tgcaggctga agatgtggca gtttattant gtcagcaata ttttagctat 300
 ccgctcacgt tcggacaagg gaccaagggtg gaaataaaaa 339



<210> 6
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 6
 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110

Lys

<210> 7
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 7
 caggtgcaac tagtgcagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
 agctgtaaaa ctagtagata caccttcact gaatacacca tacactgggt tagacaggcc 120
 cctggccaaa ggctggagtg gataggaggt attaatecta acaatggat tcctaactac 180
 aaccagaagt tcaagggccg ggccaccttg accgtaggca agtctgccag caccgcctac 240
 atggaactgt ccagcctgcg ctccgaggac actgcagtct actactgcgc cagaagaaga 300
 atcgccctatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgct 360
 accgtctcct ca 372

<210> 8
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 8
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15

Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30

Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45



Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60

Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 9

<211> 372

<212> DNA

<213> Homo sapiens

<400> 9

caggtgcaac tagtgcagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
agctgtaaaa ctagtagata caccttcact gaatacacca tacactgggt tagacaggcc 120
cctggccaaa ggctggagtg gataggaggt attaatccta acaatgggtat tcctaactac 180
aaccagaagt tcaagggccg ggccaccttg accgtaggca agtctgccag caccgcctac 240
atggaactgt ccagcctgcg ctccgaggac actgcagtct acttctgcgc cagaagaaga 300
atcgccctatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgtc 360
accgtctcct ca 372

<210> 10

<211> 124

<212> PRT

<213> Homo sapiens

<400> 10

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

Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
20 25 30

Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
35 40 45

Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60

Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
85 90 95

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120



<210> 11
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 11
 cagggtgcaac tagtgcaagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
 agctgtaaaa ctagtagata caccttcact gaatacacca tacactgggt tagacaggcc 120
 cctggccaaa ggctggagtg gataggaggt attaataccta acaatgggtat tcctaactac 180
 aaccagaagt tcaagggccg ggtcaccatc accgtagaca cctctgccag caccgcctac 240
 atggaactgt ccagcctgcg ctccgaggac actgcagtct actactgcgc cagaagaaga 300
 atcgccctatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgct 360
 accgtctcct ca 372

<210> 12
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 12
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

<210> 13
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 13
 cagggtgcaac tagtgcaagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
 agctgtaaaa ctagtagata caccttcact gaatacacca tacactgggt tagacaggcc 120
 cctggccaaa ggctggagtg gataggaggt attaataccta acaatgggtat tcctaactac 180
 aaccagaagt tcaagggccg ggtcaccatc accgtagaca cctctgccag caccgcctac 240
 atggaactgt ccagcctgcg ctccgaggac actgcagtct acttctgcg cagaagaaga 300
 atcgccctatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgct 360
 accgtctcct ca 372



<210> 14
 <211> 124
 <212> PRT
 <213> Homo sapiens

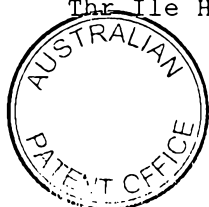
<400> 14
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

<210> 15
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 15
 caggtgcaac tagtgcagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
 agctgtaaaa ctagtggata caccttcact gaatacacca tacactgggt tagacaggcc 120
 cctggccaaa ggctggagtg gataggaggt attaatccta acaatgggtat tcctaactac 180
 aaccagaagt tcaagggccg ggtcaccatc accgtagaca cctctgccag caccgcctac 240
 atggaactgt ccagcctgcg ctccgaggac actgcagtct actactgcgc cagaagaaga 300
 atcgcttatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgctc 360
 accgtctcct ca 372

<210> 16
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 16
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Gly Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45



Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60

Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
65 70 75 80

Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95

Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110

Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 17

<211> 220

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:Chimeric light
chain variable region

<400> 17

Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala Val Ser Val Gly
1 5 10 15

Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
20 25 30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
35 40 45

Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
50 55 60

Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp Phe Asn Leu Thr
65 70 75 80

Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Asp Cys Gln Gln
85 90 95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu
100 105 110

Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp
115 120 125

Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn
130 135 140

Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu
145 150 155 160

Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp
165 170 175

Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr



180

185

190

Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser
 195 200 205

Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 210 215 220

<210> 18

<211> 453

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:chimeric heavy
 chain variable region

<400> 18

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

Val Lys Met Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr Thr
 20 25 30

Ile His Trp Val Arg Gln Ser His Gly Lys Ser Leu Glu Trp Ile Gly
 35 40 45

Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe Lys
 50 55 60

Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ser Ser Thr Ala Tyr Met
 65 70 75 80

Glu Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val Tyr Phe Cys Ala
 85 90 95

Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp Tyr
 100 105 110

Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser Ala Ser Thr Lys Gly
 115 120 125

Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr Ser Gly Gly
 130 135 140

Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro Glu Pro Val
 145 150 155 160

Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val His Thr Phe
 165 170 175

Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser Ser Val Val
 180 185 190

Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val
 195 200 205

Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
 210 215 220



Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu
225 230 235 240

Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
245 250 255

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
260 265 270

Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val
275 280 285

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser
290 295 300

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
305 310 315 320

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala
325 330 335

Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
340 345 350

Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln
355 360 365

Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
370 375 380

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
385 390 395 400

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu
405 410 415

Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
420 425 430

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
435 440 445

Leu Ser Pro Gly Lys
450

<210> 19

<211> 321

<212> DNA

<213> Homo sapiens

<400> 19

cgtactgtgg ctgcaccatc tgtcttcac tccccgccat ctgatgagca gttgaaatct 60
ggaactgcct ctgttggtg cctgctgaat aacttctatc ccagagaggc caaagtacag 120
tggaagggtg ataacgccct ccaatcgggt aactcccagg agagtgtcac agagcaggac 180
agcaaggaca gcacctacag cctcagcagc accctgacgc tgagcaaagc agactacgag 240
aaacacaaag tctacgctg cgaagtcacc catcagggcc tgagctcgcc cgtcacaaag 300
agcttcaaca ggggagagtg t 321

<210> 20



<211> 107
 <212> PRT
 <213> Homo sapiens

<400> 20

Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
 1 5 10 15

Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
 20 25 30

Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
 35 40 45

Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
 50 55 60

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

Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
 85 90 95

Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
 100 105

<210> 21
 <211> 990
 <212> DNA
 <213> Homo sapiens

<400> 21

gcctccacca agggcccacg ggtcttcccc ctggcaccct cctccaagag cacctctggg 60
 ggcacagcgg ccctgggctg cctgggtcaag gactacttcc ccgaaccggg gacgggtgtcg 120
 tggaactcag gcgccttgac cagcggcggtg cacaccttcc cggctgtcct acagtccctca 180
 ggactctact ccctcagcag cgtgggtgacc gtgccctcca gcagcttggg caccagacc 240
 tacatctgca acgtgaatca caagcccagc aacaccaagg tggacaagaa agttgagccc 300
 aaatcttggtg acaaaaactca cacatgcccc ccgtgcccag cacctgaact cctgggggga 360
 ccgtcagttct tcctcttccc cccaaaaccc aaggacaccc tcatgatctc ccggaccct 420
 gaggtcacat gcgtgggtgt ggacgtgagc cacgaagacc ctgaggtcaa gttcaactgg 480
 tacgtggacg gcgtggaggt gcataatgcc aagacaaagc cgcgaggagga gcagtacaac 540
 agcacgtacc ggggtggtcag cgtctctacc gtctctgacc aggactggct gaatggcaag 600
 gagtacaagt gcaaggtctc caacaaagcc ctcccagccc ccacgcagaa aaccatctcc 660
 aaagccaaag ggcagccccg agaaccacag gtgtacaccc tgcccccatc ccgggaggag 720
 atgaccaaga accaggtcag cctgacctgc ctgggtcaaag gcttctatcc cagcgacatc 780
 gccgtggagt gggagagcaa tgggcagccg gagaacaact acaagaccac gcctcccgtg 840
 ctggactccg acggctcctt cttcctctac agcaagctca ccgtggacaa gagcaggtgg 900
 cagcagggga acgtcttctc atgctccgtg atgcatgagg ctctgcacaa ccaactacacg 960
 cagaagagcc tctccctgtc tccgggtaaa 990

<210> 22
 <211> 330
 <212> PRT
 <213> Homo sapiens

<400> 22

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
 5 10 15



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



<210> 23
 <211> 427
 <212> DNA
 <213> Homo sapiens

<400> 23
 aagcttgccg ccaccatgga ttacacaggcc caggttctta tgttactgcc gctatgggta 60
 tctggtacct gtggggacat tgtgatgtca cagtctccat cctccctagc tgtgtcagtt 120
 ggagagaagg ttactatgag ctgcaagtcc agtcagagcc ttttatatag tcgtaatcaa 180
 aagaactact tggcctggtt ccagcagaag ccagggcagt ctcctaaact gctgattttc 240
 tgggcatcca ctagggaatc tggggtcctt gatcgcttca caggcagtggt atttgggacg 300
 gatttcaatc tcaccatcag cagtgtgcag gctgaggacc tggcagttta tgactgtcag 360
 caatatttta gctatccgct cacgttcggt gctgggacca agctggagct gaaacgtgag 420
 tggatcc 427

<210> 24
 <211> 133
 <212> PRT
 <213> Homo sapiens

<400> 24
 Met Asp Ser Gln Ala Gln Val Leu Met Leu Leu Pro Leu Trp Val Ser
 1 5 10 15
 Gly Thr Cys Gly Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala
 20 25 30
 Val Ser Val Gly Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser
 35 40 45
 Leu Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln
 50 55 60
 Lys Pro Gly Gln Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg
 65 70 75 80
 Glu Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp
 85 90 95
 Phe Asn Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr
 100 105 110
 Asp Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr
 115 120 125
 Lys Leu Glu Leu Lys
 130

<210> 25
 <211> 457
 <212> DNA
 <213> Homo sapiens

<400> 25
 aagcttgccg ccaccatggg atggagctgg gtctttctct ttctcctgtc aggaactgca 60
 ggtgtcctct ctgagggtcca gctgcaacag tctggacctg agctggtgaa gcctggggct 120
 tcagtaaaga tgtcctgcaa gacttctaga tacacattca ctgaatacac catacactgg 180
 gtgagacaga gccatggaaa gagccttgag tggattggag gtattaatcc taacaatggt 240
 attcctaact acaaccagaa gttcaagggc agggccacat tgactgtagg caagtcctcc 300



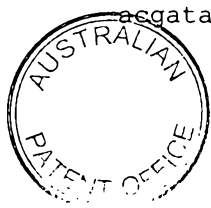
agcaccgcct acatggagct ccgcagcctg acatctgagg attctgcggt ctatttctgt 360
 gcaagaagaa gaatcgctta tggttacgac gagggccatg ctatggacta ctgggggtcaa 420
 ggaacctcag tcaccgtctc ctcaggtgag tggatcc 457

<210> 26
 <211> 143
 <212> PRT
 <213> Homo sapiens

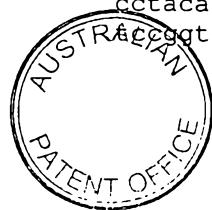
<400> 26
 Met Gly Trp Ser Trp Val Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly
 1 5 10 15
 Val Leu Ser Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys
 20 25 30
 Pro Gly Ala Ser Val Lys Met Ser Cys Lys Thr Ser Arg Tyr Thr Phe
 35 40 45
 Thr Glu Tyr Thr Ile His Trp Val Arg Gln Ser His Gly Lys Ser Leu
 50 55 60
 Glu Trp Ile Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn
 65 70 75 80
 Gln Lys Phe Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ser Ser
 85 90 95
 Thr Ala Tyr Met Glu Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val
 100 105 110
 Tyr Phe Cys Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His
 115 120 125
 Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
 130 135 140

<210> 27
 <211> 8068
 <212> DNA
 <213> Homo sapiens

<400> 27
 gaattccagc aactggcgcg ccgttactag ttattaatag taatcaatta cgggggtcatt 60
 agttcatagc ccatatatgg agttccgcgt tacataactt acggtaaagt gccgcctgg 120
 ctgaccgccc aacgaccccc gccattgac gtcaataatg acgtatgttc ccatagtaac 180
 gccaataggg actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt 240
 ggcagtacat caagtgtatc atatgccaag tacgccccct attgacgtca atgacggtaa 300
 atggccccgc tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta 360
 catctacgta ttagtcatcg ctattacat ggtgatgcgg ttttggcagt acatcaatgg 420
 gcgtggatag cggtttgact cacggggatt tccaagtctc caccctattg acgtcaatgg 480
 gagtttgatt tggcaccaaa atcaacggga ctttccaaaa tgcgtaaca actccgcccc 540
 attgacgcaa atgggcggta ggcgtgtacg gtgggaggtc tatataagca gagctcgatt 600
 agtgaaccgt cagatcgctt ggagacgcca tccacgctgt tttgacctcc atagaagaca 660
 ccgggaccga tccagcctcc gcggccggga acggtgcatt ggaacgcgga ttccccgtgc 720
 caagagtgcg gtaagtaccg cctatagagt ctataggccc acccccttgg cttcttatgc 780
 atgctatact gtttttggtt tgggggtctat acacccccgc ttccctcatgt tataggatg 840
 ggtatagctt agcctatagg tgtgggttat tgaccattat tgaccactcc cctattggtg 900
 acgatacttt ccattactaa tccataacat ggctctttgc cacaactctc tttattggct 960



atatgccaat	acactgtcct	tcagagactg	acacggactc	tgtattttta	caggatgggg	1020
tctcatttat	tatttacaaa	ttcacatata	caacaccacc	gtccccagtg	cccgcagttt	1080
ttattaaaca	taacgtggga	tctccacgcg	aatctcgggt	acgtgttccg	gacatgggct	1140
cttctccggt	agcggcggag	cttctacatc	cgagccctgc	tcccatgcct	ccagcgactc	1200
atggtcgctc	ggcagctcct	tgtcctaacc	agtggaggcc	agacttaggc	acagcacgat	1260
gccaccacc	accagtgtgc	cgcacaaggc	cgtggcggta	gggtatgtgt	ctgaaaatga	1320
gctcggggag	cgggcttgca	ccgctgacgc	atttgggaaga	cttaaggcag	cggcagaaga	1380
agatgcaggc	agctgagttg	ttgtgttctg	ataagagtca	gaggtaaactc	ccgttgcggt	1440
gctgttaacg	gtggagggca	gtgtagtctg	agcagtactc	gttgctgccg	cgcgcgccac	1500
cagacataat	agctgacaga	ctaacagact	gttcctttcc	atgggtcctt	tctgcagtca	1560
ccgtccttga	cacgcgtctc	gggaagcttg	ccgccaccat	ggattcacag	gcccaggttc	1620
ttatgttact	gccgctatgg	gtatctggta	cctgtgggga	cattgtgatg	tcacagttctc	1680
catcctccct	agctgtgtca	gttgagagaga	aggttactat	gagctgcaag	tccagtcaga	1740
gccttttata	ttctagaaat	caaaagaact	acttggcctg	gttccagcag	aagccagggc	1800
agtctcctaa	actgctgatt	ttctgggcat	ccactaggga	atctggggtc	cctgatcgct	1860
tcacaggcag	tggatttggg	acggatttca	atctcaccat	cagcagtgtg	caggctgagg	1920
acctggcagt	ttatgactgt	cagcaatatt	ttagctatcc	gctcacgttc	gggtctggga	1980
ccaagctgga	gctgaaacgt	gagtggatcc	atctgggata	agcatgctgt	tttctgtctg	2040
tccctaacat	gccctgtgat	tatgcgcaaa	caacacaccc	aagggcagaa	ctttgttact	2100
taaacaccat	cctgtttgct	tctttcctca	ggaactgtgg	ctgcaccatc	tgtcttcac	2160
ttcccgccat	ctgatgagca	gttgaaatct	ggaactgcct	ctgttggtgtg	cctgctgaat	2220
aacttctatc	ccagagaggc	caaagtacag	tggaaagggtg	ataacgccct	ccaatcgggt	2280
aactcccagg	agagtgtcac	agagcaggac	agcaaggaca	gcacctacag	cctcagcagc	2340
accctgacgc	tgagcaaagc	agactacgag	aaacacaaag	tctacgcctg	cgaagtcacc	2400
catcagggcc	tgagctcgcc	cgtcacaaag	agcttcaaca	ggggagagtgt	ttagaggagg	2460
aagtgcccc	acctgctcct	cagttccagc	ctgacccccct	cccacccctt	ggcctctgac	2520
cctttttcca	caggggacct	acccctattg	cggctcctca	gctcatcttt	cacctcacc	2580
ccctcctcct	ccttggcttt	aattatgcta	atgttgagg	agaatgaata	aataaagtga	2640
atctttgcac	ctgtgggtgga	tctaataaaa	gatatttatt	ttcattagat	atgtgtgttg	2700
gttttttgtg	tgcagtgcct	ctatctggag	gccaggtagg	gctggccttg	ggggaggggg	2760
aggccagaat	gactccaaga	gctacaggaa	ggcaggtcag	agacccccact	ggacaaacag	2820
tggctggact	ctgcaccata	acacacaatc	aacaggggag	tgagctggaa	atttgctagc	2880
gaattcttga	agacgaaagg	gcctcgtgat	acgcctat	ttataggtta	atgtcatgat	2940
aataatgggt	tcttagacgt	caggtggcac	ttttcgggga	aatgtgcgcg	gaaccctat	3000
ttgttttatt	ttctaataac	attcaaatat	gtatccgctc	atgagacaat	aacctgata	3060
aatgcttcaa	taatagttaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	3120
tattcccttt	tttgcggcat	tttgccctcc	tgtttttgct	caccagaaa	cgtgggtgaa	3180
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	3240
cagcggtaag	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	tgagcacttt	3300
taaagtctctg	ctatgtggcg	cggtattatc	ccgtgttgac	gccgggcaag	agcaactcgg	3360
tcgccgcata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	3420
tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	tgagtataa	3480
cactgcggcc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	ccgctttttt	3540
gcacaacatg	gggatcatg	taactcgcct	tgatcgttgg	gaaccggagc	tgaatgaagc	3600
cataccaaac	gacgagcgtg	acaccacgat	gcctgcagca	atggcaacaa	cgttgcgcaa	3660
actattaact	ggcgaaactac	ttactctagc	ttccgggcaa	caattaatag	actggatgga	3720
ggcggataaa	ggtgcaggac	cacttctgcg	ctcggccctt	ccggctggct	ggtttattgc	3780
tgataaatct	ggagccgggtg	agcgtgggtc	tcgcggtatc	attgcagcac	tggggccaga	3840
tggtaaagccc	tcccgtatcg	tagttatcta	cacgacgggg	agtcaggcaa	ctatggatga	3900
acgaaataga	cagatcgctg	agatagggtc	ctcactgatt	aagcattggg	aactgtcaga	3960
ccaagtttac	tcatatatac	tttagattga	tttaaaactt	catttttaat	ttaaaaggat	4020
ctaggtgaag	atcctttttg	ataatctcat	gaccaaaatc	ccttaacgtg	agttttcgtt	4080
ccactgcgcg	tcagaccccg	tagaaaagat	caaaggatct	tcttgagatc	ctttttttct	4140
cgcgctaate	tgtcgttgcg	aaacaaaaaa	accacgccta	ccagcgggtg	tttgtttgcc	4200
ggatcaagag	ctaccaactc	tttttccgaa	ggtaactggc	ttcagcagag	cgcagatacc	4260
aaatactgtc	cttctagtgt	agccgtagt	aggccaccac	ttcaagaact	ctgtagcacc	4320
gcctacatac	ctcgtctctg	taatcctgtt	accagtggct	gctgccagt	gcgataagtc	4380
gtgtcttacc	gggttggaact	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	4440
aacgggggg	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	4500
cctacagcgt	gagctatgag	aaagcggcac	gcttcccga	gggagaaagg	cggacaggta	4560
acccgtaagc	ggcagggtcg	gaacaggaga	gcgcacgagg	gagcttccag	ggggaaacgc	4620



ctggtatctt	tatagtccctg	tcgggttttcg	ccacctctga	cttgagcgtc	gattttttgtg	4680
atgctcgtca	ggggggcgga	gcctatggaa	aaacgccagc	aacgcggcct	ttttacgggtt	4740
cctggccttt	tgtctggcctt	ttgtctacat	gttcttttct	gcgttatccc	ctgattctgt	4800
ggataaccgt	attaccgcct	ttgagtgcgc	tgataccgct	cgccgcagcc	gaacgaccga	4860
gcgcagcgag	tcagtgcgcg	aggaagcgga	agagcgcctg	atgcggtatt	ttctccttac	4920
gcatctgtgc	ggatatttcac	accgcatatg	gtgcactctc	agtacaatct	gctctgatgc	4980
cgcatagtta	agccagtata	cactccgcta	tcgctacgtg	actgggtcat	ggctgcgccc	5040
cgacacccgc	caacacccgc	tgacgcgccc	tgacgggctt	gtctgctccc	ggcatccgct	5100
tacagacaag	ctgtgaccgt	ctccgggagc	tgcatgtgtc	agaggttttc	accgtcatca	5160
ccgaaacgcg	cgaggcagct	gtggaatgtg	tgtagttag	gggtgggaaa	gtccccaggc	5220
ccccagcag	gcagaagtat	gcaaaagcatg	catctcaatt	agtcagcaac	caggctcccc	5280
agcaggcaga	agtatgcaaa	gcatgcatct	caattagtcg	gcaaccatag	ccccgcccct	5340
aactccgccc	atcccgcctt	taactccgcc	cagttccgcc	cattctccgc	cccatggctg	5400
actaattttt	tttattttatg	cagaggccga	ggccgcctcg	gcctctgagc	tattccagaa	5460
gtagtgcgga	ggctttttttg	gaggcctagg	cttttgcaaa	aagctagctt	cacgtccgca	5520
caagcactca	gggcgcaagg	gctgctaaag	gaagcgggaa	acgtagaaa	ccagtccgca	5580
gaaacggtgc	tgacccccga	tgaatgtcag	ctactgggct	atctggacaa	gggaaaacgc	5640
aagcgcaaa	agaaagcag	tagcttgca	tgggcttaca	tggcgatagc	tagactgggc	5700
ggttttatgg	acagcaagcg	aaccgggaatt	gccagctggg	gcgccctctg	gtaaggttgg	5760
gaagccctgc	aaagtaaaact	ggatggcttt	cttgccgcca	aggatctgat	ggcgaggggg	5820
atcaagatct	gatcaagaga	caggatgagg	atcgtttcgc	atgattgaac	aagatggatt	5880
gcacgcaggt	tctccggccg	cttgggtgga	gaggctattc	ggctatgact	gggcacaaca	5940
gacaatcggc	tgtctctgatg	ccgccgtgtt	ccggctgtca	gcgcaggggc	gcccggttct	6000
ttttgtcaag	accgacctgt	ccggtgccct	gaatgaactg	caggacgagg	cagcgcggt	6060
atcggtggctg	gccacgacgg	gcgttccttg	cgcagctgtg	ctcgacgttg	tactgaagc	6120
gggaaggggac	tggctgctat	tgggcgaagt	gccggggcag	gatctcctgt	catctcacct	6180
tgctcctgcc	gagaaagtat	ccatcatggc	tgatgcaatg	cggcggctgc	atacgcttga	6240
tccggctacc	tgccccattcg	accaccaagc	gaaacatcgc	atcgagcgag	cacgtactcg	6300
gatggaagcc	ggctctgtcg	atcaggatga	tctggacgaa	gagcatcagg	ggctcgcgcc	6360
agccgaactg	ttcgccaggc	tcaaggcgcg	catgcccgac	ggcgaggatc	tcgtcgtgac	6420
ccatggcgat	gcctgcttgc	cgaatatcat	gggtgaaaaat	ggccgctttt	ctggattcat	6480
cgactgtggc	cggtggttg	tggcggaccg	ctatcaggac	atagcgttg	ctaccctga	6540
tattgctgaa	gagcttggcg	gcgaatgggc	tgaccgcttc	ctcgtgcttt	acggtatcgc	6600
cgtccccgat	tcgcagcgca	tcgccttcta	tcgccttctt	gacgagttct	tctgagcggg	6660
actctggggg	tcgaaatgac	cgaccaagcg	acgccaacc	tgccatcacg	agatttcgat	6720
tccaccgcgc	ccttcctatga	aaggttgggc	tctggaatcg	ttttccggga	cgccggctgg	6780
atgatccctc	agcgcgggga	tctcatgctg	gagtctctcg	cccaccccg	gctcgatccc	6840
ctcgcgagtt	ggttcagctg	ctgcctgagg	ctggacgacc	tcgcggagtt	ctaccggcag	6900
tgcaaatccg	tcggcatcca	ggaaaccagc	agcggctatc	cgcgcatcca	tgcccccgaa	6960
ctgcaggagt	ggggaggcac	gatggccgct	ttggtcccg	atctttgtga	aggaacctta	7020
cttctgtggt	gtgacataat	tggacaaact	acctacagag	atttaaagct	ctaaggtaaa	7080
tataaaattt	ttaagtgtat	aatgtgttaa	actactgatt	ctaattgttt	gtgtatttta	7140
gattccaacc	tatggaactg	atgaatggga	gcagtgtgtg	aatgccttta	atgaggaaaa	7200
cctgttttgc	tcagaagaaa	tgccatctag	tgatgatgag	gctactgctg	actctcaaca	7260
ttctactcct	ccaaaaaaga	agagaaaggt	agaagacccc	aaggactttc	cttcagaatt	7320
gctaagtttt	ttgagtcatg	ctgtgttttag	taatagaact	cttgcttgct	ttgctattta	7380
caccacaaaag	gaaaaagctg	cactgctata	caagaaaatt	atggaaaaat	attctgtaac	7440
ctttataagt	aggcataaca	gttataatca	taacatactg	ttttttctta	ctccacacag	7500
gcatagagtg	tctgctatta	ataactatgc	tcaaaaattg	tgtaccttta	gctttttaat	7560
ttgtaaaggg	gttaataagg	aatattttgat	gtatagtgcc	ttgactagag	atcataatca	7620
gccataccac	attttagtag	gttttacttg	ctttaaaaaa	cctcccacac	ctccccctga	7680
acctgaaaca	taaaatgaat	gcaattgttg	ttgttaactt	gtttattgca	gcttataatg	7740
gttacaaata	aagcaatagc	atcacaat	tcacaaataa	agcatttttt	tactgcatt	7800
ctagtgtgtg	tttgtccaaa	ctcatcaatg	tatcttatca	tgtctggatc	taataaaaaga	7860
tatttatttt	cattagatat	gtgtgtttgt	tttttgtgtg	cagtgcctct	atctggaggc	7920
caggtagggc	tggccttggg	ggagggggag	gccagaatga	ctccaagagc	tacaggagg	7980
caggtcagag	acccactgg	acaaacagtg	gctggactct	gcaccataac	acacaatcaa	8040
caggggagtg	agctggaaat	ttgctagc				8068



<211> 239
 <212> PRT
 <213> Homo sapiens

<400> 28

```

Asp Ser Gln Ala Gln Val Leu Met Leu Leu Pro Leu Trp Val Ser Gly
 1          5          10          15

Thr Cys Gly Asp Ile Val Met Ser Gln Ser Pro Ser Ser Leu Ala Val
 20          25          30

Ser Val Gly Glu Lys Val Thr Met Ser Cys Lys Ser Ser Gln Ser Leu
 35          40          45

Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys
 50          55          60

Pro Gly Gln Ser Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu
 65          70          75          80

Ser Gly Val Pro Asp Arg Phe Thr Gly Ser Gly Phe Gly Thr Asp Phe
 85          90          95

Asn Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr Asp
100          105          110

Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr Lys
115          120          125

Leu Glu Leu Lys Arg Thr Val Ala Ala Pro Ser Val Phe Ile Phe Pro
130          135          140

Pro Ser Asp Glu Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu
145          150          155          160

Leu Asn Asn Phe Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp
165          170          175

Asn Ala Leu Gln Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp
180          185          190

Ser Lys Asp Ser Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys
195          200          205

Ala Asp Tyr Glu Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln
210          215          220

Gly Leu Ser Ser Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Cys
225          230          235

```

<210> 29
 <211> 7731
 <212> DNA
 <213> Homo sapiens

<400> 29

```

ttgaagacga aagggcctcg tgatacgctt atttttatag gttaatgtca tgataataat 60
ggtttcttag acgtcaggtg gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt 120
atttttctaa atacattcaa atatgtatcc gctcatgaga caataaccct gataaatgct 180
tcaataatat tgaaaaagga agagtatgag tattcaacat ttccgtgtcg cccttattcc 240

```



cttttttgcg	gcatttttgc	ttcctgtttt	tgctcaccca	gaaacgctgg	tgaaagtaaa	300
agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	gaactggatc	tcaacagcgg	360
taagatccct	gagagttttc	gccccgaaga	acgttttcca	atgatgagca	cttttaaagt	420
tctgctatgt	ggcgcggtat	tatcccgtgt	tgacgcgggg	caagagcaac	tcggtcgccc	480
catacactat	tctcagaatg	acttggttga	gtactcacca	gtcacagaaa	agcatcttac	540
ggatggcatg	acagtaagag	aattatgcag	tgctgccata	accatgagtg	ataacactgc	600
ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	ctaaccgctt	ttttgcacaa	660
catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	720
aaacgacgag	cgtgacacca	cgatgcctgc	agcaatggca	acaacgttgc	gcaaaactatt	780
aactggcgaa	ctacttactc	tagcttcccc	gcaacaatta	atagactgga	tgagggcgga	840
taaagttgca	ggaccacttc	tgcgctcggc	ccttcgggct	ggctggttta	ttgctgataa	900
atctggagcc	ggtgagcgtg	ggtctcgcg	tatcattgca	gcactggggc	cagatggtaa	960
gccctcccgt	atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	1020
tagacagatc	gctgagatag	gtgcctcact	gattaaagcat	tggttaactgt	cagaccaagt	1080
ttactcatat	atactttaga	ttgattttaa	acttcatttt	taatttaaaa	ggatctaggt	1140
gaagatccct	tttgataatc	tcatagacaa	aatcccttaa	cgtgagtttt	cgttccactg	1200
agcgtcagac	cccgtagaaa	agatcaaagg	atcttcttga	gatccttttt	ttctgcgctg	1260
aatctgctgc	ttgcaaacaa	aaaaaccacc	gctaccagcg	gtgggtttgt	tgccggatca	1320
agagctacca	actctttttc	cgaaggtaac	tggttccagc	agagcgcaga	taccaaatac	1380
tgctcttcta	gtgtagccgt	agttaggcca	ccacttcaag	aactctgtag	caccgcctac	1440
atacctcgct	ctgctaattc	tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	1500
taccgggttg	gactcaagac	gatagtacc	ggataaggcg	cagcggtcgg	gctgaacggg	1560
gggttcgtgc	acacagccca	gcttgagcgc	aacgacctac	accgaactga	gatacctaca	1620
gcgtgagcta	tgagaaagcg	ccacgcttcc	cgaagggaga	aaggcggaca	ggtatccggt	1680
aagcggcagg	gtcggaaacg	gagagcgcac	gagggagctt	ccagggggaa	acgcctggta	1740
tctttatagt	cctgtcgggt	ttcgccacct	ctgacttgag	cgtcgatttt	tgtgatgtct	1800
gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	gcctttttac	ggttcctggc	1860
cttttgctgg	ccttttgctc	acatgttctt	tctgcgttta	tccccgtatt	ctgtggataa	1920
ccgtattacc	gcctttgagt	gagctgatac	cgctcgccgc	agccgaacga	ccgagcgcag	1980
cgagtcagtg	agcgaggaag	cggaaagcgc	cctgatgcgg	tattttctcc	ttacgcctct	2040
gtgcggtatt	tcacaccgca	tatggtgcac	tctcagtaca	atctgctctg	atgccgcata	2100
gttaagccag	tatacactcc	gctatcgcta	cgtgactggg	tcatggctgc	gccccgacac	2160
ccgccaacac	ccgctgacgc	gccctgacgg	gcttgctcgc	tcccggcatc	cgcttacaga	2220
caagctgtga	ccgtctccgg	gagctgcatg	tgctcagagg	tttcaccgtc	atcaccgaaa	2280
cgcgcgaggc	agcatgcatc	tcaattagtc	agcaaccata	gtcccgcccc	taactccgcc	2340
catcccgcgc	ctaactccgc	ccagttccgc	ccattctccg	ccccatggct	gactaatttt	2400
ttttatttat	gcagaggccg	aggccgcctc	ggcctctgag	ctattccaga	agtagtgagg	2460
aggctttttt	ggaggcctag	gcttttgcaa	aaagctagct	tacagctcag	ggctgcgatt	2520
tcgcgccaaa	cttgacggca	atcctagcgt	gaaggctggt	aggattttat	ccccgctgcc	2580
atcatggttc	gaccattgaa	ctgcacgtc	gccgtgtccc	aaaatatggg	gattggcaag	2640
aacggagacc	taccctggcc	tccgctcagg	aacgagttca	agtacttcca	aagaatgacc	2700
acaacctctt	cagtggaaag	taaacagaat	ctggtgatta	tgggtaggaa	aacctgggtc	2760
tccattcctg	agaagaatcg	acctttaaag	gacagaatta	atatagttct	cagtagagaa	2820
ctcaaagaac	caccacgagg	agctcatttt	cttgccaaaa	gtttggatga	tgctttaaga	2880
cttattgaac	aaccggaatt	ggcaagtaaa	gtagacatgg	tttgatagat	cggaggcagt	2940
tctgtttacc	aggaagccat	gaatcaacca	ggccacctca	gactctttgt	gacaaggatc	3000
atgcaggaat	ttgaaagtga	cacgtttttc	ccagaaattg	atgtggggaa	atataaactt	3060
ctcccagaat	acccaggcgt	cctctctgag	gtccaggagg	aaaaaggcat	caagtataag	3120
tttgaagtct	acgagaagaa	agactaacag	gaagatgctt	tcaagttctc	tgctccccct	3180
ctaaagctat	gcatttttat	aagaccatgg	gacttttgct	ggcttttagat	ctttgtgaag	3240
gaaccttact	tctgtggtgt	gacataattg	gacaaactac	ctacagagat	ttaaagctct	3300
aaggtaaata	taaaattttt	aagtgtataa	tgtgttaaac	tactgattct	aattgtttgt	3360
gtattttaga	ttccaacctt	tggaaactgat	gaatgggagc	agtgggtggaa	tgcttttaat	3420
gaggaaaacc	tgttttgctc	agaagaaatg	ccatctagtg	atgatgaggc	tactgctgac	3480
tctcaacatt	ctactcctcc	aaaaaagaag	agaaaggtag	aagaccccaa	ggactttcct	3540
tcagaattgc	taagtttttt	gagtcagtct	gtgttttagta	atagaactct	tgcttgcttt	3600
gctattttaca	ccacaaagga	aaaagctgca	ctgctatata	agaaaattat	ggaaaaatat	3660
tctgtaacct	ttataagtag	gcataacagt	tataatcata	acatactgtt	ttttcttact	3720
ccacacaggc	atagagtgtc	tgctattaat	aactatgctc	aaaaattgtg	tacctttagc	3780
tttttaattt	gtaaaggggt	taataaggaa	tatttgatgt	atagtgcctt	gactagagat	3840
cataatcage	cataccacat	ttgtagaggt	tttacttgct	ttaaaaaacc	tcccacacct	3900



ccccctgaac	ctgaaacata	aaatgaatgc	aattgttgtt	gttaacttgt	ttattgcagc	3960
ttataatggt	tacaaataaaa	gcaatagcat	cacaaatttc	acaaataaag	catttttttc	4020
actgcattct	agttgtggtt	tgtccaaact	catcaatgta	tcttatcatg	tctggatcta	4080
ataaaaagata	tttatttttca	ttagatatgt	gtgttgggtt	tttgtgtgca	gtgcctctat	4140
ctggaggcca	ggtagggctg	gccttggggg	agggggaggc	cagaatgact	ccaagagcta	4200
caggaaggca	ggtcagagac	cccactggac	aaacagtggc	tggactctgc	accataacac	4260
acaatcaaca	ggggagttag	ctggaaattt	gctagcgaat	tccagcacac	tggcggccgt	4320
tactagttaa	taatagtaat	caattacggg	gtcattagtt	catagcccat	atatggagtt	4380
ccgcgttaca	taacttacgg	taaatggccc	gcctggctga	ccgccaacg	acccccgcc	4440
attgacgtca	ataatgacgt	atgttcccat	agtaacgcca	atagggactt	tccattgacg	4500
tcaatgggtg	gagtattttac	ggtaaactgc	ccacttggca	gtacatcaag	tgtatcatat	4560
gccaaagtacg	ccccctattg	acgtcaatga	cggtaaattg	ccgcctggc	attatgccca	4620
gtacatgacc	ttatgggact	ttcctacttg	gcagtacatc	tacgtattag	tcacgctat	4680
taccatggtg	atgcggtttt	ggcagtacat	caatgggcgt	ggatagcggg	ttgactcacg	4740
gggatttcca	agtctccacc	ccattgacgt	caatgggagt	ttgttttggc	accaaataca	4800
acgggacttt	ccaaaatgtc	gtaacaactc	cgccccattg	acgcaaatgg	gcggtaggcg	4860
tgtacggtgg	gaggtctata	taagcagagc	tcgttttagtg	aaccgtcaga	tcgcctggag	4920
acgccatcca	cgctgttttg	acctccatag	aagacaccgg	gaccgatcca	gcctccgcgg	4980
ccgggaacgg	tgcattggaa	cgcggattcc	ccgtgccaa	agtgcgtaa	gtaccgccta	5040
tagagtctat	aggccccacc	ccttggcttc	ttatgcatgc	tatactgttt	ttggcttggg	5100
gtctatacac	ccccgccttc	tcatgttata	ggtgatggta	tagcttagcc	tataggtgtg	5160
ggttattgac	cattattgac	cactccccta	ttggtgacga	tactttccat	tactaatcca	5220
taacatggct	ctttgccaca	actctcttta	ttggctatat	gccaatcac	tgtccttcag	5280
agactgacac	ggactctgta	tttttacagg	atggggtctc	atttattatt	tacaaattca	5340
catatacaac	accaccgtcc	ccagtggccg	cagtttttat	taaacataac	gtgggatctc	5400
cacgcgaatc	tcgggtacgt	gttccggaca	tgggctcttc	tccggtagcg	gcggagcttc	5460
tacatccgag	ccctgctccc	atgcctccag	cgactcatgg	tcgctcggca	gctccttgct	5520
cctaacagtg	gaggccagag	ttaggcacag	cacgatgccc	accaccacca	gtgtgccgca	5580
caaggccgtg	gcggtagggg	atgtgtctga	aaatgagctc	ggggagcggg	cttgaccgc	5640
tgacgcattt	ggaagactta	aggcagcggc	agaagaagat	gcaggcagct	gagttgttgt	5700
gttctgataa	gagtcagagg	taactcccgt	tgcggtgctg	ttaacggtgg	agggcagtgt	5760
agtctgagca	gtactcgttg	ctgccgcgcg	cgccaccaga	cataatagct	gacagactaa	5820
cagactgttc	ctttccatgg	gtcttttctg	cagtcaccgt	ccttgacacg	cgtctcggga	5880
agcttgccgc	caccatggga	tggagctggg	tctttctctt	tctcctgtca	ggaactgcag	5940
gtgtcctctc	tgaggtccag	ctgcaacagt	ctggacctga	gctggtgaag	cctggggctt	6000
cagtaaagat	gtcctgcaag	acttctagat	acacattcac	tgaatacacc	atacactggg	6060
tgagacagag	ccatggaaa	agccttgagt	ggattggagg	tattaatcct	aacaatggta	6120
ttcctaacta	caaccagaag	ttcaagggca	gggccacatt	gactgtaggc	aagtcctcca	6180
gcaccgccta	catggagctc	cgcagcctga	catctgagga	ttctgcggtc	tatttctgtg	6240
caagaagaag	aatcgccctat	ggttacgacg	agggccatgc	tatggactac	tggggtcaag	6300
gaacctcagt	caccgtctcc	tcaggtgagt	ggatcctctg	cgctggggcc	cagctctgtc	6360
ccacaccgcg	gtcacatggc	accacctctc	ttgcagcctc	caccaagggc	ccatcgggtc	6420
tccccctggc	accctcctcc	aagagcacct	ctggggggcac	agcgccctg	ggctgcctgg	6480
tcaaggacta	cttccccgaa	ccggtgacgg	tgtcgtggaa	ctcaggcgcc	ctgaccagcg	6540
gcgtgcacac	cttcccggct	gtcctacagt	cctcaggact	ctactccctc	agcagcgtgg	6600
tgaccgtgcc	ctccagcagc	ttgggcaccc	agacctacat	ctgcaacgtg	aatcacaagc	6660
ccagcaatcac	caaggtggac	aagaaaagttg	agcccaaattc	ttgtgacaaa	actcacacat	6720
gcccacogtg	cccagcacct	gaactcctgg	ggggaccgtc	agtcttcctc	ttcccccaa	6780
aacccaagga	caccctcatg	atctcccggg	cccctgaggt	cacatgcgtg	gtggtggacg	6840
tgagccacga	agacctgag	gtcaagttca	actggtacgt	ggacggcggtg	gaggtgcata	6900
atgccaaagac	aaagccgcgg	gaggagcagt	acaacagcac	gtaccgggtg	gtcagcgtcc	6960
tcaccgtcct	gcaccaggac	tggctgaatg	gcaaggagta	caagtgcagg	gtctccaaca	7020
aagccctccc	agcccccatc	gagaaaacca	tctccaaagc	caaagggcag	ccccgagaac	7080
cacaggtgta	caccctgccc	ccatcccggg	aggagatgac	caagaaccag	gtcagcctga	7140
cctgcctggt	caaaggcttc	tatcccagcg	acatcgccgt	ggagtgggag	agcaatgggc	7200
agccggagaa	caactacaag	accacgcctc	ccgtgctgga	ctccgacggc	tccttcttcc	7260
tctacagcaa	gctcaccgtg	gacaagagca	ggtggcagca	ggggaacgtc	ttctcatgct	7320
ccgtgatgca	tgaggctctg	cacaaccact	acacgcagaa	gagcctctcc	ctgtctccgg	7380
gtaaatagagt	gcgacggccg	gcaagccccg	ctccccgggc	tctcgcggtc	gcacgaggat	7440
gcttggcacg	tacccctgt	acatacttcc	cgggcgccca	gcatggaaat	aaagcaccgg	7500
atctaataaa	agatattttat	tttcattaga	tatgtgtgtt	ggttttttgt	gtgcagtgcc	7560



tctatctgga ggccaggtag ggctggcctt gggggagggg gaggccagaa tgactccaag 7620
 agctacagga aggcaggtca gagacccac tggacaaaca gtggctggac tctgcaccat 7680
 aacacacaat caacagggga gtgagctgga aatttgctag cgaattaatt c 7731

<210> 30
 <211> 472
 <212> PRT
 <213> Homo sapiens

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



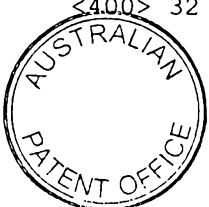
Lys Asp Thr Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val
 275 280 285
 Val Asp Val Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val
 290 295 300
 Asp Gly Val Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln
 305 310 315 320
 Tyr Asn Ser Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln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

<210> 31
 <211> 339
 <212> DNA
 <213> Homo sapiens

<400> 31
 gacattgtga tgacccaatc tccagactct ttggctgtgt ctctagggga gagggccacc 60
 atcaactgca agtccagtca gagcctttta tattctagaa atcaaaagaa ctacttggcc 120
 tggatcagc agaaaccagg acagccaccc aaactcctca tcttttgggc tagcactagg 180
 gaatctgggg tacctgatag gttcagtggc agtggggttg ggacagactt caccctcacc 240
 attagcagcc tgcaggctga agatgtggca gtttattact gtcagcaata ttttagctat 300
 ccgctcacgt tcggacaagg gaccaaggtg gaaataaaa 339

<210> 32
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 32



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

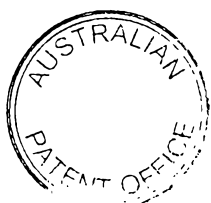
Lys

<210> 33
 <211> 113
 <212> PRT
 <213> Homo sapiens

<400> 33
 Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15
 Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30
 Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
 35 40 45
 Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60
 Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80
 Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
 85 90 95
 Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110

Lys

<210> 34
 <211> 113
 <212> PRT
 <213> Homo sapiens



<400> 34

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110

Lys

<210> 35

<211> 8068

<212> DNA

<213> Homo sapiens

<400> 35

gaattccagc aactggcgg ccgttactag ttattaatag taatcaatta cggggtcatt 60
 agttcatagc ccatatatgg agttccgcgt tacataactt acggtaaagt gccgcctgg 120
 ctgaccgccc aacgaccccc gccattgac gtcaataatg acgtatgttc ccatagtaac 180
 gccaataggg actttccatt gacgtcaatg ggtggagtat ttacggtaaa ctgcccactt 240
 ggcagtacat caagtgtatc atatgccaag tacgccccct attgacgtca atgacggtaa 300
 atggcccggc tggcattatg cccagtacat gaccttatgg gactttccta cttggcagta 360
 catctacgta ttagtcacg ctattaccat ggtgatgcgg ttttggcagt acatcaatgg 420
 gcgtggatag cggtttgact cacggggatt tccaagtctc caccocattg acgtcaatgg 480
 gagtttgttt tggcaccaaa atcaacggga ctttccaaaa tgcgtaaca actccgcccc 540
 attgacgcaa atgggcggta ggcggtgacg gtgggagggtc tatataagca gagctcgttt 600
 agtgaaccgt cagatgcctt ggagacgcca tccacgctgt tttgacctcc atagaagaca 660
 ccgggaccga tccagcctcc gcggccggga acggtgcatt ggaacgcgga ttccccgtgc 720
 caagatgacg gtaagtaccg cctatagagt ctataggccc acccccttgg cttcttatgc 780
 atgtatatac gtttttggct tggggcttat acacccccgc ttcctcatgt tataggtgat 840
 ggtatagctt agcctatagg tgtgggttat tgaccattat tgaccactcc cctattgggtg 900
 acgatacttt ccattactaa tccataacat ggctctttgc cacaactctc tttattggct 960
 atatgccaat aactgtcct tcagagactg acacggactc tgtattttta caggatgggg 1020
 tctcatttat tatttataaa ttcacatata caacaccacc gtccccagtg cccgcagttt 1080
 ttattaaaca taacgtggga tctccacgcg aatctcgggt acgtgttccg gacatgggct 1140
 cttctccggt agcggcggag cttctacatc cgagccctgc tcccatgcct ccagcgactc 1200
 atggtcgctc ggcagctcct tgctcctaac agtggaggcc agacttaggc acagcacgat 1260
 gccaccacc accagtgtgc cgcacaaggc cgtggcggtg gggatatgtg ctgaaaatga 1320
 gctcggggag cgggcttgca ccgctgacgc atttgaaga ctaaggcag cggcagaaga 1380
 agatgcaggc agctgagttg ttgtgttctg ataagagtca gaggttaactc ccgttgcggt 1440
 gctgttaacg gtggagggca gtgtagtctg agcagtactc gttgctgccg cgcgcgccac 1500
 cagacataat agctgacaga ctaacagact gttccttcc atgggtcttt tctgcagtca 1560
 ccgtccttga cagcgtctc gggaagcttg ccgccaccat ggagacagac aactcctgc 1620
 tatgggtgct gctgctctgg gttccaggtt cctccggaga cattgtgatg acccaatctc 1680
 cagactcttt ggctgtgtct ctaggggaga gggccaccat caactgcaag tccagtcaga 1740



gccttttata	ttctagaaat	caaaagaact	acttggcctg	gtatcagcag	aaaccaggac	1800
agccacccaa	actcctcatc	ttttgggcta	gcactagggg	atctggggta	cctgatagg	1860
tcagtggcag	tgggttttgg	acagacttca	ccctcaccat	tagcagcctg	caggctgaag	1920
atgtggcagt	ttattactgt	cagcaatatt	ttagctatcc	gctcacgttc	ggacaagggg	1980
ccaaggtgga	aataaaacgt	gagtggatcc	atctgggata	agcatgctgt	tttctgtctg	2040
tccctaacat	gccctgtgat	tatgcgcaaa	caacacaccc	aagggcagaa	ctttgttact	2100
taaacaccat	cctgttttgc	tctttcctca	ggaactgtgg	ctgcaccatc	tgtcttcac	2160
ttcccgcgat	ctgatgagca	gttgaaatct	ggaactgcct	ctgttgtgtg	cctgctgaat	2220
aacttctatc	ccagagaggc	caaagtacag	tggaaaggtg	ataacgcctt	ccaatcgggt	2280
aactcccagg	agagtgtcac	agagcaggac	agcaaggaca	gcacctacag	cctcagcagc	2340
accctgacgc	tgagcaaagc	agactacgag	aaacacaaag	tctacgcctg	cgaagtaccc	2400
catcaggggc	tgagctcgcc	cgtcacaaag	agcttcaaca	ggggagagt	ttagaggagg	2460
aagtgcccc	acctgctcct	cagttccagc	ctgacccctt	cccatccttt	ggcctctgac	2520
cctttttcca	caggggacct	acccctattg	cggtcctcca	gctcatcttt	cacctcacc	2580
ccctcctcct	ccttggcttt	aattatgcta	atgttggagg	agaatgaata	aataaagtga	2640
atctttgcac	ctgtggtgga	tctaataaaa	gatatttatt	ttcattagat	atgtgtgttg	2700
gttttttgtg	tgcagtgcct	ctatctggag	gccaggtagg	gctggccttg	ggggaggggg	2760
aggccagaat	gactccaaga	gctacaggaa	ggcaggtcag	agacccact	ggacaaacag	2820
tggctggact	ctgcaccata	acacacaatc	aacaggggag	tgagctggaa	atttgctagc	2880
gaattcttga	agacgaaaag	gcctcgtgat	acgcctattt	ttataggtta	atgtcatgat	2940
aataatgggt	tcttagacgt	caggtggcac	ttttcgggga	aatgtgcgcg	gaacccttat	3000
ttgtttat	ttctaaatac	attcaaatat	gtatccgctc	atgagacaat	aaccctgata	3060
aatgcttcaa	taatattgaa	aaaggaagag	tatgagtatt	caacatttcc	gtgtcgccct	3120
tattcccttt	tttgcgcat	tttgccctcc	tgtttttgct	caccagaaa	cgctggtgaa	3180
agtaaaagat	gctgaagatc	agttgggtgc	acgagtgggt	tacatcgaac	tggatctcaa	3240
cagcggtaa	atccttgaga	gttttcgccc	cgaagaacgt	tttccaatga	tgagcacttt	3300
taaagtctctg	ctatgtggcg	cggattatc	ccgtgttgac	gccgggcaag	agcaactcgg	3360
tcgccgata	cactattctc	agaatgactt	ggttgagtac	tcaccagtca	cagaaaagca	3420
tcttacggat	ggcatgacag	taagagaatt	atgcagtgtc	gccataacca	tgagtataa	3480
cactgcggcc	aacttacttc	tgacaacgat	cggaggaccg	aaggagctaa	ccgctttttt	3540
gcacaacatg	ggggatcatg	taactcgctt	tgatcggttg	gaaccggagc	tgaatgaagc	3600
cataccaaac	gacgagcgtg	acaccacgat	gcctgcagca	atggcaacaa	cgttgcgcaa	3660
actattaact	ggcgaactac	ttactctagc	ttcccggcaa	caattaatag	actggatgga	3720
ggcggataaa	gttgacagg	cacttctgcg	ctcgccctt	ccggctggct	ggtttattgc	3780
tgataaatct	ggagccggtg	agcgtgggtc	tcgcggtatc	attgcagcac	tggggccaga	3840
tggtaaagccc	tcccgatcgc	tagttatcta	cacgacgggg	agtcaggcaa	ctatggatga	3900
acgaaataga	cagatcgctg	agataggtgc	ctcactgatt	aagcatttgt	aactgtcaga	3960
ccaagtttac	tcatatatac	tttagattga	tttaaaactt	catttttaat	ttaaaaggat	4020
ctagggtgaag	atcctttttg	ataatctcat	gaccaaactc	ccttaacgtg	agttttcggt	4080
ccactgagcg	tcagaccccg	tagaaaagat	caaaggatct	tcttgagatc	ctttttttct	4140
gcgcgtaatc	tgctgcttgc	aaacaaaaaa	accaccgcta	ccagcgggtg	tttgtttgcc	4200
ggatcaagag	ctaccaactc	tttttccgaa	ggtaactggc	ttcagcagag	cgcagatacc	4260
aaatactgtc	cttctagtgt	agccgtagtt	aggccaccac	ttcaagaact	ctgtagcacc	4320
gcctacatac	ctcgctctgc	taatcctgtt	accagtggct	gctgccagt	gcgataagtc	4380
gtgtcttacc	gggttgga	caagacgata	gttaccggat	aaggcgcagc	ggtcgggctg	4440
aacggggggg	tcgtgcacac	agcccagctt	ggagcgaacg	acctacaccg	aactgagata	4500
cctcacagcg	gagctatgag	aaagcggcac	gcttcccga	gggagaaagg	cggacaggta	4560
tccggtaagc	ggcagggtcg	gaacaggaga	gcgcacgagg	gagcttccag	ggggaacgc	4620
ctgggtatctt	tatagtcctg	tcgggtttcg	ccacctctga	cttgagcgtc	gatttttgtg	4680
atgctcgta	ggggggcgga	gcctatggaa	aaacgccagc	aacgcggcct	ttttacgggt	4740
cctggccttt	tgctggcctt	ttgctcacat	gttctttcct	gcgttatccc	ctgattctgt	4800
ggataaccgt	attaccgctt	ttgagtga	tgataccgct	cgccgcagcc	gaacgaccga	4860
gcgcagcgag	tcagtga	aggaagcgga	agagcgctg	atgcggtatt	ttctccttac	4920
gcatctgtgc	ggtatttcac	accgcata	gtgcactctc	agtacaatct	gctctgatgc	4980
cgcatagtta	agccagtata	cactccgcta	tcgctacgtg	actgggtcat	ggctgcgccc	5040
cgacacccgc	caacacccgc	tgacgcgccc	tgacgggctt	gtctgctccc	ggcatccgct	5100
tacagacaag	ctgtgaccgt	ctccgggagc	tgcatgtgtc	agaggttttc	accgtcatca	5160
ccgaaacgcg	cgaggcagct	gtggaatgtg	tgctcagttg	ggtgtggaaa	gtccccaggc	5220
tccccagcag	gcagaagtat	gcaaagcatg	catctcaatt	agtcagcaac	caggctcccc	5280
agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	gcaaccatag	tcccgccctt	5340
aactccgccc	atcccgc	taactccg	cagttccg	cattctccgc	cccatggctg	5400



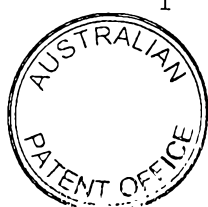
Leu Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln
 50 55 60
 Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg
 65 70 75 80
 Glu Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp
 85 90 95
 Phe Thr Leu Thr Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr
 100 105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

<210> 37
 <211> 372
 <212> DNA
 <213> Homo sapiens

<400> 37
 caggtgcaac tagtgcagtc cggcgccgaa gtgaagaaac ccggtgcttc cgtgaaagtc 60
 agctgtaaaa ctagtagata caccttcact gaatacacca tacactgggt tagacaggcc 120
 cctggccaaa ggctggagtg gataggaggt attaatccta acaatgggtat tcctaactac 180
 aaccagaagt tcaagggccg ggccaccttg accgtaggca agtctgccag caccgcctac 240
 atggaactgt ccagcctgcg ctccgaggac actgcagtct actactgcgc cagaagaaga 300
 atcgctatg gttacgacga gggccatgct atggactact ggggtcaagg aacccttgct 360
 accgtctcct ca 372

<210> 38
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 38
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15



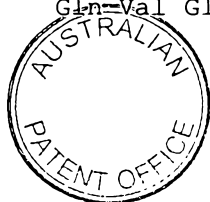
Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80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

<210> 39
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 39
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
 1 5 10 15
 Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe Thr Glu Tyr
 20 25 30
 Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
 35 40 45
 Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Phe Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

<210> 40
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 40
 Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala



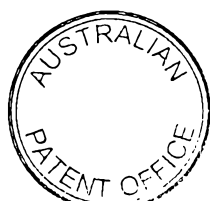
1	5	10	15
Ser Val Lys	Val Ser Cys Lys Thr	Ser Arg Tyr Thr Phe	Thr Glu Tyr
	20	25	30
Thr Ile His	Trp Val Arg Gln Ala	Pro Gly Gln Arg Leu	Glu Trp Ile
	35	40	45
Gly Gly Ile	Asn Pro Asn Asn Gly	Ile Pro Asn Tyr	Asn Gln Lys Phe
	50	55	60
Lys Gly Arg	Val Thr Ile Thr Val	Asp Thr Ser Ala	Ser Thr Ala Tyr
	65	70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

<210> 41
 <211> 124
 <212> PRT
 <213> Homo sapiens

<400> 41
Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
1 5 10 15
Ser Val Lys Val Ser Cys Lys Thr Ser Gly Tyr Thr Phe Thr Glu Tyr
20 25 30
Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu Glu Trp Ile
35 40 45
Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
50 55 60
Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
65 70 75 80
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
85 90 95
Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
100 105 110
Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
115 120

<210> 42
 <211> 7731
 <212> DNA
 <213> Homo sapiens

<400> 42



ttgaagacga	aagggcctcg	tgatacgct	atTTTTatag	gttaatgtca	tgataataat	60
ggTTTTcttag	acgtcaggtg	gcactTTTcg	gggaaatgtg	cgcggaaccc	ctatttgttt	120
atTTTTctaa	atacattcaa	atatgtatcc	gctcatgaga	caataaccct	gataaatgct	180
tcaataatat	tgaaaaagga	agagtatgag	tattcaacat	ttccgtgtcg	cccttattcc	240
ctTTTTtgcg	gcattTTTg	ttcctgTTTT	tgctcaccca	gaaacgctgg	tgaaagtaaa	300
agatgctgaa	gatcagttgg	gtgcacgagt	gggttacatc	gaactggatc	tcaacagcgg	360
taagatcctt	gagagTTTTc	gccccgaaga	acgtTTTTcca	atgatgagca	ctTTTaaagt	420
tctgctatgt	ggcgcggtat	tatcccgTgt	tgacgcggg	caagagcaac	tcggtcgccg	480
catacactat	tctcagaatg	acttggtTga	gtactcacca	gtcacagaaa	agcatcttac	540
ggatggcatg	acagtaagag	aattatgcag	tgctgccata	accatgagtg	ataaactgc	600
ggccaactta	cttctgacaa	cgatcggagg	accgaaggag	ctaaccgctt	TTTTgcacaa	660
catgggggat	catgtaactc	gccttgatcg	ttgggaaccg	gagctgaatg	aagccatacc	720
aaacgacgag	cgtgacacca	cgatgcctgc	agcaatggca	acaacgtTgc	gcaaactatt	780
aactggcgaa	ctacttactc	tagcttcccg	gcaacaatta	atagactgga	tgaggcgga	840
taaagtTgca	ggaccacttc	tgcgctcggc	ccttccggct	ggctggTTta	ttgctgataa	900
atctggagcc	ggtgagcgtg	ggtctcgcg	tatcattTga	gcactggggc	cagatggtaa	960
gccctcccg	atcgtagtta	tctacacgac	ggggagtcag	gcaactatgg	atgaacgaaa	1020
tagacagatc	gctgagatag	gtgcctcact	gattaagcat	tggttaactgt	cagaccaagt	1080
ttactcatat	atactTTTaga	ttgattTaaa	acttcatttt	taattTaaaa	ggatctaggt	1140
gaagatcctt	TTtgataatc	tcatgaccaa	aatcccttaa	cgtgagTTTT	cgttccactg	1200
agcgtcagac	cccgtagaaa	agatcaaagg	atcttctTga	gatcctTTTT	ttctgcgcgt	1260
aatctgctgc	ttgcaaacaa	aaaaaccacc	gctaccagcg	gtggtTtTgt	tgccggatca	1320
agagctacca	actctTTTTc	cgaaggtaac	tggttccagc	agagcgcaga	taccaaatac	1380
tgctctteta	gtgtagccgt	agttaggcca	ccacttcaag	aactctgtag	caccgcctac	1440
atacctcgct	ctgctaatec	tgttaccagt	ggctgctgcc	agtggcgata	agtcgtgtct	1500
taccgggttg	gactcaagac	gatagttacc	ggataaggcg	cagcggtcgg	gctgaacgga	1560
gggttcgtgc	acacagccca	gcttggagcg	aacgacctac	accgaactga	gatactaca	1620
gcgtgagcta	tgagaaagcg	ccacgctTcc	cgaagggaga	aaggcggaca	ggtatccggt	1680
aagcggcagg	gtcggaaacag	gagagcgcac	gagggagctt	ccagggggaa	acgcctggta	1740
tctttatagt	cctgtcggtg	ttcgccacct	ctgacttgag	cgtcgatttt	tgtgatgctc	1800
gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	gcctTTTTac	ggttcctggc	1860
cttttgctgg	ccttttgctc	acatgttctt	tcctgcgtta	tcctctgatt	ctgtggataa	1920
ccgtattacc	gcctttgagt	gagctgatac	cgctcgccgc	agccgaacga	ccgagcgcag	1980
cgagtcagtg	agcgaggaag	cggaagagcg	cctgatgcgg	tattttctcc	ttacgcactc	2040
gtgcggtatt	tcacaccgca	tatggtgcac	tctcagtaca	atctgctctg	atgccgcata	2100
gttaagccag	tatacactcc	gctatcgcta	cgtgactggg	tcatggctgc	gccccgacac	2160
ccgccaaacac	ccgtgacgc	gccctgacgg	gcttgtctgc	tcctggcatc	cgcttacaga	2220
caagctgtga	ccgtctccgg	gagctgcatg	tgtcagaggt	tttcaccgctc	atcaccgaaa	2280
cgcgcgaggc	agcatgcate	tcaattagtc	agcaaccata	gtcccgcctc	taactccgcc	2340
catcccgccc	ctaactccgc	ccagttccgc	ccattctccg	ccccatggct	gactaatttt	2400
ttttatttat	gcagaggccg	aggccgcctc	ggcctctgag	ctattccaga	agtagtgagg	2460
aggctTTTTt	ggaggcctag	gcttttgcaa	aaagctagct	tacagctcag	ggctgcgatt	2520
tcgcgccaaa	cttgacggca	atcctagcgt	gaaggctgg	aggattttat	ccccgctgcc	2580
atcatggttc	gaccattgaa	ctgcctcgtc	gccgtgtccc	aaaatatggg	gattggcaag	2640
aacggagacc	taccctggcc	tcgctcagg	aacgagttca	agtacttcca	aagaatgacc	2700
acaacctctt	cagtggaaag	taaacagaat	ctggtgatta	tgggtaggaa	aacctggctc	2760
tccattcctg	agaagaatcg	acctTTaaag	gacagaatta	atatagttct	cagtagagaa	2820
ctcaaagaac	caccacgagg	agctcatttt	cttgccaaaa	gtttggatga	tgctttaaga	2880
cttattgaac	aaccggaatt	ggcaagtaaa	gtagacatgg	tttgatagat	cggaggcagt	2940
tctgtttacc	aggaagccat	gaatcaacca	ggccacctca	gactctttgt	gacaaggatc	3000
atgcaggaat	ttgaaagtga	cacgtTTTTc	ccagaaattg	atttggggaa	atataaactt	3060
ctcccagaat	accagggcgt	cctctctgag	gtccaggagg	aaaaaggcat	caagtataag	3120
tttgaagtct	acgagaagaa	agactaacag	gaagatgctt	tcaagttctc	tgtccctc	3180
ctaaagctat	gcattTTTTat	aagaccatgg	gacttttgct	ggctttagat	ctttgtgaag	3240
gaaccttact	tctgtggtgt	gacataattg	gacaaactac	ctacagagat	ttaaagctct	3300
aaggtaaata	taaaattttt	aagtgtataa	tgtgttaaac	tactgattct	aattgtttgt	3360
gtatttttaga	ttccaaccta	tggaaactgat	gaatgggagc	agtgtgggaa	tgcttttaat	3420
gaggaaaacc	tgTTTTgctc	agaagaaatg	ccatctagtg	atgatgaggc	tactgctgac	3480
tctcaacatt	ctactcctcc	aaaaaagaag	agaaaggtag	aagaccccaa	ggactttcct	3540
tcagaattgc	taagtTTTTt	gagtcatgct	gtgttttagta	atagaactct	tgcttgcttt	3600
gctattttaca	ccacaaagga	aaaagctgca	ctgctatata	agaaaattat	ggaaaaatat	3660



tctgtaacct	ttataagtag	gcataacagt	tataatcata	acataactgtt	ttttcttact	3720
ccacacaggc	atagagtgtc	tgctattaat	aactatgctc	aaaaattgtg	taccttttagc	3780
tttttaattt	gtaaaggggt	taataaggaa	tatttgatgt	atagtgcctt	gactagagat	3840
cataatcagc	cataccacat	ttgtagaggt	tttacttgct	ttaaaaaacc	tcccacacct	3900
ccccctgaac	ctgaaacata	aaatgaatgc	aattggtggt	gttaacttgt	ttattgcagc	3960
ttataatggt	tacaaataaa	gcaatagcat	cacaaatttc	acaaataaag	catttttttc	4020
actgcattct	agttgtggtt	tgtccaaact	catcaatgta	tcttatcatg	tctggatcta	4080
ataaaagata	tttattttca	ttagatatgt	gtgttggttt	tttgtgtgca	gtgcctctat	4140
ctggaggcca	ggtagggctg	gccttggggg	agggggaggc	cagaatgact	ccaagagcta	4200
caggaaggca	ggtcagagac	cccactggac	aaacagtggc	tggactctgc	accataacac	4260
acaatcaaca	ggggagttag	ctggaaaatt	gctagcgaat	tccagcacac	tggcgccgt	4320
tactagtatt	taatagtaat	caattacggg	gtcattagtt	catagcccat	atatggagtt	4380
ccgcgttaca	taacttacgg	taaatggccc	gcctggctga	ccgccaacg	acccccgcc	4440
attgacgtca	ataatgacgt	atgttcccat	agtaacgcca	atagggactt	tccattgacg	4500
tcaatgggtg	gagtatttac	ggtaaactgc	ccacttgcca	gtacatcaag	tgtatcatat	4560
gccaaagtac	ccccctattg	acgtcaatga	cggtaaattg	cccgcctggc	attatgcca	4620
gtacatgacc	ttatgggact	ttcctacttg	gcagtacatc	tacgtattag	tcatcgctat	4680
taccatgggt	atgcgggttt	ggcagtagac	caatgggcgt	ggatagcgg	ttgactcacg	4740
gggatttcca	agtctccacc	ccattgacgt	caatgggagt	ttgttttggc	acaaaaatca	4800
acgggacttt	ccaaaatgtc	gtaacaactc	cgccccattg	acgcaaattg	gcggtagggc	4860
tgtacggtgg	gaggtctata	taagcagagc	tcgttttagt	aaccgtcaga	tcgcctggag	4920
acgccatcca	cgtgtttttg	acctccatag	aagacaccgg	gaccgatcca	gcctccgcgg	4980
ccgggaacgg	tgcattggaa	cgcggattcc	ccgtgccaag	agtgcagtaa	gtaccgcta	5040
tagagtctat	aggcccaccc	ccttggtctc	ttatgcatgc	tatactgttt	ttggcttggg	5100
gtctatacac	ccccgcttcc	tcatgttata	ggtagtggtg	tagcttagcc	tataggtgtg	5160
ggttattgac	cattattgac	cactccccta	ttggtgacga	tactttccat	tactaatcca	5220
taacatggct	ctttgccaca	actctcttta	ttggtctatat	gccaaatacac	tgtccttcag	5280
agactgacac	ggactctgta	tttttacagg	atggggtctc	atttattatt	tacaaattca	5340
catatacaac	accaccgtcc	ccagtgcctg	cagtttttat	taaacataac	gtgggatctc	5400
caecggaatc	tcgggtacgt	gttccggaca	tgggctcttc	tccggtagcg	gcggagcttc	5460
tacatccgag	ccctgctccc	atgcctccag	cgactcatgg	tcgctcggca	gctccttget	5520
cctaacagtg	gaggccagac	ttaggcacag	cacgatgccc	accaccacca	gtgtgccgca	5580
caaggccgtg	gcggtaggg	atgtgtctga	aaatgagctc	ggggagcggg	cttgaccgc	5640
tgacgcattt	ggaagactta	aggcagcggc	agaagaagat	gcaggcagct	gagttgttgt	5700
gttctgataa	gagtcagagg	taactccgt	tgcggtgctg	ttaacggtgg	agggcagtgt	5760
agtctgagca	gtactcgttg	ctgccgcgcg	cgccaccaga	cataatagct	gacagactaa	5820
cagactgttc	ctttccatgg	gtcttttctg	cagtcaccgt	ccttgacacg	cgtctcggga	5880
agcttgccgc	caccatggac	tggacctggc	gcgtgttttg	cctgctcgcc	gtggctcctg	5940
gggcccacag	ccaggtgcaa	ctggtgcagt	ccggcgccga	agtgaagaaa	cccgtgctt	6000
ccgtgaaagt	cagctgtaaa	actagtagat	acaccttcac	tgaatacacc	atacactggg	6060
ttagacaggc	ccctggccaa	aggctggagt	ggataggagg	tattaatcct	aacaatggta	6120
ttcctaacta	caaccagaag	ttcaagggcc	gggccacctt	gaccgtaggc	aagtctgcca	6180
gcaccgccta	catggaactg	tccagcctgc	gctccgagga	cactgcagtc	tactactgcg	6240
ccagaagaag	aatcgcctat	ggttacgacg	agggccatgc	tatggactac	tggggtcaag	6300
gaaccccttg	cacogtctcc	tcaggtagt	ggatcctctg	cgcttgggcc	cagctctgtc	6360
ccacaccggc	gtcacatggc	accacctctc	ttgcagcctc	caccaagggc	ccatcggtct	6420
tccccctggc	accctcctcc	aagagcacct	ctgggggcac	agcggccctg	ggctgcctgg	6480
tcaaggacta	cttccccgaa	ccggtgacgg	tgtcgtggaa	ctcaggcgcc	ctgaccagcg	6540
gcgtgcacac	cttcccggt	gtcctacagt	cctcaggact	ctactccctc	agcagcgtgg	6600
tgaccgtgcc	ctccagcagc	ttgggcaccc	agacctacat	ctgcaacgtg	aatcacaagc	6660
ccagcaacac	caaggtggac	aagaaagtgt	agcccaaatc	ttgtgacaaa	actcacacat	6720
gcccaccgtg	cccagcacct	gaactcctgg	ggggaccgtc	agtcttcttc	ttcccccaa	6780
aacccaagga	caccctcatg	atctcccgga	cccctgaggt	cacatgcgtg	gtgggtggacg	6840
tgagccacga	agaccctgag	gtcaagttca	actggtacgt	ggacggcgtg	gaggtgcata	6900
atgccaaagc	aaagccgcgg	gaggagcagt	acaacagcac	gtaccgggtg	gtcagcgtcc	6960
tcaccgtcct	gcaccaggac	tggctgaatg	gcaaggagta	caagtgcaag	gtctccaaca	7020
aagccctccc	agcccccatc	gagaaaacca	tctccaaagc	caaagggcag	ccccgagaac	7080
cacaggtgta	caccctgccc	ccatcccggg	aggagatgac	caagaaccag	gtcagcctga	7140
cctgcctggt	caaaggcttc	tatcccagcg	acatcgccgt	ggagtgggag	agcaatgggc	7200
agccggagaa	caactacaag	accacgcctc	ccgtgctgga	ctccgacggc	tccttcttcc	7260
tctacagcaa	gctcaccgtg	gacaagagca	ggtggcagca	ggggaacgtc	ttctcatgct	7320



```

ccgtgatgca tgaggctctg cacaaccact acacgcagaa gagcctctcc ctgtctccgg 7380
gtaaatgagt gcgacggccg gcaagccccg ctccccgggc tctcgcggtc gcacgaggat 7440
gcttggcacg taccacctgt acatacttcc cgggcgcccc gcatggaaat aaagcaccgg 7500
atctaataaa agatatttat ttccattaga tatgtgtgtt gggtttttgt gtgcagtgcc 7560
tctatctgga ggccaggtag ggctggcctt gggggagggg gaggccagaa tgactccaag 7620
agctacagga aggcaggtag gagacccac tggacaaaca gtggctggac tctgcaccat 7680
aacacacaat caacagggga gtgagctgga aatttgctag cgaattaatt c 7731

```

<210> 43

<211> 472

<212> PRT

<213> Homo sapiens

<400> 43

```

Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala Pro Gly
  1              5              10              15

```

```

Ala His Ser Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys
          20              25              30

```

```

Pro Gly Ala Ser Val Lys Val Ser Cys Lys Thr Ser Arg Tyr Thr Phe
          35              40              45

```

```

Thr Glu Tyr Thr Ile His Trp Val Arg Gln Ala Pro Gly Gln Arg Leu
  50              55              60

```

```

Glu Trp Ile Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn
  65              70              75              80

```

```

Gln Lys Phe Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ala Ser
          85              90              95

```

```

Thr Ala Tyr Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val
 100              105              110

```

```

Tyr Tyr Cys Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His
 115              120              125

```

```

Ala Met Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Ser
 130              135              140

```

```

Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys Ser Thr
 145              150              155              160

```

```

Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr Phe Pro
 165              170              175

```

```

Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu Thr Ser Gly Val
 180              185              190

```

```

His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu Tyr Ser Leu Ser
 195              200              205

```

```

Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile
 210              215              220

```

```

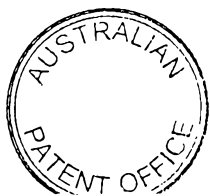
Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val
 225              230              235              240

```

```

Glu Pro Lys Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala

```



245										250					255				
Pro	Glu	Leu	Leu	Gly	Gly	Pro	Ser	Val	Phe	Leu	Phe	Pro	Pro	Lys	Pro				
			260					265					270						
Lys	Asp	Thr	Leu	Met	Ile	Ser	Arg	Thr	Pro	Glu	Val	Thr	Cys	Val	Val				
		275					280					285							
Val	Asp	Val	Ser	His	Glu	Asp	Pro	Glu	Val	Lys	Phe	Asn	Trp	Tyr	Val				
	290					295					300								
Asp	Gly	Val	Glu	Val	His	Asn	Ala	Lys	Thr	Lys	Pro	Arg	Glu	Glu	Gln				
305					310					315					320				
Tyr	Asn	Ser	Thr	Tyr	Arg	Val	Val	Ser	Val	Leu	Thr	Val	Leu	His	Gln				
				325					330					335					
Asp	Trp	Leu	Asn	Gly	Lys	Glu	Tyr	Lys	Cys	Lys	Val	Ser	Asn	Lys	Ala				
			340					345					350						
Leu	Pro	Ala	Pro	Ile	Glu	Lys	Thr	Ile	Ser	Lys	Ala	Lys	Gly	Gln	Pro				
		355					360						365						
Arg	Glu	Pro	Gln	Val	Tyr	Thr	Leu	Pro	Pro	Ser	Arg	Glu	Glu	Met	Thr				
	370					375					380								
Lys	Asn	Gln	Val	Ser	Leu	Thr	Cys	Leu	Val	Lys	Gly	Phe	Tyr	Pro	Ser				
385					390					395				400					
Asp	Ile	Ala	Val	Glu	Trp	Glu	Ser	Asn	Gly	Gln	Pro	Glu	Asn	Asn	Tyr				
			405					410						415					
Lys	Thr	Thr	Pro	Pro	Val	Leu	Asp	Ser	Asp	Gly	Ser	Phe	Phe	Leu	Tyr				
			420					425					430						
Ser	Lys	Leu	Thr	Val	Asp	Lys	Ser	Arg	Trp	Gln	Gln	Gly	Asn	Val	Phe				
		435					440					445							
Ser	Cys	Ser	Val	Met	His	Glu	Ala	Leu	His	Asn	His	Tyr	Thr	Gln	Lys				
	450					455					460								
Ser	Leu	Ser	Leu	Ser	Pro	Gly	Lys												
465					470														

<210> 44
 <211> 25
 <212> DNA
 <213> Homo sapiens

<400> 44
 accgtctcct caggtgagtg gatcc

25

<210> 45
 <211> 14
 <212> DNA
 <213> Homo sapiens

<400> 45
 cctctcttgc agcc

14



<210> 46
 <211> 14
 <212> DNA
 <213> Homo sapiens

<400> 46
 cctctcttgc agcc

14

<210> 47
 <211> 4
 <212> PRT
 <213> Homo sapiens

<400> 47
 Thr Val Ser Ser
 1

<210> 48
 <211> 4
 <212> PRT
 <213> Homo sapiens

<400> 48
 Ser Thr Lys Gly
 1

<210> 49
 <211> 27
 <212> DNA
 <213> Homo sapiens

<400> 49
 accgtctcct cagcctccac caagggc

27

<210> 50
 <211> 8
 <212> PRT
 <213> Homo sapiens

<400> 50
 Thr Val Ser Ser Ser Thr Lys Gly
 1 5

<210> 51
 <211> 27
 <212> DNA
 <213> Homo sapiens

<400> 51
 accgtctcct cagcctccac caagggc

27

<210> 52
 <211> 9



<212> PRT
 <213> Homo sapiens

<400> 52
 Thr Val Ser Ser Ala Ser Thr Lys Gly
 1 5

<210> 53
 <211> 22
 <212> DNA
 <213> Homo sapiens

<400> 53
 gaaataaaaac gtgagtggat cc 22

<210> 54
 <211> 27
 <212> DNA
 <213> Homo sapiens

<400> 54
 cttcttttctt caggaactgt ggctgca 27

<210> 55
 <211> 4
 <212> PRT
 <213> Homo sapiens

<400> 55
 Thr Val Ala Ala
 1

<210> 56
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 56
 gaaataaaaac gaactgtggc tgca 24

<210> 57
 <211> 7
 <212> PRT
 <213> Homo sapiens

<400> 57
 Glu Ile Lys Thr Val Ala Ala
 1 5

<210> 58
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 58



gaaataaaac gaactgtggc tgca

24

<210> 59

<211> 8

<212> PRT

<213> Homo sapiens

<400> 59

Glu Ile Lys Arg Thr Val Ala Ala

1

5

<210> 60

<211> 20

<212> PRT

<213> Homo sapiens

<400> 60

Met Asp Ser Gln Ala Gln Val Leu Met Leu Leu Leu Trp Val Ser

1

5

10

15

Gly Thr Cys Gly

20

<210> 61

<211> 19

<212> PRT

<213> Homo sapiens

<400> 61

Met Gly Trp Ser Trp Val Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly

1

5

10

15

Val Leu Ser

<210> 62

<211> 9

<212> DNA

<213> Homo sapiens

<400> 62

gccgccacc

9

<210> 63

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:PRIMER

<400> 63

cagaaagctt gccgccacca tggattcaca ggcccag

37

<210> 64



<211> 6
<212> PRT
<213> Homo sapiens

<400> 64
Met Asp Ser Gln Ala Gln
1 5

<210> 65
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 65
ccgaggatcc actcacgttt cagctccagc ttggt 35

<210> 66
<211> 37
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 66
cagaaagctt gccgccacca tgggatggag ctgggtc 37

<210> 67
<211> 6
<212> PRT
<213> Homo sapiens

<400> 67
Met Gly Trp Ser Trp Val
1 5

<210> 68
<211> 35
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 68
ccgaggatcc actcacctga ggagacggtg actga 35

<210> 69
<211> 36
<212> DNA
<213> Artificial Sequence

<220>



<223> Description of Artificial Sequence:primer

<400> 69

gtcatcaciaa tgtctccgga ggaacctgga acccag

36

<210> 70

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 70

ctccggagac attgtgatga cccaatctc

29

<210> 71

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 71

gaatataaaa ggctctgact ggacttgacag ttgatgggtg ccctc

45

<210> 72

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 72

cagtcagagc cttttatatt ctagaatca aaagaactac ttggcctggt atcagcagaa 60
accaggacag cc 72

<210> 73

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 73

accccagatt ccctagtgt agcccaaaag atgaggagtt tggg

44

<210> 74

<211> 67

<212> DNA

<213> Artificial Sequence

<220>



<223> Description of Artificial Sequence:primer

<400> 74

tagcactagg gaatctgggg tacctgatag gttcagtggc agtggggttg ggacagactt 60
caccctc 67

<210> 75

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 75

gtcccttgtc cgaacgtgag cggatagcta aaatattgct gacagtaata aac 53

<210> 76

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 76

gtcacgttc ggacaaggga ccaaggtgga aat 33

<210> 77

<211> 72

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 77

cagtcagagc cttttatatt ctagaaatca aaagaactac ttggcctggt tccagcagaa 60
accaggacag cc 72

<210> 78

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 78

gtcccttgtc cgaacgtgag cggatagcta aaatattgct gacagtcata aactgcc 57

<210> 79

<211> 34

<212> DNA

<213> Artificial Sequence



<220>
<223> Description of Artificial Sequence:primer

<400> 79
cccaaactcc tcattctattg ggctagcact aggg 34

<210> 80
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 80
ccctagtgt agcccaatag atgaggagtt tggg 34

<210> 81
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 81
tacgcaaacc gcctctc 17

<210> 82
<211> 18
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

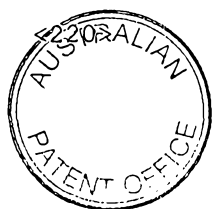
<400> 82
gagtgcacca tatgcggt 18

<210> 83
<211> 16
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:primer

<400> 83
aacagctatg accatg 16

<210> 84
<211> 17
<212> DNA
<213> Artificial Sequence



<223> Description of Artificial Sequence:primer

<400> 84

gttttcccag tcacgac

17

<210> 85

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 85

gtgtattcag tgaagggtga tctactagtt ttacagctga ctttcac

47

<210> 86

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 86

tagtagatatac accttcactg aatacaccat aactgggtt agacaggccc ctg

53

<210> 87

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 87

cccttgaact tctggttgta gtttaggaata ccattgtag gattaatacc tcctatccac 60
tccagccttt g 71

<210> 88

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 88

taactacaac cagaagttca agggccgggc caccttgacc gtaggcaagt ctgccagcac 60
cgctacatg g 71

<210> 89

<211> 63

<212> DNA

<213> Artificial Sequence



<220>

<223> Description of Artificial Sequence:primer

<400> 89

gcatggccct cgtcgtaacc ataggcgatt cttcttctgg cgcagtagta gactgcagtg 60
tcc 63

<210> 90

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 90

ctatggttac gacgagggcc atgctatgga ctactggggt caaggaac 48

<210> 91

<211> 71

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 91

taactacaac cagaagttca agggccgggt caccatcacc gtagacacct ctgccagcac 60
cgctacatg g 71

<210> 92

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 92

ggacactgca gtctacttct gcgccag 27

<210> 93

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 93

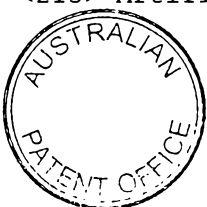
tacgcaaacc gcctctc 17

<210> 94

<211> 18

<212> DNA

<213> Artificial Sequence



<220>

<223> Description of Artificial Sequence:primer

<400> 94

gagtgacacca tatgcggt

18

<210> 95

<211> 75

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 95

cctttggcca ggggcctgtc taaccacgtg tatgggtgat tcagtgaagg tgtatccact 60
agtttccact agttt 75

<210> 96

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 96

gtcaccgtcc ttgacacgcg tctcggga

28

<210> 97

<211> 17

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 97

ttggaggagg gtgccag

17

<210> 98

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 98

gagacattgt gacccaatct cc

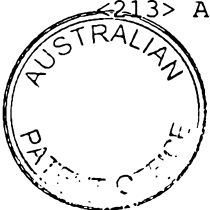
22

<210> 99

<211> 25

<212> DNA

<213> Artificial Sequence



<220>

<223> Description of Artificial Sequence:primer

<400> 99

gacagtcata aactgccaca tcttc

25

<210> 100

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 100

ttgacacgcg tctcgggaag ctt

23

<210> 101

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:primer

<400> 101

ggcgcagagg atccactcac ct

22



EDITORIAL NOTE - NO 40322/99

Sequence listing pages 1-64 is part of the description.

Claim pages are to follow.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: Boehringer Ingelheim International GmbH
- (B) STREET: Rheinstrasse
- (C) CITY: Ingelheim am Rhein
- (E) COUNTRY: Germany
- (F) POSTAL CODE (ZIP): 55216
- (G) TELEPHONE: ++49-6132-772770
- (H) TELEFAX: ++49-6132-774377

(ii) TITLE OF INVENTION: FAP alpha-specific antibody with improved producibility

(iii) NUMBER OF SEQUENCES: 101

(iv) 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 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GACATTGTGA TGACCCAATC TCCAGACTCT TTGGCTGTGT CTCTAGGGGA GAGGGCCACC	60
ATCAACTGCA AGTCCAGTCA GAGCCTTTTA TATTCTAGAA ATCAAAAGAA CTA CTTGGCC	120
TGGTATCAGC AGAAACCAGG ACAGCCACCC AAACCTCTCA TCTTTTGGGC TAGCACTAGG	180
GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT CACCCTCACC	240
ATTAGCAGCC TGCAGGCTGA AGATGTGGCA GTTTATTACT GTCAGCAATA TTTTAGCTAT	300
CCGCTCACGT TCGGACAAGG GACCAAGGTG GAAATAAAA	339

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

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

```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 339 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

GACATTGTGA TGACCCAATC TCCAGACTCT TTGGCTGTGT CTCTAGGGGA GAGGGCCACC      60
ATCAACTGCA AGTCCAGTCA GAGCCTTTTA TATTCTAGAA ATCAAAAGAA CTACTTGGCC      120
TGGTTCCAGC AGAAACCAGG ACAGCCACCC AACTCCTCA TCTTTTGGGC TAGCACTAGG      180
GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT CACCCTCACC      240
ATTAGCAGCC TGCAGGCTGA AGATGTGGCA GTTTATGACT GTCAACAATA TTTTAGCTAT      300
CCGCTCACGT TCGGACAAGG GACCAAGGTG GAAATAAAA                                339

```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1           5           10           15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
          20           25           30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
          35           40           45

Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
          50           55           60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
65           70           75           80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
          85           90           95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
          100          105          110

Lys

```

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 339 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

GACATTGTGA TGACCCAATC TCCAGACTCT TTGGCTGTGT CTCTAGGGGA GAGGGCCACC      60
ATCAACTGCA AGTCCAGTCA GAGCCTTTTA TATTCTAGAA ATCAAAAGAA CTA CT TGGCC      120
TGGTATCAGC AGAAACCAGG ACAGCCACCC AAACTCCTCA TCTATTGGGC TAGCACTAGG      180
GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT CACCCTCACC      240
ATTAGCAGCC TGCAGGCTGA AGATGTGGCA GTTTATTACT GTCAGCAATA TTTTAGCTAT      300

```

CCGCTCACGT TCGGACAAGG GACCAAGGTG GAAATAAAA

339

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
 1 5 10 15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
 20 25 30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
 35 40 45

Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
 50 55 60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
 65 70 75 80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
 85 90 95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
 100 105 110

Lys

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGGTGCAAC TAGTGCAAGT CCGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC 60

AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TAACTGGGT TAGACAGGCC 120

CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC 180

AACCAGAAAGT TCAAGGGCCG GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC 240
 ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA 300
 ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTGTC 360
 ACCGTCTCCT CA 372

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC	60
AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACTCTGGGT TAGACAGGCC	120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC	180
AACCAGAAGT TCAAGGGCCG GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC	240
ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA	300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTTGTC	360
ACCGTCTCCT CA	372

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGGTGCAAC	TAGTGCAGTC	CGGCGCCGAA	GTGAAGAAAC	CCGGTGCTTC	CGTGAAAGTC	60
AGCTGTAAAA	CTAGTAGATA	CACCTTCACT	GAATACACCA	TACACTGGGT	TAGACAGGCC	120
CCTGGCCAAA	GGCTGGAGTG	GATAGGAGGT	ATTAATCCTA	ACAATGGTAT	TCCTAACTAC	180
AACCAGAAGT	TCAAGGGCCG	GGTCACCATC	ACCGTAGACA	CCTCTGCCAG	CACCGCCTAC	240
ATGGAAGTGT	CCAGCCTGCG	CTCCGAGGAC	ACTGCAGTCT	ACTACTGCGC	CAGAAGAAGA	300
ATCGCCTATG	GTTACGACGA	GGGCCATGCT	ATGGACTACT	GGGGTCAAGG	AACCCCTGTGTC	360
ACCGTCTCCT	CA					372

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 124 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC      60
AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TAACTGGGT TAGACAGGCC      120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC      180
AACCAGAAGT TCAAGGGCCG GGTCACCATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC      240
ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA      300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCCCTGTC      360
ACCGTCTCCT CA                                     372

```

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

```

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

```

85					90					95					
Ala	Arg	Arg	Arg	Ile	Ala	Tyr	Gly	Tyr	Asp	Glu	Gly	His	Ala	Met	Asp
100					105					110					
Tyr	Trp	Gly	Gln	Gly	Thr	Leu	Val	Thr	Val	Ser	Ser				
115					120										

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGGTGCAAC	TAGTGCAAGTC	CGGCGCCGAA	GTGAAGAAAC	CCGGTGCTTC	CGTGAAAGTC	60
AGCTGTAAAA	CTAGTGGAATA	CACCTTCACT	GAATACACCA	TACACTGGGT	TAGACAGGCC	120
CCTGGCCAAA	GGCTGGAGTG	GATAGGAGGT	ATTAATCCTA	ACAATGGTAT	TCCTAACTAC	180
AACCAGAAGT	TCAAGGGCCG	GGTCACCATC	ACCGTAGACA	CCTCTGCCAG	CACCGCCTAC	240
ATGGAAGTGT	CCAGCCTGCG	CTCCGAGGAC	ACTGCAGTCT	ACTACTGCGC	CAGAAGAAGA	300
ATCGCCTATG	GTTACGACGA	GGGCCATGCT	ATGGACTACT	GGGGTCAAGG	AACCCTTGTC	360
ACCGTCTCCT	CA					372

(2) INFORMATION FOR SEO ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

				165						170					175				
Ser	Thr	Tyr	Ser	Leu	Ser	Ser	Thr	Leu	Thr	Leu	Ser	Lys	Ala	Asp	Tyr				
			180					185					190						
Glu	Lys	His	Lys	Val	Tyr	Ala	Cys	Glu	Val	Thr	His	Gln	Gly	Leu	Ser				
		195					200					205							
Ser	Pro	Val	Thr	Lys	Ser	Phe	Asn	Arg	Gly	Glu	Cys								
	210					215					220								

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 453 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

```

Thr Val Pro Ser Ser Ser Leu Gly Thr Gln Thr Tyr Ile Cys Asn Val
 195                                200                                205

Asn His Lys Pro Ser Asn Thr Lys Val Asp Lys Lys Val Glu Pro Lys
 210                                215                                220

Ser Cys Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala Pro Glu Leu
 225                                230                                235                                240

Leu Gly Gly Pro Ser Val Phe Leu Phe Pro Pro Lys Pro Lys Asp Thr
                                245                                250                                255

Leu Met Ile Ser Arg Thr Pro Glu Val Thr Cys Val Val Val Asp Val
                                260                                265                                270

Ser His Glu Asp Pro Glu Val Lys Phe Asn Trp Tyr Val Asp Gly Val
 275                                280                                285

Glu Val His Asn Ala Lys Thr Lys Pro Arg Glu Glu Gln Tyr Asn Ser
 290                                295                                300

Thr Tyr Arg Val Val Ser Val Leu Thr Val Leu His Gln Asp Trp Leu
 305                                310                                315                                320

Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala Leu Pro Ala
                                325                                330                                335

Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro Arg Glu Pro
                                340                                345                                350

Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr Lys Asn Gln
 355                                360                                365

Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser Asp Ile Ala
 370                                375                                380

Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr Lys Thr Thr
 385                                390                                395                                400

Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr Ser Lys Leu
                                405                                410                                415

Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe Ser Cys Ser
                                420                                425                                430

Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys Ser Leu Ser
 435                                440                                445

Leu Ser Pro Gly Lys
 450

```

(2) INFORMATION FOR SEQ ID NO: 19:

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

- (ii) MOLECULE TYPE: cDNA

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

```

CGTACTGTGG CTGCACCATC TGTCTTCATC TTCCCGCCAT CTGATGAGCA GTTGAAATCT      60
GGAAGTGCCT CTGTTGTGTG CCTGCTGAAT AACTTCTATC CCAGAGAGGC CAAAGTACAG      120
TGGAAGGTGG ATAACGCCCT CCAATCGGGT AACTCCCAGG AGAGTGTAC AGAGCAGGAC      180
AGCAAGGACA GCACCTACAG CCTCAGCAGC ACCCTGACGC TGAGCAAAGC AGACTACGAG      240
AAACACAAAG TCTACGCCTG CGAAGTCACC CATCAGGGCC TGAGCTCGCC CGTCACAAAG      300
AGCTTCAACA GGGGAGAGTG T                                             321

```

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: peptide

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

```

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

```

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 990 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

```

GCCTCCACCA AGGGCCCATC GGTCTTCCCC CTGGCACCCCT CCTCCAAGAG CACCTCTGGG      60
GGCACAGCGG CCCTGGGCTG CCTGGTCAAG GACTACTTCC CCGAACCGGT GACGGTGTCTG      120
TGGAATCAG GCGCCCTGAC CAGCGGCGTG CACACCTTCC CGGCTGTCCT ACAGTCCTCA      180
GGACTCTACT CCCTCAGCAG CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC      240
TACATCTGCA ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGTTGAGCCC      300
AAATCTTGTG ACAAAACTCA CACATGCCCA CCGTGCCCAG CACCTGAACT CCTGGGGGGA      360
CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC TCATGATCTC CCGGACCCCT      420
GAGGTCACAT GCGTGGTGGT GGACGTGAGC CACGAAGACC CTGAGGTCAA GTTCAACTGG      480
TACGTGGACG GCGTGGAGGT GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC      540
AGCACGTACC GGGTGGTCAG CGTCCTCACC GTCCTGCACC AGGACTGGCT GAATGGCAAG      600
GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCCAGCCC CCATCGAGAA AACCATCTCC      660
AAAGCCAAAG GGCAGCCCCG AGAACCACAG GTGTACACCC TGCCCCCATC CCGGGAGGAG      720
ATGACCAAGA ACCAGGTCAG CCTGACCTGC CTGGTCAAAG GCTTCTATCC CAGCGACATC      780
GCCGTGGAGT GGGAGAGCAA TGGGCAGCCG GAGAACAACCT ACAAGACCAC GCCTCCCGTG      840
CTGGACTCCG ACGGCTCCTT CTTCTCTAC AGCAAGCTCA CCGTGGACAA GAGCAGGTGG      900
CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA CCACTACACG      960
CAGAAGAGCC TCTCCCTGTC TCCGGGTAAA      990

```

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 330 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser Ser Lys
1           5           10          15
Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys Asp Tyr

```

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

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 427 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

```

AAGCTTGCCG CCACCATGGA TTCACAGGCC CAGGTTCTTA TGTTACTGCC GCTATGGGTA      60
TCTGGTACCT GTGGGGACAT TGTGATGTCA CAGTCTCCAT CCTCCCTAGC TGTGTCAGTT      120
GGAGAGAAGG TTACTATGAG CTGCAAGTCC AGTCAGAGCC TTTTATATAG TCGTAATCAA      180
AAGAACTACT TGGCCTGGTT CCAGCAGAAG CCAGGGCAGT CTCCTAAACT GCTGATTTTC      240
TGGGCATCCA CTAGGGAATC TGGGGTCCCT GATCGCTTCA CAGGCAGTGG ATTTGGGACG      300
GATTTCAATC TCACCATCAG CAGTGTGCAG GCTGAGGACC TGGCAGTTTA TGA CTGTCAG      360
CAATATTTTA GCTATCCGCT CACGTTCCGT GCTGGGACCA AGCTGGAGCT GAAACGTGAG      420
TGGATCC                                           427
  
```

(2) INFORMATION FOR SEQ ID NO: 24:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 133 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

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

Phe Asn Leu Thr Ile Ser Ser Val Gln Ala Glu Asp Leu Ala Val Tyr
 100 105 110

Asp Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Ala Gly Thr
 115 120 125

Lys Leu Glu Leu Lys
 130

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 457 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

AAGCTTGCCG CCACCATGGG ATGGAGCTGG GTCTTTCTCT TTCTCCTGTC AGGAACTGCA	60
GGTGTCTCT CTGAGGTCCA GCTGCAACAG TCTGGACCTG AGCTGGTGAA GCCTGGGGCT	120
TCAGTAAAGA TGTCTTGCAA GACTTCTAGA TACACATTCA CTGAATACAC CATACTG	180
GTGAGACAGA GCCATGGAAA GAGCCTTGAG TGGATTGGAG GTATTAATCC TAACAATGGT	240
ATTCTTAAT ACAACCAGAA GTTCAAGGGC AGGGCCACAT TGACTGTAGG CAAGTCCTCC	300
AGCACCGCCT ACATGGAGCT CCGCAGCCTG ACATCTGAGG ATTCTGCGGT CTATTTCTGT	360
GCAAGAAGAA GAATCGCCTA TGGTTACGAC GAGGGCCATG CTATGGACTA CTGGGGTCAA	420
GGAACCTCAG TCACCGTCTC CTCAGGTGAG TGGATCC	457

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 143 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Met Gly Trp Ser Trp Val Phe Leu Phe Leu Leu Ser Gly Thr Ala Gly	
1 5 10 15	
Val Leu Ser Glu Val Gln Leu Gln Gln Ser Gly Pro Glu Leu Val Lys	
20 25 30	

Pro Gly Ala Ser Val Lys Met Ser Cys Lys Thr Ser Arg Tyr Thr Phe
 35 40 45

Thr Glu Tyr Thr Ile His Trp Val Arg Gln Ser His Gly Lys Ser Leu
 50 55 60

Glu Trp Ile Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn
 65 70 75 80

Gln Lys Phe Lys Gly Arg Ala Thr Leu Thr Val Gly Lys Ser Ser Ser
 85 90 95

Thr Ala Tyr Met Glu Leu Arg Ser Leu Thr Ser Glu Asp Ser Ala Val
 100 105 110

Tyr Phe Cys Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His
 115 120 125

Ala Met Asp Tyr Trp Gly Gln Gly Thr Ser Val Thr Val Ser Ser
 130 135 140

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 8068 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

GAATTCCAGC AACTGGCGG CCGTTACTAG TTATTAATAG TAATCAATTA CGGGGTCATT	60
AGTTCATAGC CCATATATGG AGTTCGCGT TACATAACTT ACGGTAAATG GCCCGCCTGG	120
CTGACCGCCC AACGACCCCC GCCCATTGAC GTCAATAATG ACGTATGTTC CCATAGTAAC	180
GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCACTT	240
GGCAGTACAT CAAGTGATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA	300
ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA	360
CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG TTTTGGCAGT ACATCAATGG	420
GCGTGGATAG CGGTTTGACT CACGGGGATT TCCAAGTCTC CACCCCATTG ACGTCAATGG	480
GAGTTTGTTT TGGCACCAA ATCAACGGGA CTTTCCAAA TGTCGTAACA ACTCCGCCCC	540
ATTGACGCAA ATGGGCGGTA GCGGTGTACG GTGGGAGGTC TATATAAGCA GAGCTCGTTT	600
AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC ATAGAAGACA	660
CCGGGACCGA TCCAGCCTCC GCGGCCGGA ACGGTGCATT GGAACGCGGA TTCCCCGTGC	720

CAAGAGTGAC	GTAAGTACCG	CCTATAGAGT	CTATAGGCC	ACCCCTTGG	CTTCTTATGC	780
ATGCTATACT	GTTTTTGGCT	TGGGGTCTAT	ACACCCCGC	TTCCTCATGT	TATAGGTGAT	840
GGTATAGCTT	AGCCTATAGG	TGTGGGTAT	TGACCATTAT	TGACCACTCC	CCTATTGGTG	900
ACGATACTTT	CCATTACTAA	TCCATAACAT	GGCTCTTTGC	CACAACTCTC	TTTATTGGCT	960
ATATGCCAAT	ACACTGTCCT	TCAGAGACTG	ACACGGACTC	TGTATTTTTA	CAGGATGGGG	1020
TCTCATTTAT	TATTTACAAA	TTCACATATA	CAACACCACC	GTCCCAGTG	CCCGCAGTTT	1080
TTATTAAACA	TAACGTGGGA	TCTCCACGCG	AATCTCGGGT	ACGTGTTCCG	GACATGGGCT	1140
CTTCTCCGGT	AGCGGCGGAG	CTTCTACATC	CGAGCCCTGC	TCCCATGCCT	CCAGCGACTC	1200
ATGGTCGCTC	GGCAGCTCCT	TGCTCCTAAC	AGTGGAGGCC	AGACTTAGGC	ACAGCACGAT	1260
GCCCAACCACC	ACCAGTGTGC	CGCACAAGGC	CGTGGCGGTA	GGGTATGTGT	CTGAAAATGA	1320
GCTCGGGGAG	CGGGCTTGCA	CCGCTGACGC	ATTTGGAAGA	CTTAAGGCAG	CGGCAGAAGA	1380
AGATGCAGGC	AGCTGAGTTG	TTGTGTTCTG	ATAAGAGTCA	GAGGTAAGTC	CCGTTGCGGT	1440
GCTGTAAACG	GTGGAGGGCA	GTGTAGTCTG	AGCAGTACTC	GTTGCTGCCG	CGCGCGCCAC	1500
CAGACATAAT	AGCTGACAGA	CTAACAGACT	GTTCCCTTTCC	ATGGGTCTTT	TCTGCAGTCA	1560
CCGTCCTTGA	CACGCGTCTC	GGGAAGCTTG	CCGCCACCAT	GGATTCACAG	GCCCAGGTTC	1620
TTATGTTACT	GCCGCTATGG	GTATCTGGTA	CCTGTGGGGA	CATTGTGATG	TCACAGTCTC	1680
CATCCTCCCT	AGCTGTGTCA	GTTGGAGAGA	AGGTTACTAT	GAGCTGCAAG	TCCAGTCAGA	1740
GCCTTTTATA	TTCTAGAAAT	CAAAAGAACT	ACTTGGCCTG	GTTCCAGCAG	AAGCCAGGGC	1800
AGTCTCCTAA	ACTGCTGATT	TTCTGGGCAT	CCACTAGGGA	ATCTGGGGTC	CCTGATCGCT	1860
TCACAGGCAG	TGGATTTGGG	ACGGATTTC	ATCTCACC	CAGCAGTGTG	CAGGCTGAGG	1920
ACCTGGCAGT	TTATGACTGT	CAGCAATATT	TTAGCTATCC	GCTCACGTTC	GGTGTGGGA	1980
CCAAGCTGGA	GCTGAAACGT	GAGTGGATCC	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	2040
TCCCTAACAT	GCCCTGTGAT	TATGCGCAA	CAACACACCC	AAGGCAGAA	CTTTGTTACT	2100
TAAACACCAT	CCTGTTTGCT	TCTTTCCTCA	GGAAGTGTGG	CTGCACCATC	TGTCTTCATC	2160
TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	2220
AACTTCTATC	CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	2280
AACTCCCAGG	AGAGTGTCAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	2340
ACCCTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG	TCTACGCCTG	CGAAGTCACC	2400
CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG	AGCTTCAACA	GGGGAGAGTG	TTAGAGGGAG	2460
AAGTGCCCCC	ACCTGCTCCT	CAGTTCCAGC	CTGACCCCT	CCCATCCTTT	GGCCTCTGAC	2520
CCTTTTTC	CAGGGGACCT	ACCCCTATTG	CGGTCCTCCA	GCTCATCTTT	CACCTCACCC	2580

CCCTCCTCCT	CCTTGGCTTT	AATTATGCTA	ATGTTGGAGG	AGAATGAATA	AATAAAGTGA	2640
ATCTTTGCAC	CTGTGGTGGA	TCTAATAAAA	GATATTTATT	TTCATTAGAT	ATGTGTGTTG	2700
GTTTTTTGTG	TGCAGTGCCT	CTATCTGGAG	GCCAGGTAGG	GCTGGCCTTG	GGGGAGGGGG	2760
AGGCCAGAAT	GACTCCAAGA	GCTACAGGAA	GGCAGGTCAG	AGACCCCACT	GGACAAACAG	2820
TGGCTGGACT	CTGCACCATA	ACACACAATC	AACAGGGGAG	TGAGCTGGAA	ATTTGCTAGC	2880
GAATTCTTGA	AGACGAAAGG	GCCTCGTGAT	ACGCCTATTT	TTATAGGTTA	ATGTCATGAT	2940
AATAATGGTT	TCTTAGACGT	CAGGTGGCAC	TTTTCGGGGA	AATGTGCGCG	GAACCCCTAT	3000
TTGTTTATTT	TTCTAAATAC	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	3060
AATGCTTCAA	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT	3120
TATTCCCTTT	TTTGCGGCAT	TTTGCCTTCC	TGTTTTTGCT	CACCCAGAAA	CGCTGGTGAA	3180
AGTAAAAGAT	GCTGAAGATC	AGTTGGGTGC	ACGAGTGGGT	TACATCGAAC	TGGATCTCAA	3240
CAGCGGTAAG	ATCCTTGAGA	GTTTTCGCCC	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	3300
TAAAGTTCTG	CTATGTGGCG	CGGTATTATC	CCGTGTTGAC	GCCGGGCAAG	AGCAACTCGG	3360
TCGCCGCATA	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA	3420
TCTTACGGAT	GGCATGACAG	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA	TGAGTGATAA	3480
CACTGCGGCC	AACTTACTTC	TGACAACGAT	CGGAGGACCG	AAGGAGCTAA	CCGCTTTTTT	3540
GCACAACATG	GGGGATCATG	TAACTCGCCT	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	3600
CATACCAAAC	GACGAGCGTG	ACACCACGAT	GCCTGCAGCA	ATGGCAACAA	CGTTGCGCAA	3660
ACTATTAACT	GGCGAACTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA	3720
GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT	GGTTTATTGC	3780
TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC	ATTGCAGCAC	TGGGGCCAGA	3840
TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	3900
ACGAAATAGA	CAGATCGCTG	AGATAGGTGC	CTCACTGATT	AAGCATTGGT	AACTGTCAGA	3960
CCAAGTTTAC	TCATATATAC	TTTAGATTGA	TTTAAAACTT	CATTTTTAAT	TTAAAAGGAT	4020
CTAGGTGAAG	ATCCTTTTTG	ATAATCTCAT	GACCAAAATC	CCTTAACGTG	AGTTTTCGTT	4080
CCACTGAGCG	TCAGACCCCG	TAGAAAAGAT	CAAAGGATCT	TCTTGAGATC	CTTTTTTTCT	4140
GCGCGTAATC	TGCTGCTTGC	AAACAAAAAA	ACCACCGCTA	CCAGCGGTGG	TTTGTTTGCC	4200
GGATCAAGAG	CTACCAACTC	TTTTTCCGAA	GGTAACTGGC	TTCAGCAGAG	CGCAGATACC	4260
AAATACTGTC	CTTCTAGTGT	AGCCGTAGTT	AGGCCACCAC	TTCAAGAACT	CTGTAGCACC	4320
GCCTACATAC	CTCGCTCTGC	TAATCCTGTT	ACCAGTGGCT	GCTGCCAGTG	GCGATAAGTC	4380

GTGTCTTACC	GGGTTGGACT	CAAGACGATA	GTTACCGGAT	AAGGCGCAGC	GGTCGGGCTG	4440
AACGGGGGGT	TCGTGCACAC	AGCCCAGCTT	GGAGCGAACG	ACCTACACCG	AACTGAGATA	4500
CCTACAGCGT	GAGCTATGAG	AAAGCGCCAC	GCTTCCCGAA	GGGAGAAAGG	CGGACAGGTA	4560
TCCGGTAAGC	GGCAGGGTCG	GAACAGGAGA	GCGCACGAGG	GAGCTTCCAG	GGGGAAACGC	4620
CTGGTATCTT	TATAGTCCTG	TCGGGTTTCG	CCACCTCTGA	CTTGAGCGTC	GATTTTTGTG	4680
ATGCTCGTCA	GGGGGGCGGA	GCCTATGGAA	AAACGCCAGC	AACGCGGCCT	TTTTACGGTT	4740
CCTGGCCTTT	TGCTGGCCTT	TTGCTCACAT	GTTCTTTCTT	GCGTTATCCC	CTGATTCTGT	4800
GGATAACCGT	ATTACCGCCT	TTGAGTGAGC	TGATACCGCT	CGCCGCAGCC	GAACGACCGA	4860
GCGCAGCGAG	TCAGTGAGCG	AGGAAGCGGA	AGAGCGCCTG	ATGCGGTATT	TTCTCCTTAC	4920
GCATCTGTGC	GGTATTTTAC	ACCGCATATG	GTGCACTCTC	AGTACAATCT	GCTCTGATGC	4980
CGCATAGTTA	AGCCAGTATA	CACTCCGCTA	TCGCTACGTG	ACTGGGTCAT	GGCTGCGCCC	5040
CGACACCCGC	CAACACCCGC	TGACGCGCCC	TGACGGGCTT	GTCTGCTCCC	GGCATCCGCT	5100
TACAGACAAG	CTGTGACCGT	CTCCGGGAGC	TGCATGTGTC	AGAGGTTTTT	ACCGTCATCA	5160
CCGAAACGCG	CGAGGCAGCT	GTGGAATGTG	TGTCAGTTAG	GGTGTGGAAA	GTCCCCAGGC	5220
TCCCCAGCAG	GCAGAAGTAT	GCAAAGCATG	CATCTCAATT	AGTCAGCAAC	CAGGCTCCCC	5280
AGCAGGCAGA	AGTATGCAAA	GCATGCATCT	CAATTAGTCA	GCAACCATAG	TCCCGCCCCCT	5340
AACTCCGCCC	ATCCCGCCCC	TAActCCGCC	CAGTTCCGCC	CATTCTCCGC	CCCATGGCTG	5400
ACTAATTTTT	TTTATTTTATG	CAGAGGCCGA	GGCCGCCTCG	GCCTCTGAGC	TATTCAGAA	5460
GTAGTGAGGA	GGCTTTTTTTG	GAGGCCTAGG	CTTTTGCAAA	AAGCTAGCTT	CACGCTGCCG	5520
CAAGCACTCA	GGGCGCAAGG	GCTGCTAAAG	GAAGCGGAAC	ACGTAGAAAG	CCAGTCCGCA	5580
GAAACGGTGC	TGACCCCGGA	TGAATGTCAG	CTACTGGGCT	ATCTGGACAA	GGGAAAACGC	5640
AAGCGCAAAG	AGAAAGCAGG	TAGCTTGAGC	TGGGCTTACA	TGGCGATAGC	TAGACTGGGC	5700
GGTTTTATGG	ACAGCAAGCG	AACCGGAATT	GCCAGCTGGG	GCGCCCTCTG	GTAAGGTTGG	5760
GAAGCCCCTGC	AAAGTAAACT	GGATGGCTTT	CTTGCCGCCA	AGGATCTGAT	GGCGCAGGGG	5820
ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	AAGATGGATT	5880
GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	GGGCACAACA	5940
GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	GCCCGGTTCT	6000
TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	CAGCGCGGCT	6060
ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	TCACTGAAGC	6120
GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	6180
TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	ATACGCTTGA	6240

TCCGGCTACC	TGCCCATTTCG	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	CACGTACTCG	6300
GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	GGCTCGCGCC	6360
AGCCGAAGTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCCGAC	GGCGAGGATC	TCGTCTGTAC	6420
CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	CTGGATTTCAT	6480
CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	CTACCCGTGA	6540
TATTGCTGAA	GAGCTTGCGG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	ACGGTATCGC	6600
CGCTCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	6660
ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCCAACC	TGCCATCACG	AGATTTTCGAT	6720
TCCACCGCCG	CCTTCTATGA	AAGGTGGGC	TTCGGAATCG	TTTTCCGGGA	CGCCGGCTGG	6780
ATGATCCTCC	AGCGCGGGGA	TCTCATGCTG	GAGTTCCTTCG	CCCACCCCGG	GCTCGATCCC	6840
CTCGCGAGTT	GGTTCAGCTG	CTGCCTGAGG	CTGGACGACC	TCGCGGAGTT	CTACCGGCAG	6900
TGCAAATCCG	TCGGCATCCA	GGAAACCAGC	AGCGGCTATC	CGCGCATCCA	TGCCCCCGAA	6960
CTGCAGGAGT	GGGGAGGCAC	GATGGCCGCT	TTGGTCCCGG	ATCTTTGTGA	AGGAACCTTA	7020
CTTCTGTGGT	GTGACATAAT	TGGACAAACT	ACCTACAGAG	ATTTAAAGCT	CTAAGGTAAA	7080
TATAAAATTT	TTAAGTGTAT	AATGTGTAA	ACTACTGATT	CTAATTGTTT	GTGTATTTTA	7140
GATTCCAACC	TATGGAACTG	ATGAATGGGA	GCAGTGGTGG	AATGCCTTTA	ATGAGGAAAA	7200
CCTGTTTTGC	TCAGAAGAAA	TGCCATCTAG	TGATGATGAG	GCTACTGCTG	ACTCTCAACA	7260
TTCTACTCCT	CCAAAAAAGA	AGAGAAAGGT	AGAAGACCCC	AAGGACTTTC	CTTCAGAATT	7320
GCTAAGTTTT	TTGAGTCATG	CTGTGTTTAG	TAATAGAACT	CTTGCTTGCT	TTGCTATTTA	7380
CACCACAAAG	GAAAAAGCTG	CACTGCTATA	CAAGAAAATT	ATGGAAAAAT	ATTCTGTAAC	7440
CTTTATAAGT	AGGCATAACA	GTTATAATCA	TAACATACTG	TTTTTCTTCTTA	CTCCACACAG	7500
GCATAGAGTG	TCTGCTATTA	ATAACTATGC	TCAAAAATTG	TGTACCTTTA	GCTTTTTAAT	7560
TTGTAAAGGG	GTTAATAAGG	AATATTTGAT	GTATAGTGCC	TTGACTAGAG	ATCATAATCA	7620
GCCATACCAC	ATTTGTAGAG	GTTTTACTTG	CTTTAAAAAA	CCTCCCACAC	CTCCCCCTGA	7680
ACCTGAAACA	TAAAATGAAT	GCAATTGTTG	TTGTTAACTT	GTTTATTGCA	GCTTATAATG	7740
GTTACAAATA	AAGCAATAGC	ATCACAAATT	TCACAAATAA	AGCATTTTTT	TCACTGCATT	7800
CTAGTTGTGG	TTTGTCCAAA	CTCATCAATG	TATCTTATCA	TGTCTGGATC	TAATAAAAGA	7860
TATTTATTTT	CATTAGATAT	GTGTGTTGGT	TTTTTGTGTG	CAGTGCCTCT	ATCTGGAGGC	7920
CAGGTAGGGC	TGGCCTTGGG	GGAGGGGGAG	GCCAGAATGA	CTCCAAGAGC	TACAGGAAGG	7980
CAGGTCAGAG	ACCCCACTGG	ACAAACAGTG	GCTGGACTCT	GCACCATAAC	ACACAATCAA	8040

CAGGGGAGTG AGCTGGAAAT TTGCTAGC

8068

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 239 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7731 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

TTGAAGACGA AAGGGCCTCG TGATACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT	60
GGTTTCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT	120
ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT	180
TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCG CCCTTATTCC	240
CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA	300
AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG	360
TAAGATCCTT GAGAGTTTTT CCCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTTAAAGT	420
TCTGCTATGT GGC GCGGTAT TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG	480
CATACACTAT TCTCAGAAATG ACTTGGTTGA GTACTCACCA GTCACAGAAA AGCATCTTAC	540
GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC	600
GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTTGCACAA	660
CATGGGGGAT CATGTAACTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC	720
AAACGACGAG CGTGACACCA CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAACCTATT	780
AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA	840
TAAAGTTGCA GGACCACTTC TGCCTCGGC CCTTCCGGCT GGCTGGTTTA TTGCTGATAA	900
ATCTGGAGCC GGTGAGCGTG GGTCTCGCGG TATCATTGCA GCACTGGGGC CAGATGGTAA	960
GCCCTCCCGT ATCGTAGTTA TCTACACGAC GGGGAGTCAG GCAACTATGG ATGAACGAAA	1020
TAGACAGATC GCTGAGATAG GTGCCTCACT GATTAAGCAT TGGTAACTGT CAGACCAAGT	1080
TTACTCATAT ATACTTTAGA TTGATTAAA ACTTCATTTT TAATTTAAAA GGATCTAGGT	1140
GAAGATCCTT TTTGATAATC TCATGACCAA AATCCCTTAA CGTGAGTTTT CGTTCCACTG	1200
AGCGTCAGAC CCCGTAGAAA AGATCAAAGG ATCTTCTTGA GATCCTTTTT TTCTGCGCGT	1260
AATCTGCTGC TTGCAAACAA AAAAACCACC GCTACCAGCG GTGGTTTGTT TGCCGGATCA	1320
AGAGCTACCA ACTCTTTTTC CGAAGGTAAC TGGCTTCAGC AGAGCGCAGA TACCAAATAC	1380

TGTCCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG	AACTCTGTAG	CACCGCCTAC	1440
ATACCTCGCT	CTGCTAATCC	TGTTACCAGT	GGCTGCTGCC	AGTGGCGATA	AGTCGTGTCT	1500
TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	CAGCGGTCGG	GCTGAACGGG	1560
GGGTTCGTGC	ACACAGCCCA	GCTTGGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	1620
GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	1680
AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	1740
TCTTTATAGT	CCTGTCGGGT	TTCGCCACCT	CTGACTTGAG	CGTCGATTTT	TGTGATGCTC	1800
GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTTTAC	GGTTCCTGGC	1860
CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	1920
CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	AGCCGAACGA	CCGAGCGCAG	1980
CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCTGATGCGG	TATTTTCTCC	TTACGCATCT	2040
GTGCGGTATT	TCACACCGCA	TATGGTGCAC	TCTCAGTACA	ATCTGCTCTG	ATGCCGCATA	2100
GTTAAGCCAG	TATACACTCC	GCTATCGCTA	CGTGACTGGG	TCATGGCTGC	GCCCCGACAC	2160
CCGCCAACAC	CCGCTGACGC	GCCCTGACGG	GCTTGTCTGC	TCCCGGCATC	CGCTTACAGA	2220
CAAGCTGTGA	CCGTCTCCGG	GAGCTGCATG	TGTCAGAGGT	TTTCACCGTC	ATCACCAGAA	2280
CGCGCGAGGC	AGCATGCATC	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACTCCGCC	2340
CATCCCGCCC	CTAACTCCGC	CCAGTTCCGC	CCATTCTCCG	CCCCATGGCT	GACTAATTTT	2400
TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	GGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	2460
AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTAGCT	TACAGCTCAG	GGCTGCGATT	2520
TCGCGCCAAA	CTTGACGGCA	ATCCTAGCGT	GAAGGCTGGT	AGGATTTTAT	CCCCGCTGCC	2580
ATCATGGTTC	GACCATTGAA	CTGCATCGTC	GCCGTGTCCC	AAAATATGGG	GATTGGCAAG	2640
AACGGAGACC	TACCCTGGCC	TCCGCTCAGG	AACGAGTTCA	AGTACTTCCA	AAGAATGACC	2700
ACAACCTCTT	CAGTGGAAGG	TAAACAGAAT	CTGGTGATTA	TGGGTAGGAA	AACCTGGTTC	2760
TCCATTCCCTG	AGAAGAATCG	ACCTTTAAAG	GACAGAATTA	ATATAGTTCT	CAGTAGAGAA	2820
CTCAAAGAAC	CACCACGAGG	AGCTCATTTT	CTTGCCAAAA	GTTTGGATGA	TGCCTTAAGA	2880
CTTATTGAAC	AACCGGAATT	GGCAAGTAAA	GTAGACATGG	TTTGGATAGT	CGGAGGCAGT	2940
TCTGTTTACC	AGGAAGCCAT	GAATCAACCA	GGCCACCTCA	GACTCTTTGT	GACAAGGATC	3000
ATGCAGGAAT	TTGAAAAGTG	CACGTTTTTC	CCAGAAATTG	ATTTGGGGAA	ATATAAACTT	3060
CTCCCAGAAT	ACCCAGGCGT	CCTCTCTGAG	GTCCAGGAGG	AAAAGGCAT	CAAGTATAAG	3120
TTTGAAGTCT	ACGAGAAGAA	AGACTAACAG	GAAGATGCTT	TCAAGTTCTC	TGCTCCCCTC	3180

CTAAAGCTAT	GCATTTTTAT	AAGACCATGG	GACTTTTGCT	GGCTTTAGAT	CTTTGTGAAG	3240
GAACCTTACT	TCTGTGGTGT	GACATAATTG	GACAAACTAC	CTACAGAGAT	TTAAAGCTCT	3300
AAGGTAAATA	TAAAATTTTT	AAGTGTATAA	TGTGTTAAAC	TACTGATTCT	AATTGTTTGT	3360
GTATTTTAGA	TTCCAACCTA	TGGAACTGAT	GAATGGGAGC	AGTGGTGGAA	TGCCTTTAAT	3420
GAGGAAAACC	TGTTTTGCTC	AGAAGAAATG	CCATCTAGTG	ATGATGAGGC	TACTGCTGAC	3480
TCTCAACATT	CTACTCCTCC	AAAAAAGAAG	AGAAAGGTAG	AAGACCCCAA	GGACTTTCCT	3540
TCAGAATTGC	TAAGTTTTTT	GAGTCATGCT	GTGTTTAGTA	ATAGAACTCT	TGCTTGCTTT	3600
GCTATTTACA	CCACAAAGGA	AAAAGCTGCA	CTGCTATACA	AGAAAATTAT	GGAAAAATAT	3660
TCTGTAACCT	TTATAAGTAG	GCATAACAGT	TATAATCATA	ACATACTGTT	TTTTCTTACT	3720
CCACACAGGC	ATAGAGTGTC	TGCTATTAAT	AACTATGCTC	AAAAATTGTG	TACCTTTAGC	3780
TTTTTAATTT	GTAAAGGGGT	TAATAAGGAA	TATTTGATGT	ATAGTGCCTT	GACTAGAGAT	3840
CATAATCAGC	CATACCACAT	TTGTAGAGGT	TTTACTTGCT	TTAAAAACC	TCCCACACCT	3900
CCCCCTGAAC	CTGAAACATA	AAATGAATGC	AATTGTTGTT	GTAACTTGT	TTATTGCAGC	3960
TTATAATGGT	TACAAATAAA	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTC	4020
ACTGCATTCT	AGTTGTGGTT	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGGATCTA	4080
ATAAAAAGATA	TTTATTTTCA	TTAGATATGT	GTGTTGGTTT	TTTGTGTGCA	GTGCCTCTAT	4140
CTGGAGGCCA	GGTAGGGCTG	GCCTTGGGGG	AGGGGGAGGC	CAGAATGACT	CCAAGAGCTA	4200
CAGGAAGGCA	GGTCAGAGAC	CCCCTGGGAC	AAACAGTGGC	TGGACTCTGC	ACCATAACAC	4260
ACAATCAACA	GGGGAGTGAG	CTGGAAATTT	GCTAGCGAAT	TCCAGCACAC	TGGCGGCCGT	4320
TACTAGTTAT	TAATAGTAAT	CAATTACGGG	GTCATTAGTT	CATAGCCCAT	ATATGGAGTT	4380
CCGCGTTACA	TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCCAAG	ACCCCCGCCC	4440
ATTGACGTCA	ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	4500
TCAATGGGTG	GAGTATTTAC	GGTAACTGTC	CCACTTGCCA	GTACATCAAG	TGTATCATAT	4560
GCCAAGTACG	CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCGCCTGGC	ATTATGCCCA	4620
GTACATGACC	TTATGGGACT	TTCTACTTGG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	4680
TACCATGGTG	ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	4740
GGGATTTCCA	AGTCTCCACC	CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAATCA	4800
ACGGGACTTT	CCAAAATGTC	GTAACAACTC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	4860
TGTACGGTGG	GAGGTCTATA	TAAGCAGAGC	TCGTTTAGTG	AACCGTCAGA	TCGCCTGGAG	4920
ACGCCATCCA	CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	4980
CCGGGAACGG	TGCATTGGAA	CGCGGATTCC	CCGTGCCAAG	AGTGACGTAA	GTACCGCCTA	5040

TAGAGTCTAT	AGGCCCCACCC	CCTTGGCTTC	TTATGCATGC	TATACTGTTT	TTGGCTTGGG	5100
GTCTATACAC	CCCCGCTTCC	TCATGTTATA	GGTGATGGTA	TAGCTTAGCC	TATAGGTGTG	5160
GGTTATTGAC	CATTATTGAC	CACTCCCCTA	TTGGTGACGA	TACTTTCCAT	TACTAATCCA	5220
TAACATGGCT	CTTTGCCACA	ACTCTCTTTA	TTGGCTATAT	GCCAATACAC	TGTCCTTCAG	5280
AGACTGACAC	GGACTCTGTA	TTTTTACAGG	ATGGGGTCTC	ATTTATTATT	TACAAATTCA	5340
CATATACAAC	ACCACCGTCC	CCAGTGCCCG	CAGTTTTTAT	TAAACATAAC	GTGGGATCTC	5400
CACGCGAATC	TCGGGTACGT	GTTCCGGACA	TGGGCTCTTC	TCCGGTAGCG	GCGGAGCTTC	5460
TACATCCGAG	CCCTGCTCCC	ATGCCTCCAG	CGACTCATGG	TCGCTCGGCA	GCTCCTTGCT	5520
CCTAACAGTG	GAGGCCAGAC	TTAGGCACAG	CACGATGCCC	ACCACCACCA	GTGTGCCGCA	5580
CAAGGCCGTG	GCGGTAGGGT	ATGTGTCTGA	AAATGAGCTC	GGGAGCGGG	CTTGCAACGC	5640
TGACGCATTT	GGAAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTTGT	5700
GTTCTGATAA	GAGTCAGAGG	TAACCTCCGT	TGCGGTGCTG	TTAACGGTGG	AGGGCAGTGT	5760
AGTCTGAGCA	GTA CTGTTG	CTGCCGCGCG	CGCCACCAGA	CATAATAGCT	GACAGACTAA	5820
CAGACTGTTC	CTTTCCATGG	GTCTTTTCTG	CAGTCACCGT	CCTTGACACG	CGTCTCGGGA	5880
AGCTTGCCGC	CACCATGGGA	TGGAGCTGGG	TCTTTCTCTT	TCTCCTGTCA	GGAAGTGCAG	5940
GTGTCTCTC	TGAGGTCCAG	CTGCAACAGT	CTGGACCTGA	GCTGGTGAAG	CCTGGGGCTT	6000
CAGTAAAGAT	GTCCTGCAAG	ACTTCTAGAT	ACACATTAC	TGAATACACC	ATACACTGGG	6060
TGAGACAGAG	CCATGGAAAAG	AGCCTTGAGT	GGATTGGAGG	TATTAATCCT	AACAATGGTA	6120
TTCTTAATA	CAACCAGAAG	TTCAAGGGCA	GGGCCACATT	GACTGTAGGC	AAGTCCTCCA	6180
GCACCGCCTA	CATGGAGCTC	CGCAGCCTGA	CATCTGAGGA	TTCTGCGGTC	TATTTCTGTG	6240
CAAGAAGAAG	AATCGCCTAT	GGTTACGACG	AGGGCCATGC	TATGGACTAC	TGGGGTCAAG	6300
GAACCTCAGT	CACCGTCTCC	TCAGGTGAGT	GGATCCTCTG	CGCCTGGGCC	CAGCTCTGTC	6360
CCACACCGCG	GTCACATGGC	ACCACCTCTC	TTGCAGCCTC	CACCAAGGGC	CCATCGGTCT	6420
TCCCCCTGGC	ACCCTCCTCC	AAGAGCACCT	CTGGGGGCAC	AGCGGCCCTG	GGCTGCCTGG	6480
TCAAGGACTA	CTTCCCCGAA	CCGGTGACGG	TGTCGTGGAA	CTCAGGCGCC	CTGACCAGCG	6540
GCGTGACAC	CTTCCCGGCT	GTCCTACAGT	CCTCAGGACT	CTACTCCCTC	AGCAGCGTGG	6600
TGACCGTGCC	CTCCAGCAGC	TTGGGCACCC	AGACCTACAT	CTGCAACGTG	AATCACAAGC	6660
CCAGCAACAC	CAAGGTGGAC	AAGAAAGTTG	AGCCCAAATC	TTGTGACAAA	ACTCACACAT	6720
GCCCACCGTG	CCCAGCACCT	GAACCTCTGG	GGGGACCGTC	AGTCTTCCTC	TTCCCCCAA	6780
AACCCAAGGA	CACCTCATG	ATCTCCCGGA	CCCCTGAGGT	CACATGCGTG	GTGGTGACG	6840

```

TGAGCCACGA AGACCCTGAG GTCAAGTTCA ACTGGTACGT GGACGGCGTG GAGGTGCATA      6900
ATGCCAAGAC AAAGCCGCGG GAGGAGCAGT ACAACAGCAC GTACCGGGTG GTCAGCGTCC      6960
TCACCGTCCT GCACCAGGAC TGGCTGAATG GCAAGGAGTA CAAGTGCAAG GTCTCCAACA      7020
AAGCCCTCCC AGCCCCCATC GAGAAAACCA TCTCCAAAGC CAAAGGGCAG CCCCAGAGAAC      7080
CACAGGTGTA CACCCTGCCC CCATCCCGGG AGGAGATGAC CAAGAACCAG GTCAGCCTGA      7140
CCTGCCTGGT CAAAGGCTTC TATCCCAGCG ACATCGCCGT GGAGTGGGAG AGCAATGGGC      7200
AGCCGGAGAA CAACTACAAG ACCACGCCTC CCGTGCTGGA CTCCGACGGC TCCTTCTTCC      7260
TCTACAGCAA GCTCACCGTG GACAAGAGCA GGTGGCAGCA GGGGAACGTC TTCTCATGCT      7320
CCGTGATGCA TGAGGCTCTG CACAACCACT ACACGCAGAA GAGCCTCTCC CTGTCTCCGG      7380
GTAAATGAGT GCGACGGCCG GCAAGCCCCG CTCCCCGGGC TCTCGCGGTC GCACGAGGAT      7440
GCTTGGCACG TACCCCTGT ACATACTTCC CGGGCGCCCA GCATGGAAAT AAAGCACCGG      7500
ATCTAATAAA AGATATTTAT TTTCATTAGA TATGTGTGTT GGTTTTTTGT GTGCAGTGCC      7560
TCTATCTGGA GGCCAGGTAG GGCTGGCCTT GGGGGAGGGG GAGGCCAGAA TGA CTCCAAG      7620
AGCTACAGGA AGGCAGGTCA GAGACCCAC TGGACAAACA GTGGCTGGAC TCTGCACCAT      7680
AACACACAAT CAACAGGGGA GTGAGCTGGA AATTTGCTAG CGAATTAATT C              7731

```

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

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

```

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

[illegible]

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 339 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA	GAGGGCCACC	60
ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA	ATCAAAAGAA	CTACTTGGCC	120
TGGTATCAGC	AGAAACCAGG	ACAGCCACCC	AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	180
GAATCTGGGG	TACCTGATAG	GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	240
ATTAGCAGCC	TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT	300
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA			339

(2) INFORMATION FOR SEO ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 113 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

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

```

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
   35                               40                               45

Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
   50                               55                               60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
  65                               70                               75                               80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
      85                               90                               95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
    100                               105                               110

Lys

```

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 113 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
1           5           10           15

Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser Leu Leu Tyr Ser
      20           25           30

Arg Asn Gln Lys Asn Tyr Leu Ala Trp Phe Gln Gln Lys Pro Gly Gln
   35                               40                               45

Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg Glu Ser Gly Val
   50                               55                               60

Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp Phe Thr Leu Thr
  65                               70                               75                               80

Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Asp Cys Gln Gln
      85                               90                               95

Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
    100                               105                               110

Lys

```

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 113 amino acids

- (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

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

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8068 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

```

GAATTCAGC ACACTGGCGG CCGTTACTAG TTATTAATAG TAATCAATTA CGGGGTCATT      60
AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT ACGGTAAATG GCGGCCTGG      120
CTGACCGCCC AACGACCCCC GCCCATTGAC GTCAATAATG ACGTATGTTC CCATAGTAAC      180
GCCAATAGGG ACTTTCCATT GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCACTT      240
GGCAGTACAT CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA      300
ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA CTTGGCAGTA      360
  
```

CATCTACGTA	TTAGTCATCG	CTATTACCAT	GGTGATGCGG	TTTTGGCAGT	ACATCAATGG	420
GCGTGGATAG	CGGTTTGACT	CACGGGGATT	TCCAAGTCTC	CACCCCATTG	ACGTCAATGG	480
GAGTTTGTTT	TGGCACCAAA	ATCAACGGGA	CTTTCCAAAA	TGTCGTAACA	ACTCCGCCCC	540
ATTGACGCAA	ATGGGCGGTA	GGCGTGACG	GTGGGAGGTC	TATATAAGCA	GAGCTCGTTT	600
AGTGAACCGT	CAGATCGCCT	GGAGACGCCA	TCCACGCTGT	TTTGACCTCC	ATAGAAGACA	660
CCGGGACCGA	TCCAGCCTCC	GCGGCCGGGA	ACGGTGCAAT	GGAACGCGGA	TTCCCCGTGC	720
CAAGAGTGAC	GTAAGTACCG	CCTATAGAGT	CTATAGGCCC	ACCCCCTTGG	CTTCTTATGC	780
ATGCTATACT	GTTTTTGCT	TGGGGTCTAT	ACACCCCGCG	TTCCTCATGT	TATAGGTGAT	840
GGTATAGCTT	AGCCTATAGG	TGTGGGTTAT	TGACCATTAT	TGACCACTCC	CCTATTGGTG	900
ACGATACTTT	CCATTACTAA	TCCATAACAT	GGCTCTTTGC	CACAACTCTC	TTTATTGGCT	960
ATATGCCAAT	AACTGTCTCT	TCAGAGACTG	ACACGGACTC	TGTATTTTTA	CAGGATGGGG	1020
TCTCATTTAT	TATTTACAAA	TTCACATATA	CAACACCACC	GTCCCCAGTG	CCCGCAGTTT	1080
TTATTAAACA	TAACGTGGGA	TCTCCACGCG	AATCTCGGGT	ACGTGTTCCG	GACATGGGCT	1140
CTTCTCCGGT	AGCGGCGGAG	CTTCTACATC	CGAGCCCTGC	TCCCATGCCT	CCAGCGACTC	1200
ATGGTCGCTC	GGCAGCTCCT	TGCTCCTAAC	AGTGGAGGCC	AGACTTAGGC	ACAGCACGAT	1260
GCCCACCACC	ACCAGTGTGC	CGCACAAGGC	CGTGGCGGTA	GGGTATGTGT	CTGAAAATGA	1320
GCTCGGGGAG	CGGGCTTGCA	CCGCTGACGC	ATTTGGAAGA	CTTAAGGCAG	CGGCAGAAGA	1380
AGATGCAGGC	AGCTGAGTTG	TTGTGTTCTG	ATAAGAGTCA	GAGGTAATCT	CCGTTGCGGT	1440
GCTGTTAACG	GTGGAGGGCA	GTGTAGTCTG	AGCAGTACTC	GTTGCTGCCG	CGCGCGCCAC	1500
CAGACATAAT	AGCTGACAGA	CTAACAGACT	GTTCTTTTCC	ATGGGTCTTT	TCTGCAGTCA	1560
CCGTCCTTGA	CACGCGTCTC	GGAAGCTTG	CCGCCACCAT	GGAGACAGAC	AACTCCTGC	1620
TATGGGTGCT	GCTGCTCTGG	GTTCCAGGTT	CCTCCGGAGA	CATTGTGATG	ACCCAATCTC	1680
CAGACTCTTT	GGCTGTGTCT	CTAGGGGAGA	GGGCCACCAT	CAACTGCAAG	TCCAGTCAGA	1740
GCCTTTTATA	TTCTAGAAAT	CAAAAGAACT	ACTTGGCCTG	GTATCAGCAG	AAACCAGGAC	1800
AGCCACCCAA	ACTCCTCATC	TTTTGGGCTA	GCACTAGGGA	ATCTGGGGTA	CCTGATAGGT	1860
TCAGTGGCAG	TGGGTTTGGG	ACAGACTTCA	CCCTCACCAT	TAGCAGCCTG	CAGGCTGAAG	1920
ATGTGGCAGT	TTATTACTGT	CAGCAATATT	TTAGCTATCC	GCTCACGTTC	GGACAAGGGA	1980
CCAAGGTGGA	AATAAAACGT	GAGTGGATCC	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	2040
TCCCTAACAT	GCCCTGTGAT	TATGCGCAAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT	2100
TAAACACCAT	CCTGTTTGCT	TCTTTCTCTA	GGAAGTGTGG	CTGCACCATC	TGTCTTCATC	2160

TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	2220
AACCTTCTATC	CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	2280
AACCTCCAGG	AGAGTGTCAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	2340
ACCCTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG	TCTACGCCTG	CGAAGTCACC	2400
CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG	AGCTTCAACA	GGGGAGAGTG	TTAGAGGGAG	2460
AAGTGCCCCC	ACCTGCTCCT	CAGTTCCAGC	CTGACCCCCCT	CCCATCCTTT	GGCCTCTGAC	2520
CCTTTTTTCCA	CAGGGGACCT	ACCCCTATTG	CGGTCCTCCA	GCTCATCTTT	CACCTCACCC	2580
CCCTCCTCCT	CCTTGGCTTT	AATTATGCTA	ATGTTGGAGG	AGAATGAATA	AATAAAGTGA	2640
ATCTTTGCAC	CTGTGGTGGA	TCTAATAAAA	GATATTTATT	TTCATTAGAT	ATGTGTGTTG	2700
GTTTTTTGTG	TGCAGTGCCT	CTATCTGGAG	GCCAGGTAGG	GCTGGCCTTG	GGGGAGGGGG	2760
AGGCCAGAAT	GACTCCAAGA	GCTACAGGAA	GGCAGGTCAG	AGACCCCACT	GGACAAACAG	2820
TGGCTGGACT	CTGCACCATA	ACACACAATC	AACAGGGGAG	TGAGCTGGAA	ATTTGCTAGC	2880
GAATTCTTGA	AGACGAAAGG	GCCTCGTGAT	ACGCCTATTT	TTATAGGTTA	ATGTCATGAT	2940
AATAATGGTT	TCTTAGACGT	CAGGTGGCAC	TTTTCGGGGA	AATGTGCGCG	GAACCCCTAT	3000
TTGTTTATTT	TTCTAAATAC	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	3060
AATGCTTCAA	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT	3120
TATTCCTTTT	TTGCGGCAT	TTTGCCTTCC	TGTTTTTGCT	CACCCAGAAA	CGCTGGTGAA	3180
AGTAAAAGAT	GCTGAAGATC	AGTTGGGTGC	ACGAGTGGGT	TACATCGAAC	TGGATCTCAA	3240
CAGCGGTAAG	ATCCTTGAGA	GTTTTCGCCC	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	3300
TAAAGTTCTG	CTATGTGGCG	CGGTATTATC	CCGTGTTGAC	GCCGGGCAAG	AGCAACTCGG	3360
TCGCCGCATA	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA	3420
TCTTACGGAT	GGCATGACAG	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA	TGAGTGATAA	3480
CACTGCGGCC	AACTTACTTC	TGACAACGAT	CGGAGGACCG	AAGGAGCTAA	CCGCTTTTTT	3540
GCACAACATG	GGGGATCATG	TAACTCGCCT	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	3600
CATACCAAAC	GACGAGCGTG	ACACCACGAT	GCCTGCAGCA	ATGGCAACAA	CGTTGCGCAA	3660
ACTATTAACT	GGCGAACTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA	3720
GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT	GGTTTATTGC	3780
TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC	ATTGCAGCAC	TGGGGCCAGA	3840
TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	3900
ACGAAATAGA	CAGATCGCTG	AGATAGGTGC	CTCACTGATT	AAGCATTGGT	AACTGTCAGA	3960
CCAAGTTTAC	TCATATATAC	TTTAGATTGA	TTTAAAACTT	CATTTTAAAT	TTAAAAGGAT	4020

CTAGGTGAAG ATCCTTTTTG ATAATCTCAT GACCAAAATC CCTTAACGTG AGTTTTCGTT	4080
CCACTGAGCG TCAGACCCCG TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT	4140
GCGCGTAATC TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC	4200
GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAAGTGGC TTCAGCAGAG CGCAGATACC	4260
AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC TTCAAGAACT CTGTAGCACC	4320
GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT GCTGCCAGTG GCGATAAGTC	4380
GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTGGGGCTG	4440
AACGGGGGGT TCGTGACAC AGCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA	4500
CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA	4560
TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC	4620
CTGGTATCTT TATAGTCCTG TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG	4680
ATGCTCGTCA GGGGGGCGGA GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT	4740
CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT	4800
GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATACCGCT CGCCGAGCC GAACGACCGA	4860
GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG ATGCGGTATT TTCTCCTTAC	4920
GCATCTGTGC GGTATTTTAC ACCGCATATG GTGCACTCTC AGTACAATCT GCTCTGATGC	4980
CGCATAGTTA AGCCAGTATA CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC	5040
CGACACCCGC CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT	5100
TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTT ACCGTCATCA	5160
CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC	5220
TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGCTCCCC	5280
AGCAGGCAGA AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT	5340
AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG	5400
ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC TATTCCAGAA	5460
GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG CTTTTGCAAA AAGCTAGCTT CACGCTGCCG	5520
CAAGCACTCA GGGCGCAAGG GCTGCTAAAG GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA	5580
GAAACGGTGC TGACCCCGGA TGAATGTCAG CTACTGGGCT ATCTGGACAA GGGAAAACGC	5640
AAGCGCAAAG AGAAAGCAGG TAGCTTGCAG TGGGCTTACA TGGCGATAGC TAGACTGGGC	5700
GGTTTTATGG ACAGCAAGCG AACCAGGAATT GCCAGCTGGG GCGCCCTCTG GTAAGGTTGG	5760
GAAGCCCTGC AAAGTAAACT GGATGGCTTT CTTGCCGCCA AGGATCTGAT GCGCAGGGG	5820

ATCAAGATCT	GATCAAGAGA	CAGGATGAGG	ATCGTTTCGC	ATGATTGAAC	AAGATGGATT	5880
GCACGCAGGT	TCTCCGGCCG	CTTGGGTGGA	GAGGCTATTC	GGCTATGACT	GGGCACAACA	5940
GACAATCGGC	TGCTCTGATG	CCGCCGTGTT	CCGGCTGTCA	GCGCAGGGGC	GCCCGGTTCT	6000
TTTTGTCAAG	ACCGACCTGT	CCGGTGCCCT	GAATGAACTG	CAGGACGAGG	CAGCGCGGCT	6060
ATCGTGGCTG	GCCACGACGG	GCGTTCCTTG	CGCAGCTGTG	CTCGACGTTG	TCACTGAAGC	6120
GGGAAGGGAC	TGGCTGCTAT	TGGGCGAAGT	GCCGGGGCAG	GATCTCCTGT	CATCTCACCT	6180
TGCTCCTGCC	GAGAAAGTAT	CCATCATGGC	TGATGCAATG	CGGCGGCTGC	ATACGCTTGA	6240
TCCGGCTACC	TGCCCATTCT	ACCACCAAGC	GAAACATCGC	ATCGAGCGAG	CACGTACTCG	6300
GATGGAAGCC	GGTCTTGTCG	ATCAGGATGA	TCTGGACGAA	GAGCATCAGG	GGCTCGCGCC	6360
AGCCGAACTG	TTCGCCAGGC	TCAAGGCGCG	CATGCCCCGAC	GGCGAGGATC	TCGTCGTGAC	6420
CCATGGCGAT	GCCTGCTTGC	CGAATATCAT	GGTGGAAAAT	GGCCGCTTTT	CTGGATTCAT	6480
CGACTGTGGC	CGGCTGGGTG	TGGCGGACCG	CTATCAGGAC	ATAGCGTTGG	CTACCCGTGA	6540
TATTGCTGAA	GAGCTTGGCG	GCGAATGGGC	TGACCGCTTC	CTCGTGCTTT	ACGGTATCGC	6600
CGTCCCCGAT	TCGCAGCGCA	TCGCCTTCTA	TCGCCTTCTT	GACGAGTTCT	TCTGAGCGGG	6660
ACTCTGGGGT	TCGAAATGAC	CGACCAAGCG	ACGCCCAACC	TGCCATCACG	AGATTTTCGAT	6720
TCCACCGCCG	CCTTCTATGA	AAGGTTGGGC	TTCGGAATCG	TTTTCCGGGA	CGCCGGCTGG	6780
ATGATCCTCC	AGCGCGGGGA	TCTCATGCTG	GAGTTCTTCG	CCCACCCCGG	GCTCGATCCC	6840
CTCGCGAGTT	GGTTCAGCTG	CTGCCTGAGG	CTGGACGACC	TCGCGGAGTT	CTACCGGCAG	6900
TGCAAATCCG	TCGGCATCCA	GGAAACCAGC	AGCGGCTATC	CGCGCATCCA	TGCCCCCGAA	6960
CTGCAGGAGT	GGGGAGGCAC	GATGGCCGCT	TTGGTCCCGG	ATCTTTGTGA	AGGAACCTTA	7020
CTTCTGTGGT	GTGACATAAT	TGGACAAACT	ACCTACAGAG	ATTTAAAGCT	CTAAGGTAAA	7080
TATAAAATTT	TTAAGTGTAT	AATGTGTTAA	ACTACTGATT	CTAATTGTTT	GTGTATTTTA	7140
GATTCCAACC	TATGGAACTG	ATGAATGGGA	GCAGTGGTGG	AATGCCTTTA	ATGAGGAAAA	7200
CCTGTTTTGC	TCAGAAGAAA	TGCCATCTAG	TGATGATGAG	GCTACTGCTG	ACTCTCAACA	7260
TTCTACTCCT	CCAAAAAAGA	AGAGAAAAGT	AGAAGACCCC	AAGGACTTTC	CTTCAGAATT	7320
GCTAAGTTTT	TTGAGTCATG	CTGTGTTTAG	TAATAGAACT	CTTGCTTGCT	TTGCTATTTA	7380
CACCACAAAG	GAAAAAGCTG	CACTGCTATA	CAAGAAAATT	ATGGAAAAAT	ATTCTGTAAAC	7440
CTTTATAAGT	AGGCATAACA	GTTATAATCA	TAACATACTG	TTTTTTCTTA	CTCCACACAG	7500
GCATAGAGTG	TCTGCTATTA	ATAACTATGC	TCAAAAATTG	TGTACCTTTA	GCTTTTTTAAT	7560
TTGTAAAGGG	GTTAATAAGG	AATATTTGAT	GTATAGTGCC	TTGACTAGAG	ATCATAATCA	7620
GCCATACCAC	ATTTGTAGAG	GTTTTACTTG	CTTTAAAAAA	CCTCCACAC	CTCCCCCTGA	7680

```

ACCTGAAACA TAAATGAAT GCAATTGTTG TTGTTAACTT GTTTATTGCA GCTTATAATG      7740
GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT      7800
CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC TAATAAAAGA      7860
TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG CAGTGCCTCT ATCTGGAGGC      7920
CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG GCCAGAATGA CTCCAAGAGC TACAGGAAGG      7980
CAGGTCAGAG ACCCCACTGG ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA      8040
CAGGGGAGTG AGCTGGAAAT TTGCTAGC                                         8068

```

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 234 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

```

Met Glu Thr Asp Thr Leu Leu Leu Trp Val Leu Leu Leu Trp Val Pro
1              5              10              15

Gly Ser Ser Gly Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala
                20              25              30

Val Ser Leu Gly Glu Arg Ala Thr Ile Asn Cys Lys Ser Ser Gln Ser
          35              40              45

Leu Leu Tyr Ser Arg Asn Gln Lys Asn Tyr Leu Ala Trp Tyr Gln Gln
          50              55              60

Lys Pro Gly Gln Pro Pro Lys Leu Leu Ile Phe Trp Ala Ser Thr Arg
65              70              75              80

Glu Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Phe Gly Thr Asp
          85              90              95

Phe Thr Leu Thr Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr
          100             105             110

Tyr Cys Gln Gln Tyr Phe Ser Tyr Pro Leu Thr Phe Gly Gln Gly Thr
          115             120             125

Lys Val Glu Ile Lys Arg Val Phe Ile Phe Pro Pro Ser Asp Glu Gln
          130             135             140

Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe Tyr
          145             150             155             160

Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln Ser

```

38

	165		170		175
Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser Thr					
	180		185		190
Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu Lys					
	195		200		205
His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser Pro					
	210		215		220
Val Thr Lys Ser Phe Asn Arg Gly Glu Cys					
	225		230		

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 372 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC CGTGAAAGTC	60
AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA TACTACTGGGT TAGACAGGCC	120
CCTGGCCAAA GGCTGGAGTG GATAGGAGGT ATTAATCCTA ACAATGGTAT TCCTAACTAC	180
AACCAGAAGT TCAAGGGCCG GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC	240
ATGGAAGTGT CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA	300
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG AACCTTGTC	360
ACCGTCTCCT CA	372

(2) INFORMATION FOR SEQ ID NO: 38:

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

(ii) MOLECULE TYPE: peptide

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

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

39

```

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

```

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 124 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

```

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

```

(2) INFORMATION FOR SEQ ID NO: 40:

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

(ii) MOLECULE TYPE: peptide

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

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

(2) INFORMATION FOR SEQ ID NO: 41:

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

(ii) MOLECULE TYPE: peptide

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

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

Gly Gly Ile Asn Pro Asn Asn Gly Ile Pro Asn Tyr Asn Gln Lys Phe
 50 55 60
 Lys Gly Arg Val Thr Ile Thr Val Asp Thr Ser Ala Ser Thr Ala Tyr
 65 70 75 80
 Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
 85 90 95
 Ala Arg Arg Arg Ile Ala Tyr Gly Tyr Asp Glu Gly His Ala Met Asp
 100 105 110
 Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
 115 120

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 7731 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

TTGAAGACGA AAGGGCCTCG TGATACGCCT ATTTTATAG GTTAATGTCA TGATAATAAT 60
 GGTTCCTTAG ACGTCAGGTG GCACTTTTCG GGGAAATGTG CGCGGAACCC CTATTTGTTT 120
 ATTTTCTAA ATACATTCAA ATATGTATCC GCTCATGAGA CAATAACCCT GATAAATGCT 180
 TCAATAATAT TGAAAAAGGA AGAGTATGAG TATTCAACAT TTCCGTGTCT CCCTTATTC 240
 CTTTTTTGCG GCATTTTGCC TTCCTGTTTT TGCTCACCCA GAAACGCTGG TGAAAGTAAA 300
 AGATGCTGAA GATCAGTTGG GTGCACGAGT GGGTTACATC GAACTGGATC TCAACAGCGG 360
 TAAGATCCTT GAGAGTTTTT GCCCGAAGA ACGTTTTCCA ATGATGAGCA CTTTAAAGT 420
 TCTGCTATGT GGC GCGGTAT TATCCCGTGT TGACGCCGGG CAAGAGCAAC TCGGTCGCCG 480
 CATACACTAT TCTCAGAATG ACTTGTTGA GTACTACCA GTCACAGAAA AGCATCTTAC 540
 GGATGGCATG ACAGTAAGAG AATTATGCAG TGCTGCCATA ACCATGAGTG ATAACACTGC 600
 GGCCAACTTA CTTCTGACAA CGATCGGAGG ACCGAAGGAG CTAACCGCTT TTTGCACAA 660
 CATGGGGGAT CATGTAAGTC GCCTTGATCG TTGGGAACCG GAGCTGAATG AAGCCATACC 720
 AAACGACGAG CGTGACACCA CGATGCCTGC AGCAATGGCA ACAACGTTGC GCAAACATTT 780
 AACTGGCGAA CTACTTACTC TAGCTTCCCG GCAACAATTA ATAGACTGGA TGGAGGCGGA 840
 TAAAGTTGCA GGACCACTTC TCGCTCGGC CCTCCGGCT GGCTGGTTTA TTGCTGATAA 900

ATCTGGAGCC	GGTGAGCGTG	GGTCTCGCGG	TATCATTGCA	GCACTGGGGC	CAGATGGTAA	960
GCCCTCCCGT	ATCGTAGTTA	TCTACACGAC	GGGGAGTCAG	GCAACTATGG	ATGAACGAAA	1020
TAGACAGATC	GCTGAGATAG	GTGCCTCACT	GATTAAGCAT	TGGTAACTGT	CAGACCAAGT	1080
TTACTCATAT	ATACTTTAGA	TTGATTTAAA	ACTTCATTTT	TAATTTAAAA	GGATCTAGGT	1140
GAAGATCCTT	TTTGATAATC	TCATGACCAA	AATCCCTTAA	CGTGAGTTTT	CGTTCCACTG	1200
AGCGTCAGAC	CCCGTAGAAA	AGATCAAAGG	ATCTTCTTGA	GATCCTTTTT	TTCTGCGCGT	1260
AATCTGCTGC	TTGCAAACAA	AAAAACCACC	GCTACCAGCG	GTGGTTTGTT	TGCCGGATCA	1320
AGAGCTACCA	ACTCTTTTTC	CGAAGGTAAC	TGGCTTCAGC	AGAGCGCAGA	TACCAAATAC	1380
TGTCCTTCTA	GTGTAGCCGT	AGTTAGGCCA	CCACTTCAAG	AACTCTGTAG	CACCGCCTAC	1440
ATACCTCGCT	CTGCTAATCC	TGTTACCAGT	GGCTGCTGCC	AGTGGCGATA	AGTCGTGTCT	1500
TACCGGGTTG	GACTCAAGAC	GATAGTTACC	GGATAAGGCG	CAGCGGTCGG	GCTGAACGGG	1560
GGGTTTCGTG	ACACAGCCCA	GCTTGGAGCG	AACGACCTAC	ACCGAACTGA	GATACCTACA	1620
GCGTGAGCTA	TGAGAAAGCG	CCACGCTTCC	CGAAGGGAGA	AAGGCGGACA	GGTATCCGGT	1680
AAGCGGCAGG	GTCGGAACAG	GAGAGCGCAC	GAGGGAGCTT	CCAGGGGGAA	ACGCCTGGTA	1740
TCTTTATAGT	CCTGTCGGGT	TTCGCCACCT	CTGACTTGAG	CGTCGATTTT	TGTGATGCTC	1800
GTCAGGGGGG	CGGAGCCTAT	GGAAAAACGC	CAGCAACGCG	GCCTTTTTAC	GGTTCCTGGC	1860
CTTTTGCTGG	CCTTTTGCTC	ACATGTTCTT	TCCTGCGTTA	TCCCCTGATT	CTGTGGATAA	1920
CCGTATTACC	GCCTTTGAGT	GAGCTGATAC	CGCTCGCCGC	AGCCGAACGA	CCGAGCGCAG	1980
CGAGTCAGTG	AGCGAGGAAG	CGGAAGAGCG	CCTGATGCGG	TATTTTCTCC	TTACGCATCT	2040
GTGCGGTATT	TCACACCGCA	TATGGTGCAC	TCTCAGTACA	ATCTGCTCTG	ATGCCGCATA	2100
GTTAAGCCAG	TATACACTCC	GCTATCGCTA	CGTGACTGGG	TCATGGCTGC	GCCCCGACAC	2160
CCGCCAACAC	CCGCTGACGC	GCCCTGACGG	GCTTGTCTGC	TCCCGGCATC	CGCTTACAGA	2220
CAAGCTGTGA	CCGTCTCCGG	GAGCTGCATG	TGTCAGAGGT	TTTCACCGTC	ATCACCMAAA	2280
CGCGCGAGGC	AGCATGCATC	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAACCTCGCC	2340
CATCCCGCCC	CTAACTCCGC	CCAGTTCCGC	CCATTCTCCG	CCCCATGGCT	GACTAATTTT	2400
TTTTATTTAT	GCAGAGGCCG	AGGCCGCCTC	GGCCTCTGAG	CTATTCCAGA	AGTAGTGAGG	2460
AGGCTTTTTT	GGAGGCCTAG	GCTTTTGCAA	AAAGCTAGCT	TACAGCTCAG	GGCTGCGATT	2520
TCGCGCCAAA	CTTGACGGCA	ATCCTAGCGT	GAAGGCTGGT	AGGATTTTAT	CCCCGCTGCC	2580
ATCATGGTTC	GACCATTGAA	CTGCATCGTC	GCCGTGTCCC	AAAATATGGG	GATTGGCAAG	2640
AACGGAGACC	TACCCTGGCC	TCCGCTCAGG	AACGAGTTCA	AGTACTTCCA	AAGAATGACC	2700
ACAACCTCTT	CAGTGGAAGG	TAAACAGAAT	CTGGTGATTA	TGGGTAGGAA	AACCTGGTTC	2760

TCCATTCTG	AGAAGAATCG	ACCTTTAAAG	GACAGAATTA	ATATAGTTCT	CAGTAGAGAA	2820
CTCAAAGAAC	CACCACGAGG	AGCTCATTTT	CTTGCCAAAA	GTTTGGATGA	TGCCTTAAGA	2880
CTTATTGAAC	AACCGGAATT	GGCAAGTAAA	GTAGACATGG	TTTGGATAGT	CGGAGGCAGT	2940
TCTGTTTACC	AGGAAGCCAT	GAATCAACCA	GGCCACCTCA	GACTCTTTGT	GACAAGGATC	3000
ATGCAGGAAT	TTGAAAGTGA	CACGTTTTTC	CCAGAAATTG	ATTTGGGGAA	ATATAAACTT	3060
CTCCCAGAAT	ACCCAGGCGT	CCTCTCTGAG	GTCCAGGAGG	AAAAAGGCAT	CAAGTATAAG	3120
TTTGAAGTCT	ACGAGAAGAA	AGACTAACAG	GAAGATGCTT	TCAAGTTCTC	TGCTCCCCTC	3180
CTAAAGCTAT	GCATTTTTAT	AAGACCATGG	GACTTTTGCT	GGCTTTAGAT	CTTTGTGAAG	3240
GAACCTTACT	TCTGTGGTGT	GACATAATTG	GACAACTAC	CTACAGAGAT	TTAAAGCTCT	3300
AAGGTAAATA	TAAAATTTTT	AAGTGTATAA	TGTGTTAAAC	TACTGATTCT	AATTGTTTGT	3360
GTATTTTAGA	TTCCAACCTA	TGGAAGTGAT	GAATGGGAGC	AGTGGTGGAA	TGCCTTTAAT	3420
GAGGAAAACC	TGTTTTGCTC	AGAAGAAATG	CCATCTAGTG	ATGATGAGGC	TACTGCTGAC	3480
TCTCAACATT	CTACTCCTCC	AAAAAAGAAG	AGAAAGGTAG	AAGACCCCAA	GGACTTTCCT	3540
TCAGAATTGC	TAAGTTTTTT	GAGTCATGCT	GTGTTTAGTA	ATAGAACTCT	TGCTTGCTTT	3600
GCTATTTACA	CCACAAAGGA	AAAAGCTGCA	CTGCTATACA	AGAAAATTAT	GGAAAAATAT	3660
TCTGTAACCT	TTATAAGTAG	GCATAACAGT	TATAATCATA	ACATACTGTT	TTTTCTTACT	3720
CCACACAGGC	ATAGAGTGTC	TGCTATTAAT	AACTATGCTC	AAAAATTGTG	TACCTTTAGC	3780
TTTTTAATTT	GTAAAGGGGT	TAATAAGGAA	TATTTGATGT	ATAGTGCCTT	GACTAGAGAT	3840
CATAATCAGC	CATACCACAT	TTGTAGAGGT	TTTACTTGCT	TTAAAAAACC	TCCCACACCT	3900
CCCCCTGAAC	CTGAAACATA	AAATGAATGC	AATTGTTGTT	GTTAACCTGT	TTATTGCAGC	3960
TTATAATGGT	TACAAATAAA	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTC	4020
ACTGCATTCT	AGTTGTGGTT	TGTCCAAACT	CATCAATGTA	TCTTATCATG	TCTGGATCTA	4080
ATAAAAGATA	TTTATTTTCA	TTAGATATGT	GTGTTGGTTT	TTTGTGTGCA	GTGCCTCTAT	4140
CTGGAGGCCA	GGTAGGGCTG	GCCTTGGGGG	AGGGGGAGGC	CAGAATGACT	CCAAGAGCTA	4200
CAGGAAGGCA	GGTCAGAGAC	CCCACTGGAC	AAACAGTGGC	TGGACTCTGC	ACCATAACAC	4260
ACAATCAACA	GGGGAGTGAG	CTGGAAATTT	GCTAGCGAAT	TCCAGCACAC	TGGCGGCCGT	4320
TACTAGTTAT	TAATAGTAAT	CAATTACGGG	GTCATTAGTT	CATAGCCCAT	ATATGGAGTT	4380
CCGCGTTACA	TAACTTACGG	TAAATGGCCC	GCCTGGCTGA	CCGCCCAACG	ACCCCGCCCC	4440
ATTGACGTCA	ATAATGACGT	ATGTTCCCAT	AGTAACGCCA	ATAGGGACTT	TCCATTGACG	4500
TCAATGGGTG	GAGTATTTAC	GGTAAACTGC	CCACTTGGCA	GTACATCAAG	TGTATCATAT	4560

GCCAAGTACG	CCCCCTATTG	ACGTCAATGA	CGGTAAATGG	CCCGCCTGGC	ATTATGCCCA	4620
GTACATGACC	TTATGGGACT	TTCCTACTTG	GCAGTACATC	TACGTATTAG	TCATCGCTAT	4680
TACCATGGTG	ATGCGGTTTT	GGCAGTACAT	CAATGGGCGT	GGATAGCGGT	TTGACTCACG	4740
GGGATTTC	AGTCTCCACC	CCATTGACGT	CAATGGGAGT	TTGTTTTGGC	ACCAAAATCA	4800
ACGGGACTTT	CCAAAATGTC	GTAACAAC	CGCCCCATTG	ACGCAAATGG	GCGGTAGGCG	4860
TGTACGGTGG	GAGGTCTATA	TAAGCAGAGC	TCGTTTAGTG	AACCGTCAGA	TCGCCTGGAG	4920
ACGCCATCCA	CGCTGTTTTG	ACCTCCATAG	AAGACACCGG	GACCGATCCA	GCCTCCGCGG	4980
CCGGGAACGG	TGCATTGGAA	CGCGGATTCC	CCGTGCCAAG	AGTGACGTAA	GTACCGCCTA	5040
TAGAGTCTAT	AGGCCCCACCC	CCTTGCTTTC	TTATGCATGC	TATACTGTTT	TTGGCTTGGG	5100
GTCTATACAC	CCCCGCTTCC	TCATGTTATA	GGTGATGGTA	TAGCTTAGCC	TATAGGTGTG	5160
GGTTATTGAC	CATTATTGAC	CACTCCCCTA	TTGGTGACGA	TACTTTCCAT	TACTAATCCA	5220
TAACATGGCT	CTTTGCCACA	ACTCTCTTTA	TTGGCTATAT	GCCAATACAC	TGTCCTTCAG	5280
AGACTGACAC	GGACTCTGTA	TTTTTACAGG	ATGGGGTCTC	ATTTATTATT	TACAAATTCA	5340
CATATACAAC	ACCACCGTCC	CCAGTGCCCC	CAGTTTTTAT	TAAACATAAC	GTGGGATCTC	5400
CACGCGAATC	TCGGGTACGT	GTTCCGGACA	TGGGCTCTTC	TCCGGTAGCG	GCGGAGCTTC	5460
TACATCCGAG	CCCTGCTCCC	ATGCCTCCAG	CGACTCATGG	TCGCTCGGCA	GCTCCTTGCT	5520
CCTAACAGTG	GAGGCCAGAC	TTAGGCACAG	CACGATGCCC	ACCACCACCA	GTGTGCCGCA	5580
CAAGGCCGTG	GCGGTAGGGT	ATGTGTCTGA	AAATGAGCTC	GGGGAGCGGG	CTTGACCGC	5640
TGACGCATTT	GGAAGACTTA	AGGCAGCGGC	AGAAGAAGAT	GCAGGCAGCT	GAGTTGTTGT	5700
GTTCTGATAA	GAGTCAGAGG	TAACCTCCGT	TGCGGTGCTG	TTAACGGTGG	AGGGCAGTGT	5760
AGTCTGAGCA	GTACTCGTTG	CTGCCGCGCG	CGCCACCAGA	CATAATAGCT	GACAGACTAA	5820
CAGACTGTTC	CTTCCATGG	GTCTTTTCTG	CAGTCACCGT	CCTTGACACG	CGTCTCGGGA	5880
AGCTTGCCGC	CACCATGGAC	TGGACCTGGC	GCGTGTTTTG	CCTGCTCGCC	GTGGCTCCTG	5940
GGGCCCACAG	CCAGGTGCAA	CTGGTGCA	CCGGCGCCGA	AGTGAAGAAA	CCCGGTGCTT	6000
CCGTGAAAGT	CAGCTGTAAA	ACTAGTAGAT	ACACCTTCAC	TGAATACACC	ATACACTGGG	6060
TTAGACAGGC	CCCTGGCCAA	AGGCTGGAGT	GGATAGGAGG	TATTAATCCT	AACAATGGTA	6120
TTCCTAACTA	CAACCAGAAG	TTCAAGGGCC	GGGCCACCTT	GACCGTAGGC	AAGTCTGCCA	6180
GCACCGCCTA	CATGGA	TCCAGCCTGC	GCTCCGAGGA	CACTGCAGTC	TACTACTGCG	6240
CCAGAAGAAG	AATCGCCTAT	GGTTACGACG	AGGGCCATGC	TATGGACTAC	TGGGGTCAAG	6300
GAACCCTTGT	CACCGTCTCC	TCAGGTGAGT	GGATCCTCTG	CGCCTGGGCC	CAGCTCTGTC	6360
CCACACCGCG	GTCACATGGC	ACCACCTCTC	TTGCAGCCTC	CACCAAGGGC	CCATCGGTCT	6420

```

TCCCCCTGGC ACCCTCCTCC AAGAGCACCT CTGGGGGCAC AGCGGCCCTG GGCTGCCTGG      6480
TCAAGGACTA CTTCCCCGAA CCGGTGACGG TGTCGTGGAA CTCAGGCGCC CTGACCAGCG      6540
GCGTGACAC CTTCCCGGCT GTCCTACAGT CCTCAGGACT CTACTCCCTC AGCAGCGTGG      6600
TGACCGTGCC CTCCAGCAGC TTGGGCACCC AGACCTACAT CTGCAACGTG AATCACAAGC      6660
CCAGCAACAC CAAGGTGGAC AAGAAAGTTG AGCCCAAATC TTGTGACAAA ACTCACACAT      6720
GCCCACCGTG CCCAGCACCT GAACTCCTGG GGGGACCGTC AGTCTTCCTC TTCCCCCCTA      6780
AACCCAAGGA CACCCTCATG ATCTCCCGGA CCCCTGAGGT CACATGCGTG GTGGTGGACG      6840
TGAGCCACGA AGACCTGAG GTCAAGTTCA ACTGGTACGT GGACGGCGTG GAGGTGCATA      6900
ATGCCAAGAC AAAGCCGCGG GAGGAGCAGT ACAACAGCAC GTACCGGGTG GTCAGCGTCC      6960
TCACCGTCCT GCACCAGGAC TGGCTGAATG GCAAGGAGTA CAAGTGCAAG GTCTCCAACA      7020
AAGCCCTCCC AGCCCCCATC GAGAAAAACCA TCTCCAAAGC CAAAGGGCAG CCCCAGAAC      7080
CACAGGTGTA CACCCTGCCC CCATCCCGGG AGGAGATGAC CAAGAACCAG GTCAGCCTGA      7140
CCTGCCTGGT CAAAGGCTTC TATCCCAGCG ACATCGCCGT GGAGTGGGAG AGCAATGGGC      7200
AGCCGGAGAA CAACTACAAG ACCACGCCTC CCGTGCTGGA CTCCGACGGC TCCTTCTTCC      7260
TCTACAGCAA GCTCACCGTG GACAAGAGCA GGTGGCAGCA GGGGAACGTC TTCTCATGCT      7320
CCGTGATGCA TGAGGCTCTG CACAACCACT ACACGCAGAA GAGCCTCTCC CTGTCTCCGG      7380
GTAAATGAGT GCGACGGCCG GCAAGCCCCG CTCCCCGGGC TCTCGCGGTC GCACGAGGAT      7440
GCTTGGCACG TACCCCTGT ACATACTTCC CGGGCGCCCA GCATGGAAAT AAAGCACCGG      7500
ATCTAATAAA AGATATTTAT TTTCATTAGA TATGTGTGTT GGTTTTTTGT GTGCAGTGCC      7560
TCTATCTGGA GGCCAGGTAG GGCTGGCCTT GGGGGAGGGG GAGGCCAGAA TGA CTCCAAG      7620
AGCTACAGGA AGGCAGGTCA GAGACCCAC TGGACAAACA GTGGCTGGAC TCTGCACCAT      7680
AACACACAAT CAACAGGGGA GTGAGCTGGA AATTTGCTAG CGAATTAATT C      7731

```

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 472 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

Met Asp Trp Thr Trp Arg Val Phe Cys Leu Leu Ala Val Ala Pro Gly

46

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

47

Asp Trp Leu Asn Gly Lys Glu Tyr Lys Cys Lys Val Ser Asn Lys Ala
 340 345 350
 Leu Pro Ala Pro Ile Glu Lys Thr Ile Ser Lys Ala Lys Gly Gln Pro
 355 360 365
 Arg Glu Pro Gln Val Tyr Thr Leu Pro Pro Ser Arg Glu Glu Met Thr
 370 375 380
 Lys Asn Gln Val Ser Leu Thr Cys Leu Val Lys Gly Phe Tyr Pro Ser
 385 390 395 400
 Asp Ile Ala Val Glu Trp Glu Ser Asn Gly Gln Pro Glu Asn Asn Tyr
 405 410 415
 Lys Thr Thr Pro Pro Val Leu Asp Ser Asp Gly Ser Phe Phe Leu Tyr
 420 425 430
 Ser Lys Leu Thr Val Asp Lys Ser Arg Trp Gln Gln Gly Asn Val Phe
 435 440 445
 Ser Cys Ser Val Met His Glu Ala Leu His Asn His Tyr Thr Gln Lys
 450 455 460
 Ser Leu Ser Leu Ser Pro Gly Lys
 465 470

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

ACCGTCTCCT CAGGTGAGTG GATCC

25

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

CCTCTCTTGC AGCC

14

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 14 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

CCTCTCTTGC AGCC

14

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Thr Val Ser Ser
1

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Ser Thr Lys Gly
1

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

ACCGTCTCCT CAGCCTCCAC CAAGGGC

27

(2) INFORMATION FOR SEQ ID NO: 50:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 8 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Thr Val Ser Ser Ser Thr Lys Gly
1 5

(2) INFORMATION FOR SEQ ID NO: 51:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 27 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

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

ACCGTCTCCT CAGCCTCCAC CAAGGGC

27

(2) INFORMATION FOR SEQ ID NO: 52:

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

(ii) MOLECULE TYPE: peptide

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

Thr Val Ser Ser Ala Ser Thr Lys Gly

1

5

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

GAAATAAAAC GTGAGTGGAT CC

22

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

CTTCTTTCCT CAGGAACTGT GGCTGCA

27

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 4 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Thr Val Ala Ala

1

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

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

GAAATAAAAC GAACTGTGGC TGCA

24

(2) INFORMATION FOR SEQ ID NO: 57:

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

(ii) MOLECULE TYPE: peptide

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

Glu Ile Lys Thr Val Ala Ala
1 5

(2) INFORMATION FOR SEQ ID NO: 58:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 24 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

GAAATAAAAC GAACTGTGGC TGCA

24

(2) INFORMATION FOR SEQ ID NO: 59:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 8 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

Glu Ile Lys Arg Thr Val Ala Ala

1 5

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

Met	Asp	Ser	Gln	Ala	Gln	Val	Leu	Met	Leu	Leu	Leu	Leu	Trp	Val	Ser
1				5				10						15	
Gly Thr Cys Gly															
20															

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

Met	Gly	Trp	Ser	Trp	Val	Phe	Leu	Phe	Leu	Leu	Ser	Gly	Thr	Ala	Gly
1				5				10						15	
Val Leu Ser															

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

GCCGCCACC

53

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

CAGAAAGCTT GCCGCCACCA TGGATTCACA GGCCCAG

37

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

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

Met Asp Ser Gln Ala Gln
1 5

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 35 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

CCGAGGATCC ACTCACGTTT CAGCTCCAGC TTGGT

35

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

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

CAGAAAGCTT GCCGCCACCA TGGGATGGAG CTGGGTC

37

- (2) INFORMATION FOR SEQ ID NO: 67:

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

- (ii) MOLECULE TYPE: peptide

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

Met Gly Trp Ser Trp Val
1 5

- (2) INFORMATION FOR SEQ ID NO: 68:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 35 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

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

CCGAGGATCC ACTCACCTGA GGAGACGGTG ACTGA

35

- (2) INFORMATION FOR SEQ ID NO: 69:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 36 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
 - (A) DESCRIPTION: /desc = "PRIMER"

55

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

GTTCATCACAA TGTCTCCGGA GGAACCTGGA ACCCAG

36

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CTCCGGAGAC ATTGTGATGA CCCAATCTC

29

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:

CTCCGGAGAC ATTGTGATGA CCCAATCTC

29

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CAGTCAGAGC CTTTTATATT CTAGAAATCA AAAGAACTAC TTGGCCTGGT ATCAGCAGAA

60

ACCAGGACAG CC

72

(2) INFORMATION FOR SEQ ID NO: 73:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

ACCCCAGATT CCCTAGTGCT AGCCCAAAG ATGAGGAGTT TGGG

44

(2) INFORMATION FOR SEQ ID NO: 74:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 67 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

TAGCACTAGG GAATCTGGGG TACCTGATAG GTTCAGTGGC AGTGGGTTTG GGACAGACTT

60

CACCCTC

67

(2) INFORMATION FOR SEQ ID NO: 75:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

GTCCCTTGTC CGAACGTGAG CGGATAGCTA AAATATTGCT GACAGTAATA AAC

53

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

GCTCACGTTC GGACAAGGGA CCAAGGTGGA AAT

33

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 72 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CAGTCAGAGC CTTTTATATT CTAGAAATCA AAAGAACTAC TTGGCCTGGT TCCAGCAGAA

60

ACCAGGACAG CC

72

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 57 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

GTCCCTTGTC CGAACGTGAG CGGATAGCTA AAATATTGCT GACAGTCATA AACTGCC

57

(2) INFORMATION FOR SEQ ID NO: 79:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 34 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CCCAAACCTCC TCATCTATTG GGCTAGCACT AGGG

34

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CCCTAGTGCT AGCCCAATAG ATGAGGAGTT TGGG

34

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:

TACGCAAACC GCCTCTC

17

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

GAGTGCACCA TATGCGGT

18

59

(2) INFORMATION FOR SEQ ID NO: 83:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

AACAGCTATG ACCATG

16

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

GTTTCCCGAG TCACGAC

17

(2) INFORMATION FOR SEQ ID NO: 85:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 47 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

GTGTATTCAG TGAAGGTGTA TCTACTAGTT TTACAGCTGA CTTTCAC

47

(2) INFORMATION FOR SEQ ID NO: 86:

(i) SEQUENCE CHARACTERISTICS:

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

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

TAGTAGATAC ACCTTCACTG AATACACCAT ACACTGGGTT AGACAGGCCC CTG 53

- (2) INFORMATION FOR SEQ ID NO: 87:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

CCCTTGAAC TCTGGTTGTA GTTAGGAATA CCATTGTTAG GATTAATACC TCCTATCCAC 60

TCCAGCCTTT G 71

- (2) INFORMATION FOR SEQ ID NO: 88:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 71 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

TAACTACAAC CAGAAGTTCA AGGGCCGGGC CACCTTGACC GTAGGCAAGT CTGCCAGCAC 60

CGCCTACATG G 71

- (2) INFORMATION FOR SEQ ID NO: 89:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

GCATGGCCCT CGTCGTAACC ATAGGCGATT CTTCTTCTGG CGCAGTAGTA GACTGCAGTG 60
TCC 63

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CTATGGTTAC GACGAGGGCC ATGCTATGGA CTACTGGGGT CAAGGAAC 48

(2) INFORMATION FOR SEQ ID NO: 91:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 71 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

TAACTACAAC CAGAAGTTCA AGGGCCGGGT CACCATCACC GTAGACACCT CTGCCAGCAC 60
CGCCTACATG G 71

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 27 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

GGACACTGCA GTCTACTTCT GCGCCAG

27

(2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

TACGCAAACC GCCTCTC

17

(2) INFORMATION FOR SEQ ID NO: 94:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:

GAGTGCACCA TATGCGGT

18

(2) INFORMATION FOR SEQ ID NO: 95:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 76 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "PRIMER"

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

CCTTTGGCCA GGGGCCTGTC TAACCCAGTG TATGGTGTAT TCAGTGAAGG TGCTATCCAC

60

TAGTTTCCAC TAGTTT

76

63

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

GTCACCGTCC TTGACACGCG TCTCGGGA

28

(2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

TTGGAGGAGG GTGCCAG

17

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "PRIMER"

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

GAGACATTGT GACCCAATCT CC

22

(2) INFORMATION FOR SEQ ID NO: 99:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

GACAGTCATA AACTGCCACA TCTTC

25

- (2) INFORMATION FOR SEQ ID NO: 100:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

TTGACACGCG TCTCGGAAG CTT

23

- (2) INFORMATION FOR SEQ ID NO: 101:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 22 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "PRIMER"

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

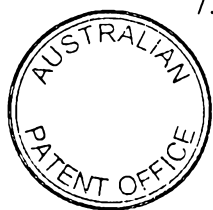
GGCGCAGAGG ATCCACTCAC CT

22

- 115 -

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An antibody protein having the complementary determining regions of the monoclonal antibody F19 (ATCC Accession No. HB 8269), said antibody protein specifically binding to fibroblast activation protein, characterised in that it has framework modifications resulting in the improved producibility in host cells as compared to a chimeric antibody having the variable regions of F19 and foreign constant regions wherein the amino acid in Kabat position 87 of the light chain region is not asparagine.
2. An antibody protein characterised in that it has a variable light chain region and a variable heavy chain region according to claim 1, each joined to a human constant region.
3. The antibody protein of claim 2, wherein said human constant region of the light chain is a human kappa constant region.
4. The antibody protein of claim 2, wherein said human constant region of the heavy chain is a human gamma-1 constant region.
5. An antibody protein according to any one of claim 1 to 4, characterised in that its expression level in crude media samples as determined by ELISA and/or purified antibody yield exceeds the expression level and/or purification yield of the chimeric antibody without framework modifications by at least a factor of 10.
6. An antibody protein according to any one of claim 1 to 4, characterised in that its expression level in crude media samples as determined by ELISA and/or purified antibody yield exceeds the expression level and/or purification yield of the chimeric antibody without framework modifications by at least a factor of 20.
7. An antibody protein according to any one of claim 1 to 4, characterised in that its



- 116 -

expression level in crude media samples as determined by ELISA and/or purified antibody yield exceeds the expression level and/or purification yield of the chimeric antibody without framework modifications by at least a factor of 100.

5 8. An antibody protein according to any one of claims 1 to 7, characterised in that it displays improved producibility in eukaryotic cells.

9. The antibody protein according to claim 8 wherein said eukaryotic cell is a Chinese hamster ovary cell (CHO cell).

10

10. The antibody protein of any one of claims 1 to 9, wherein the amino acid in Kabat position 87 of the light chain region is selected from aromatic or aliphatic amino acids.

15 11. The antibody protein of claim 10, wherein said aromatic amino acid in Kabat position 87 of the light chain region is a tyrosine or phenylalanine.

20 12. The antibody protein according to any one of claims 1 to 11, wherein the amino acid in Kabat position 36 of the light chain region is selected from aromatic amino acids.

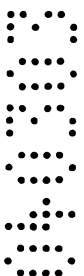
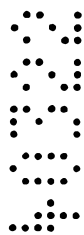
13. An antibody protein according to any one of claims 1 to 12 that contains the variable region of the light chain as set forth in SEQ ID NO: 2.

25 14. An antibody protein of claim 13 characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1.

15. An antibody protein according to any one of claims 1 to 12 that contains the variable region of the light chain as set forth in SEQ ID NO: 6.

30

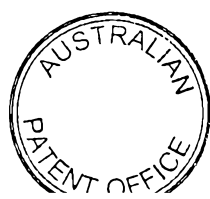
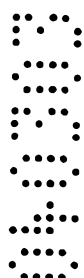
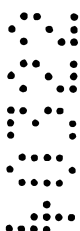
16. An antibody protein of claim 15 characterised in that the variable region of the light



- 117 -

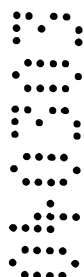
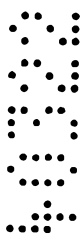
chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 5.

17. An antibody protein according to any one of claims 1 to 16 containing a variable region of the heavy chain as set forth in any one of SEQ ID NOS: 8, 10, 12, 14.
- 5 18. An antibody protein according to claim 17 characterised in that the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NOS: 7, 9, 11, 13.
- 10 19. An antibody protein according to any one of claims 1 to 13 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NO: 12.
- 15 20. The antibody protein of claim 19 characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 11.
- 20 21. An antibody protein according to claim 19 or claim 20 containing the constant region of the light chain as set forth in SEQ ID NO: 20 and the constant region of the heavy chain as set forth in SEQ ID NO: 22.
- 25 22. An antibody protein according to any one of claims 1 to 12, 17 or 18 containing the variable region of the light chain as set forth in SEQ ID NO: 2 and the variable region of the heavy chain as set forth in SEQ ID NO: 8.
- 30 23. The antibody protein of claim 22 characterised in that the variable region of the light chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 1 and the variable region of the heavy chain is encoded by a nucleotide sequence as set forth in SEQ ID NO: 7.



- 118 -

24. A nucleotide sequence encoding an antibody protein according to any one of claims 1 to 23.
25. A recombinant DNA vector that contains a nucleotide sequence of claim 24.
26. The recombinant DNA vector of claim 25, said vector being an expression vector.
27. A host cell carrying a vector according to claims 25 or 26.
28. The host cell of claim 27, wherein said host cell is a eukaryotic cell.
29. The host cell of claim 28, wherein said eukaryotic host cell is a mammalian cell.
30. The host cell of claim 29, wherein said host cell is a CHO or a COS cell.
31. A method of producing antibody proteins according to any one of claims 1 to 23, said method comprising the steps of:
 - (a) cultivating a host cell according to any one of claims 27 to 30 under conditions where said antibody protein is expressed by said host cell, and
 - (b) isolating said antibody protein.
32. The method of claim 31, wherein said host cell is a mammalian cell, preferably a CHO or COS cell.
33. The method of claim 31 or 32, wherein said host cell is cotransfected with two plasmids carrying the expression units for light and heavy chains respectively.
34. An antibody protein according to any one of claims 1 to 23, wherein said antibody protein is conjugated to a therapeutic agent.
35. The antibody protein of claim 34, wherein said therapeutic agent is a therapeutic



- 119 -

agent selected from the group consisting of radioisotopes, toxins, toxoids, inflammatory agents and chemotherapeutic agents.

36. The antibody protein of claim 35, wherein said radioisotope is a β -emitting radioisotope.

37. The antibody protein of claim 36, wherein said radioisotope is selected from the group consisting of $^{186}\text{Rhenium}$, $^{188}\text{Rhenium}$, $^{131}\text{Iodine}$ and $^{90}\text{Yttrium}$.

38. An antibody protein according to any one of claims 1 to 23, characterised in that it is labelled.

39. The antibody protein of claim 38, wherein said label is a detectable marker.

40. The antibody protein of claim 39, wherein the detectable marker is a detectable marker selected from the group consisting of enzymes, dyes, radioisotopes, digoxigenin, and biotin.

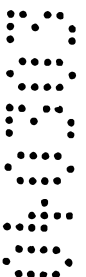
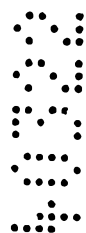
41. An antibody protein according to any one of claims 1 to 23 conjugated to an imageable agent.

42. The antibody protein of claim 41, wherein the imageable agent is a radioisotope.

43. The antibody protein of claim 42, wherein said radioisotope is a γ -emitting radioisotope.

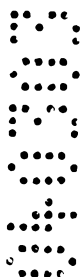
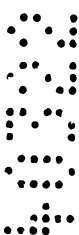
44. The antibody protein of claim 43, wherein said radioisotope is ^{125}I .

45. A pharmaceutical composition containing an antibody protein according to any one of claims 1 to 23 and a pharmaceutically acceptable carrier.



- 120 -

46. A pharmaceutical composition containing an antibody protein according to any one of claims 34 to 37 and a pharmaceutically acceptable carrier.
47. A pharmaceutical composition containing an antibody protein according to any one of claims 41 to 44 and a pharmaceutically acceptable carrier.
48. The pharmaceutical composition of claims 45 to 47, for use in the treatment of imaging of tumors, wherein said tumors are associated with activated stromal fibroblasts, preferably wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancers, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.
49. An antibody according to any one of claims 1 to 23 or 34 to 37 for use in therapy.
50. An antibody according to any one of claims 38 to 44 for use in diagnosis.
51. An antibody protein according to any one of claims 1 to 23 for use in the treatment of cancer.
52. An antibody protein according to any one of claims 34 to 37 for use in the treatment of cancer.
53. An antibody protein according to any one of claims 41 to 44 for use in imaging activated stromal fibroblasts.
54. Use of an antibody protein according to any one of claims 38 to 40 for detecting the presence of activated stromal fibroblasts in a sample.
55. An antibody protein according to any one of claims 38 to 40 for use in detecting the presence of activated stromal fibroblasts in a tumor.



56. A method of treating tumors, wherein the tumor is associated with activated stromal fibroblasts capable of specifically forming a complex with antibody proteins according to any one of claims 1 to 23, or 34 to 37, which comprises
5 contacting the tumor with an amount of said antibody proteins effective to treat the tumor.

57. Use of an antibody protein according to any one of claims 1 to 23, or 34 to 37 for the manufacture of a medicament for the treatment of tumors wherein said tumor is
10 associated with activated stromal fibroblasts capable of specifically forming a complex with said antibody protein.

58. The method of claim 56 or use of claim 57, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancer, head and neck cancers, ovarian cancers, lung
15 cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.

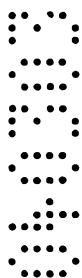
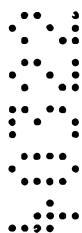
59. The method of claim 56 or the use of claim 57, wherein the contacting is effected *in vitro*.
20

60. The method of claim 56 or the use of claim 57, wherein the contacting is effected *in vivo*.

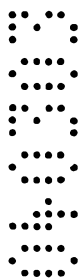
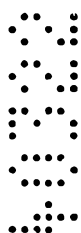
61. A method of detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor, characterised in that:
25

(a) a sample, possibly containing activated stromal fibroblasts, is contacted with an antibody protein according to any one of claims 1 to 23 or 38 to 40 under conditions suitable for the formation of a complex between said antibody and antigen,

30 (b) detecting the presence of said complex, thereby detecting the presence of activated stromal fibroblasts in wound healing, inflammation or a tumor.



62. The method of claim 61, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancers, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.
63. The method of claim 61 or 62, wherein the antibody protein is a protein according to any one of claims 38 to 40.
64. A method of imaging the presence of activated stromal fibroblasts in a healing wound, inflamed skin or a tumor, in a human patient, characterised in that:
- (a) an antibody protein according to any one of claims 1 to 23 conjugated to an imageable agent is administered to a human patient under conditions suitable for the formation of an antibody-antigen complex,
 - (b) imaging any complex formed in this manner.
65. The method of claim 64, wherein the tumor is a tumor having cancer cells selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancers, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.
66. A method of detecting tumor-stroma, characterised in that:
- (a) a suitable sample is contacted with an antibody protein according to any one of claims 1 to 23, under conditions suitable for the formation of an antibody-antigen complex,
 - (b) detecting the presence of any complex so formed,
 - (c) relating the presence of said complex to the presence of tumor-stroma.
67. The method of claim 66, wherein said antibody is labelled with a detectable marker.



- 123 -

68. A method of imaging tumor-stroma in a human patient, which comprises:

(a) administering to the patient an antibody protein according to any one of claims 41 to 44, under conditions suitable for the formation of an antibody-antigen complex,

5 (b) imaging any complex so formed, and thereby imaging the presence of tumor-stroma in a human patient.

69. An antibody protein containing an amino acid sequence as set forth in SEQ ID NO: 2.

10

70. An antibody protein according to claim 69 further containing an amino acid sequence as set forth in SEQ ID NO: 12.

15

71. An antibody protein according to claim 69 or 70 further containing an amino acid sequence as set forth in SEQ ID NO: 20 and an amino acid sequence as set forth in SEQ ID NO: 22.

72. A DNA molecule coding for an antibody protein according to any one of claims 69 to 71.

20

73. A host cell carrying a DNA molecule according to claim 72.

74. A method of producing an antibody protein of any one of claims 69 to 71, said method comprising the steps of

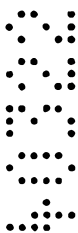
25

(a) cultivating the host cell of claim 73 under conditions where said antibody protein is expressed by said host cell and

(b) isolating said protein.

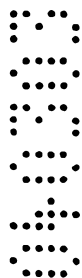
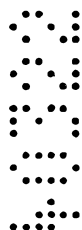
75. An antibody protein according to any one of claims 69 to 71 which is conjugated to a radioisotope, preferably ^{131}I , ^{125}I , ^{186}Re , ^{188}Re , or ^{90}Y .

30



- 124 -

76. A pharmaceutical composition comprising an antibody protein according to any one of claims 69 to 71, or 75, and a pharmaceutically acceptable carrier.
- 5 77. Use of an antibody protein according to any one of claims 1 to 23, or 34 to 37 for the manufacture of a medicament for the treatment of cancer.
78. Use of an antibody protein according to any one of claims 41 to 44 for the manufacture of a diagnostic composition for imaging activated stromal fibroblasts.
- 10 79. Use of an antibody protein according to any one of claims 38 to 40 for the manufacture of a diagnostic composition for detecting the presence of activated stromal fibroblasts in a sample.
- 15 80. The use of pharmaceutical composition of claims 45 to 47 in the manufacture of a medicament for the treatment of tumors wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancers, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancer of the brain.
- 20 81. The use of the pharmaceutical composition of claims 45 to 47 in the manufacture of a diagnostic imaging composition for the imaging of tumors wherein said tumors are tumors selected from the cancer group consisting of colorectal cancers, non-small cell lung cancers, breast cancers, head and neck cancers, ovarian cancers, lung cancers, bladder cancers, pancreatic cancers and metastatic cancers of the brain.
- 25



- 125 -

82. An antibody protein, composition, use, method or DNA molecule substantially as described herein with reference to Figures and/or Examples.

DATED this 20th day of February 2003

5

Boehringer Ingelheim International GmbH

by Davies Collison Cave

Patent Attorneys for the Applicant(s)

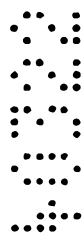


Fig. 1

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGTATCAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 2

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSLL	YSRNQKNYLA	WYQQKPGQPP
51	61	71	81	91
KLLIFWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYYCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 3

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGT <u>T</u> CAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTTTTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTAT <u>G</u> ACT	GTCA <u>A</u> CAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 4

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSL	YSRNQKNYLA	WFQQKPGQPP
51	61	71	81	91
KLLIFWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYDCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 5

1	11	21	31	41
GACATTGTGA	TGACCCAATC	TCCAGACTCT	TTGGCTGTGT	CTCTAGGGGA
51	61	71	81	91
GAGGGCCACC	ATCAACTGCA	AGTCCAGTCA	GAGCCTTTTA	TATTCTAGAA
101	111	121	131	141
ATCAAAAGAA	CTACTTGGCC	TGGTATCAGC	AGAAACCAGG	ACAGCCACCC
151	161	171	181	191
AAACTCCTCA	TCTATTGGGC	TAGCACTAGG	GAATCTGGGG	TACCTGATAG
201	211	221	231	241
GTTTCAGTGGC	AGTGGGTTTG	GGACAGACTT	CACCCTCACC	ATTAGCAGCC
251	261	271	281	291
TGCAGGCTGA	AGATGTGGCA	GTTTATTACT	GTCAGCAATA	TTTTAGCTAT
301	311	321	331	339
CCGCTCACGT	TCGGACAAGG	GACCAAGGTG	GAAATAAAA	

Fig. 6

1	11	21	31	41
DIVMTQSPDS	LAVSLGERAT	INCKSSQSL	YSRNQKNYLA	WYQQKPGQPP
51	61	71	81	91
KLLIYWASTR	ESGVPDRFSG	SGFGTDFTLT	ISSLQAEDVA	VYYCQQYFSY
101	111			
PLTFGQGTKV	EIK			

Fig. 7

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351
AACCCTTGTC ACCGTCTCCT CA
372

```

Fig. 8

```

1      11      21      31      41
QVQLVQSGAE VKKPGASVKV SCKTSRYTFT EYTIHWVRQA PGQRLEWIGG
51      61      71      81      91
INPNNGIPNY NQKFKGRATL TVGKSASTAY MELSSLRSED TAVYYCARRR
101     111     121-124
IAYGYDEGHA MDYWGQGTLV TVSS

```

Fig. 9

```

1
CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
51
CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
101
TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
151
ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
201
GGCCACCTTG ACCGTAGGCA AGTCTGCCAG CACCGCCTAC ATGGAAGTGT
251
CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA
301
ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
351
AACCCTTGTC ACCGTCTCCT CA
372

```

Fig. 10

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SCKTSRYTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKFKGRATL	TVGKSASTAY	MELSSLRSED	TAVY <u>F</u> CARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGTLV	TVSS		

Fig. 11

1				
CAGGTGCAAC	TAGTGCAGTC	CGGCGCCGAA	GTGAAGAAAC	CCGGTGCTTC
51				
CGTGAAAGTC	AGCTGTAAAA	CTAGTAGATA	CACCTTCACT	GAATACACCA
101				
TACACTGGGT	TAGACAGGCC	CCTGGCCAAA	GGCTGGAGTG	GATAGGAGGT
151				
ATTAATCCTA	ACAATGGTAT	TCCTAACTAC	AACCAGAAGT	TCAAGGGCCG
201				
<u>GGT</u> CACCA <u>ATC</u>	ACCGTAG <u>ACA</u>	<u>CCT</u> CTGCCAG	CACCGCCTAC	ATGGAACTGT
251				
CCAGCCTGCG	CTCCGAGGAC	ACTGCAGTCT	ACTACTGCGC	CAGAAGAAGA
301				
ATCGCCTATG	GTTACGACGA	GGGCCATGCT	ATGGACTACT	GGGGTCAAGG
351		372		
AACCCTTGTC	ACCGTCTCCT	CA		

Fig. 12

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SCKTSRYTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKFKGR <u>VTI</u>	TV <u>DT</u> SASTAY	MELSSLRSED	TAVYYCARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGTLV	TVSS		

Fig. 13

1
 CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
 51
 CGTGAAAGTC AGCTGTAAAA CTAGTAGATA CACCTTCACT GAATACACCA
 101
 TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
 151
 ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
 201
 GGTCACCAATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC ATGGAAGTGT
 251
 CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTTCTGCGC CAGAAGAAGA
 301
 ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
 351 372
 AACCCTTGTC ACCGTCTCCT CA

Fig. 14

1	11	21	31	41
QVQLVQSGAE	VKKPGASVKV	SKTSRYTFT	EYTIHWVRQA	PGQRLEWIGG
51	61	71	81	91
INPNNGIPNY	NQKFGR <u>VTI</u>	TV <u>DT</u> SASTAY	MELSSLRSED	TAVY <u>F</u> CARRR
101	111	121-124		
IAYGYDEGHA	MDYWGQGLV	TVSS		

Fig. 15

1
 CAGGTGCAAC TAGTGCAGTC CGGCGCCGAA GTGAAGAAAC CCGGTGCTTC
 51
 CGTGAAAGTC AGCTGTAAAA CTAGTGGATA CACCTTCACT GAATACACCA
 101
 TACACTGGGT TAGACAGGCC CCTGGCCAAA GGCTGGAGTG GATAGGAGGT
 151
 ATTAATCCTA ACAATGGTAT TCCTAACTAC AACCAGAAGT TCAAGGGCCG
 201
 GGTCACCAATC ACCGTAGACA CCTCTGCCAG CACCGCCTAC ATGGAAGTGT
 251
 CCAGCCTGCG CTCCGAGGAC ACTGCAGTCT ACTACTGCGC CAGAAGAAGA
 301
 ATCGCCTATG GTTACGACGA GGGCCATGCT ATGGACTACT GGGGTCAAGG
 351 372
 AACCCTTGTC ACCGTCTCCT CA

Fig. 16

```

1      11      21      31      41
QVQLVQSGAE VKKPGASVKV SCKTSGYTFT EYTIHWVRQA PGQRLEWIGG
51     61     71     81     91
INPNNGIPNY NQKFKGRVTI TVDTSASTAY MELSSLRSED TAVYYCARRR
101    111    121-124
IAYGYDEGHA MDYWGQGTLV TVSS

```

Fig. 17

```

1
DIVMSQSPSS LAVSVGEKVT MSCKSSQSLL YSRNQKNYLA WFQQKPGQSP
51
KLLIFWASTR ESGVPDRFTG SGFGTDFNLT ISSVQAEDLA VYDCQQYFSY
101
PLTFGAGTKL ELKRTVAAPS VFIFPPSDEQ LKSGTASVVC LLNNFYPREA
151
KVQWKVDNAL QSGNSQESVT EQDSKDSTYS LSSTLTLSKA DYEKHKVYAC
201
EVTHQGLSSP VTKSFNRGEC

```

Fig. 18

```

1
VQLQQSGPEL VKPGASVKMS CKTSRYTFTE YTIHWVRQSH GKSLEWIGGI
51
NPNNGIPNYN QKFKGRATLT VGKSSSTAYM ELRSLTSEDS AVYFCARRRI
101
AYGYDEGHAM DYWGQGTSVT VSSASTKGPS VFPLAPSSKS TSGGTAALGC
151
LVKDYFPEPV TVSWNSGALT SGVHTFPAVL QSSGLYSLSS VVTVPSSSLG
201
TQTYICNVNH KPSNTKVDKK VEPKSCDKTH TCPPCPAPEL LGGPSVFLFP
251
PKPKDTLMIS RTPEVTCVVV DVSHEDPEVK FNWYVDGVEV HNAKTKPREE
301
QYNSTYRVVS VLTVLHQDWL NGKEYKCKVS NKALPAPIEK TISKAKGQPR
351
EPQVYTLPPS REEMTKNQVS LTCLVKGFYP SDIAVEWESN GQPENNYKTT
401
PPVLDSGSGF FLYSKLTVDK SRWQQGNVFS CSVMHEALHN HYTQKSLSLS
451
PGK

```

Fig. 19

340	350	360	370	380
CGTACTGTGG	CTGCACCATC	TGTCTTCATC	TTCCCGCCAT	CTGATGAGCA
390	400	410	420	430
GTTGAAATCT	GGAAGTGCCT	CTGTTGTGTG	CCTGCTGAAT	AACTTCTATC
440	450	460	470	480
CCAGAGAGGC	CAAAGTACAG	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT
490	500	510	520	530
AACTCCCAGG	AGAGTGTCAC	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG
540	550	560	570	580
CCTCAGCAGC	ACCTTGACGC	TGAGCAAAGC	AGACTACGAG	AAACACAAAG
590	600	610	620	630
TCTACGCCTG	CGAAGTCACC	CATCAGGGCC	TGAGCTCGCC	CGTCACAAAG
640	650	660		
AGCTTCAACA	GGGGAGAGTGT			

Fig. 20

114	124	134	144	154
RTVAAPSVFI	FPPSDEQLKS	GTASVVCLLN	NFYPREAKVQ	WKVDNALQSG
164	174	184	194	204
NSQESVTEQD	SKDSTYSLSS	TLTLSKADYE	KHKVYACEVT	HQGLSSPVTK
214-220				
SFNRGEC				

Fig. 21

373
GCCTCCACCA AGGGCCCATC GGTCTTCCCC CTGGCACCCCT CCTCCAAGAG
423
CACCTCTGGG GGCACAGCGG CCCTGGGCTG CCTGGTCAAG GACTACTTCC
473
CCGAACCGGT GACGGTGTCTG TGGAACCTCAG GCGCCCTGAC CAGCGGCGTG
523
CACACCTTCC CGGCTGTCCT ACAGTCCTCA GGACTCTACT CCCTCAGCAG
573
CGTGGTGACC GTGCCCTCCA GCAGCTTGGG CACCCAGACC TACATCTGCA
623
ACGTGAATCA CAAGCCCAGC AACACCAAGG TGGACAAGAA AGTTGAGCCC
673
AAATCTTGTG ACAAACCTCA CACATGCCCA CCGTGCCCAG CACCTGAACT
723
CCTGGGGGGA CCGTCAGTCT TCCTCTTCCC CCCAAAACCC AAGGACACCC
773
TCATGATCTC CCGGACCCCT GAGGTCACAT GCGTGGTGGT GGACGTGAGC
823
CACGAAGACC CTGAGGTCAA GTTCAACTGG TACGTGGACG GCGTGGAGGT
873
GCATAATGCC AAGACAAAGC CGCGGGAGGA GCAGTACAAC AGCACGTACC
923
GGGTGGTCAG CGTCCTCACC GTCCTGCACC AGGACTGGCT GAATGGCAAG
973
GAGTACAAGT GCAAGGTCTC CAACAAAGCC CTCCCAGCCC CCATCGAGAA
1023
AACCATCTCC AAAGCCAAAG GGCAGCCCCG AGAACCACAG GTGTACACCC
1073
TGCCCCCATC CCGGGAGGAG ATGACCAAGA ACCAGGTCAG CCTGACCTGC
1123
CTGGTCAAAG GCTTCTATCC CAGCGACATC GCCGTGGAGT GGGAGAGCAA
1173
TGGGCAGCCG GAGAACAACCT ACAAGACCAC GCCTCCCGTG CTGGACTCCG
1223
ACGGCTCCTT CTTCTCTAC AGCAAGCTCA CCGTGGACAA GAGCAGGTGG
1273
CAGCAGGGGA ACGTCTTCTC ATGCTCCGTG ATGCATGAGG CTCTGCACAA
1323
CCACTACACG CAGAAGAGCC TCTCCCTGTC TCCGGGTAAA
1362

Fig. 22

125
ASTKGPSVFP LAPSSKSTSG GTAALGCLVK DYFPEPVTVS WNSGALTSGV
175
HTFPAVLQSS GLYSLSSVVT VPSSSLGTQT YICNVN HKPS NTKVDKKVEP
225
KSCDKTHTCP PCPAPELLGG PSVFLFPPKP KDTLMIS RTP EVTCVVVDVS
275
HEDPEVKFNW YVDGVEVHNA KTKPREEQYN STYRVVSVLT VLHQDWLNGK
325
EYKCKVSNKA LPAPIEKTIS KAKGQPREPQ VYTLPPSREE MTKNQVSLTC
375
LVKGFYPSDI AVEWESNGQP ENNYKTTPPV LDSDGSFFLY SKLTVDKSRW
425 454
QQGNVFSCSV MHEALHNHYT QKSLSLSPGK

10/43

Fig. 23A

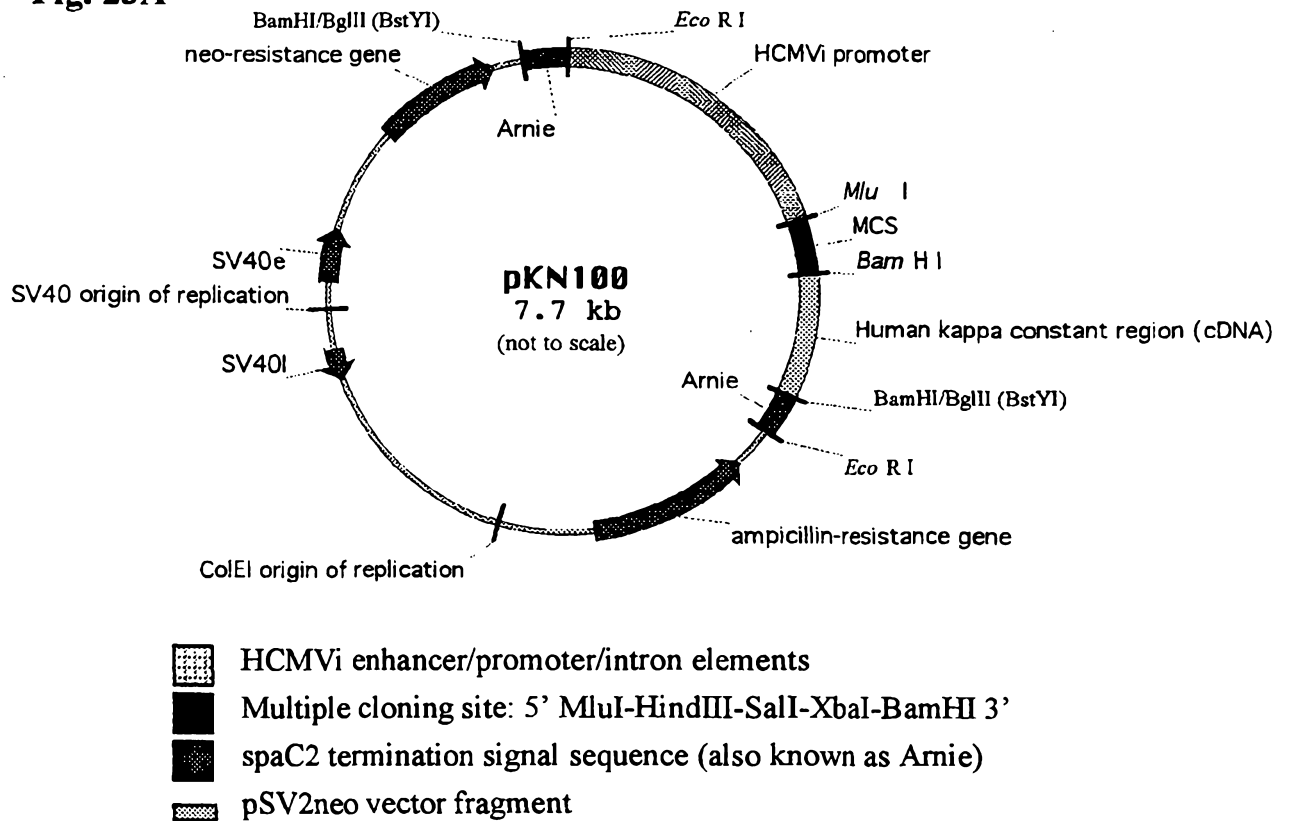


Fig. 23B

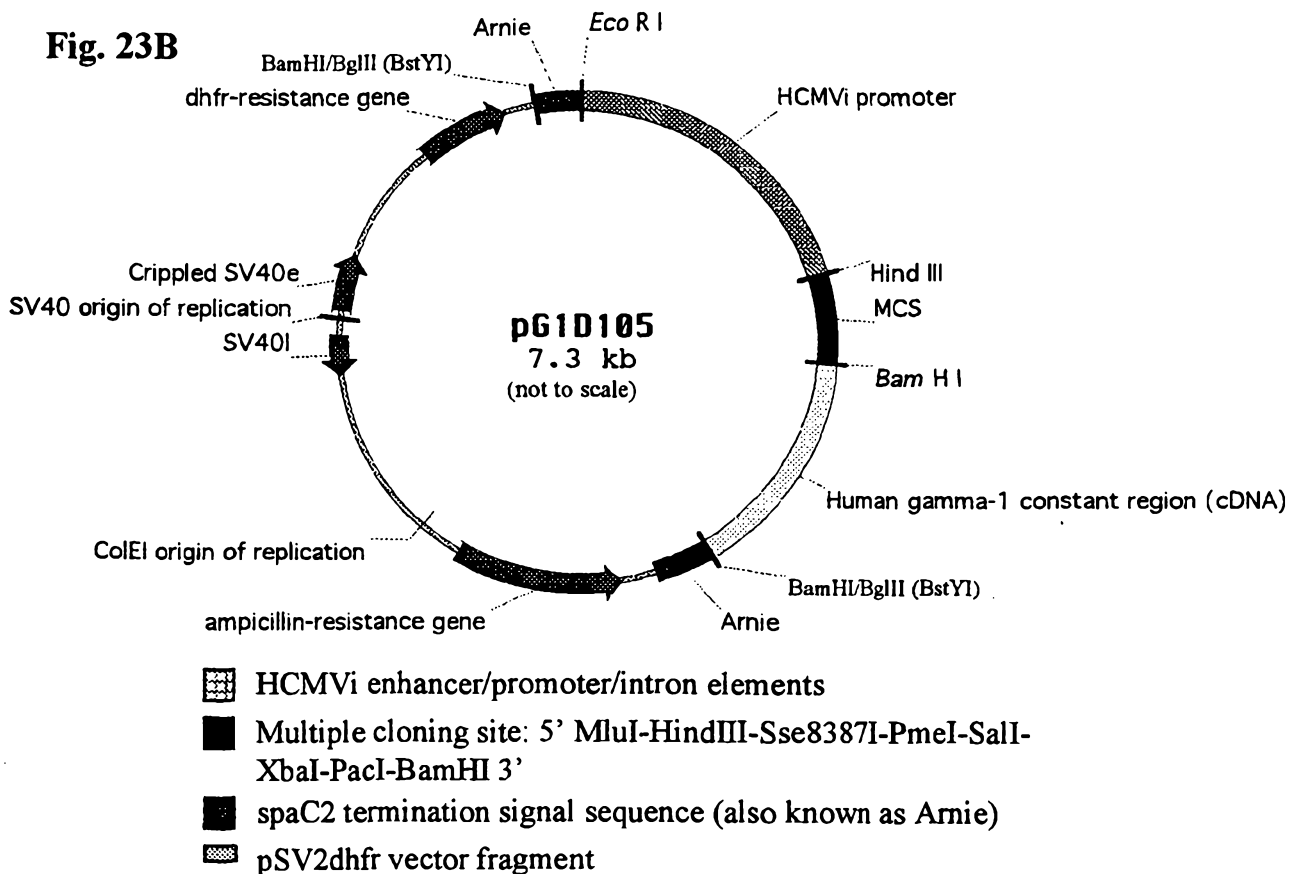


Fig. 24

HindIII
aagcttGCGGCCACCatggattcacaggcccaggttcttatgttactgccgctatgggta
1 -----+-----+-----+-----+-----+-----+
ttcgaacGGCGGTGGTaccctaagtgtccgggtccaagaatacaatgacggcgatacccat
Kozak sequence
M D S Q A Q V L M L L P L W V
tctggtagctgtggggacattgtgatgtcacagtcctccatcctccctagctgtgtcagtt
61 -----+-----+-----+-----+-----+-----+
agaccatggacacccctgtaacactacagtggtcagaggtaggagggatcgacacagtcaa
S G T C G D I V M S Q S P S S L A V S V
ggagagaagggttactatgagctgcaagtccagtcagagccttttatatagtcgtaatcaa
121 -----+-----+-----+-----+-----+-----+
cctctcttccaatgatactcgacgttcaggtcagtcctcgaaaatatatcagcattagtt
G E K V T M S C K S S Q S L L Y S R N Q
CDR 1
aagaactacttggcctgggtccagcagaagccagggcagtcctcctaaactgctgattttc
181 -----+-----+-----+-----+-----+-----+
ttcttgatgaaccggaccaaggtcgtcttcgggtcccgtcagaggatttgacgactaaaag
K N Y L A W F Q Q K P G Q S P K L L I F
tgggcatccactaggggaatctgggggtccctgatcgcttcacaggcagtggtttgggacg
241 -----+-----+-----+-----+-----+-----+
accgtaggtgatcccttagaccccagggaactagcgaagtgtccgtcacctaaacctgc
W A S T R E S G V P D R F T G S G F G T
CDR 2
gatttcaatctcaccatcagcagtggtgcaggctgaggacctggcagtttatgactgtcag
301 -----+-----+-----+-----+-----+-----+
ctaaagttagagtggtagtcgtcacacgtccgactcctggaccgtcaaatactgacagtc
D F N L T I S S V Q A E D L A V Y D C Q
caatatatttagctatccgctcacgttcgggtgctgggaccaagctggagctgaAACGTGAG
361 -----+-----+-----+-----+-----+-----+
gttataaaatcgataggcgagtgcaagccacgacctgggttcgacctcgactTTGCACTG
splice donor site
Q Y F S Y P L T F G A G T K L E L K
CDR 3
BamHI
Tggatcc
421 ----- 427
Acctagg

12/43

Fig. 25

HindIII
 AAGCTTGCCGCCACCATGGGATGGAGCTGGGTCTTTCTCTTTCTCCTGTCAGGAACTGCA
 1 -----+-----+-----+-----+-----+-----+-----+
 TTCGAACGGCGGTGGTACCCCTACCTCGACCCAGAAAGAGAAAGAGGACAGTCCTTGACGT
 Kozak sequence
 M G W S W V F L F L L S G T A

 GGTGTCCTCTCTGAGGTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCT
 61 -----+-----+-----+-----+-----+-----+-----+
 CCACAGGAGAGACTCCAGGTCGACGTTGTCAGACCTGGACTCGACCACTTCGGACCCCGA

 G V L S E V Q L Q Q S G P E L V K P G A

 TCAGTAAAGATGTCCTGCAAGACTTCTAGATACACATTCACTGAATACACCATACACTGG
 121 -----+-----+-----+-----+-----+-----+-----+
 AGTCATTTCTACAGGACGTTCTGAAGATCTATGTGTAAGTGACTTATGTGGTATGTGACC

 S V K M S C K T S R Y T F T E Y T I H W
 CDR 1

 GTGAGACAGAGCCATGGAAAGAGCCTTGAGTGGATTGGAGGTATTAATCCTAACAATGGT
 181 -----+-----+-----+-----+-----+-----+-----+
 CACTCTGTCTCGGTACCTTTCTCGGAACTCACCTAACCTCCATAATTAGGATTGTTACCA

 V R Q S H G K S L E W I , G G I N P N N G
 CDR 2

 ATTCCTAACTACAACCAGAAGTTCAAGGGCAGGGCCACATTGACTGTAGGCAAGTCCTCC
 241 -----+-----+-----+-----+-----+-----+-----+
 TAAGGATTGATGTTAGTCTTCAAGTTCCCGTCCCGGTGTAAGTACATCCGTTCCAGGAGG

 I P N Y N Q K F K G R A T L T V G K S S

 AGCACCGCCTACATGGAGCTCCGCAGCCTGACATCTGAGGATTCTGCGGTCTATTTCTGT
 301 -----+-----+-----+-----+-----+-----+-----+
 TCGTGGCGGATGTACCTCGAGGCGTCGGACTGTAGACTCCTAAGACGCCAGATAAAGACA

 S T A Y M E L R S L T S E D S A V Y F C

 GCAAGAAGAAGAATCGCCTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAA
 361 -----+-----+-----+-----+-----+-----+-----+
 CGTTCTTCTTCTTAGCGGATACCAATGCTGCTCCCGGTACGATACCTGATGACCCCACTT

 A R R R I A Y G Y D E G H A M D Y W G Q
 CDR 3

 BamHI
 GGAACCTCAGTCACCGTCTCCTCAGGTGAGTGGATCC
 421 -----+-----+-----+-----+-----+-----+-----+ 468
 CCTTGGAGTCAGTGGCAGAGGAGTCCACTCACCTAGG
 splice donor site
 G T S V T V S S

Fig. 26 /1

```

                                Spe I
1  gaattccagc acactggcgg ccggttACTAG TTATTAATAG TAATCAATTA
51  CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT
101  ACGGTAAATG GCCCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC
151  GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT
201  GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCTT GGCAGTACAT
251  CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA
301  ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA
                                SnaB I
351  CTTGGCAGTA CATCTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG
401  TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGA CT CACGGGGATT
451  TCCAAGTCTC CACCCCAT TG ACGTCAATGG GAGTTTGT TTT TGGCACC AAA
501  ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA
551  ATGGGCGGTA GCGGTGTACG GTGGGAGGTC TATATAAGCA GAGCTCGTTT
601  AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC
                                Sac II
651  ATAGAAGACA CCGGGACCGA TCCAGCCTCC GCGGCCGGA ACGGTGCATT
701  GGAACGCGGA TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGT
751  CTATAGGCC ACCCCCTTGG CTTCTTATGC ATGCTATACT GTTTTTGGCT
801  TGGGGTCTAT ACACCCCGC TTCCTCATGT TATAGGTGAT GGTATAGCTT
851  AGCCTATAGG TGTGGGTAT TGACCATTAT TGACCACTCC CCTATTGGTG
901  ACGATACTTT CCATTACTAA TCCATAACAT GGCTCTTTGC CACA ACTCTC
951  TTTATTGGCT ATATGCCAAT AACTGTCCT TCAGAGACTG ACACGGACTC
1001  TGTATTTT TA CAGGATGGGG TCTCATTTAT TATTTACAAA TTCACATATA
1051  CAACACCACC GTCCCCAGTG CCCGCAGTTT TTATTAAACA TAACGTGGGA
                                BspE I
1101  TCTCCACGCG AATCTCGGGT ACGTGTTTCCG GACATGGGCT CTTCTCCGGT
1151  AGCGGCGGAG CTTCTACATC CGAGCCCTGC TCCCATGCCT CCAGCGACTC
1201  ATGGTCGCTC GGCAGCTCCT TGCTCCTAAC AGTGGAGGCC AGACTTAGGC
1251  ACAGCACGAT GCCCACCACC ACCAGTGTGC CGCACAAGGC CGTGGCGGTA

```

1301	GGGTATGTGT	CTGAAAATGA	GCTCgggggag	cgggcttgca	ccgctgacgc
		Afl II			
1351	at tt t g g a a g a	<u>cttaagg</u> gcag	cggcagaaga	agatgcaggc	agctgagttg
1401	ttgtgttctg	ataagagtca	gaggtaactc	ccgttgcggt	gctgttaacg
1451	gtggaggggca	gtgtagtctg	agcagtactc	gttgctgccg	cgcgcgccac
1501	cagacataat	agctgacaga	ctaacagact	gttcctttcc	atgggtcttt
			Mlu I	Hind III	
1551	tctgcagtca	ccgtccttga	<u>cacgcgt</u> ctc	ggg <u>aagctt</u> G	CCGCCACCAT
					M
					Kpn I
1601	GGATTCACAG	GCCCAGGTTC	TTATGTTACT	GCCGCTATGG	GTATCT <u>GGTA</u>
	D S Q	A Q V	L M L L	P L W	V S G
1651	<u>CCTGTGGGGA</u>	CATTGTGATG	TCACAGTCTC	CATCCTCCCT	AGCTGTGTCA
	T C G D	I V M	S Q S	P S S L	A V S
1701	GTTGGAGAGA	AGGTTACTAT	GAGCTGCAAG	TCCAGTCAGA	GCCTTTTATA
	V G E	K V T M	S C K	S S Q	S L L Y
	<u>XbaI</u>			CDR 1	
1751	<u>TTCTAGAAAT</u>	CAAAAGAACT	ACTTGGCCTG	GTTCCAGCAG	AAGCCAGGGC
	<u>S R N</u>	<u>Q K N</u>	<u>Y L A W</u>	F Q Q	K P G
1801	AGTCTCCTAA	ACTGCTGATT	TTCTGGGCAT	CCACTAGGGA	ATCTGGGGTC
	Q S P K	L L I	F <u>W A</u>	<u>S T R E</u>	<u>S G V</u>
				CDR 2	
1851	CCTGATCGCT	TCACAGGCAG	TGGATTTGGG	ACGGATTTCA	ATCTCACCAT
	P D R	F T G S	G F G	T D F	N L T I
1901	CAGCAGTGTG	CAGGCTGAGG	ACCTGGCAGT	TTATGACTGT	CAGCAATATT
	S S V	Q A E	D L A V	Y D C	<u>Q Q Y</u>
1951	TTAGCTATCC	GCTCACGTTC	GGTGCTGGGA	CCAAGCTGGA	GCTGAA <u>ACGT</u>
	<u>F S Y P</u>	<u>L T</u> F	G A G	T K L E	L K <u>R</u>
	CDR 3				
	BamH I				
2001	<u>GAGTggatcc</u>	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	TCCCTAACAT
2051	GCCCTGTGAT	TATGCGCAAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT
2101	TAAACACCAT	CCTGTTTGCT	TCTTTCCT <u>CA</u>	<u>GGA</u> ACTGTGG	CTGCACCATC
				T V	A A P S
2151	TGTCTTCATC	TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT
	V F I	F P P	S D E Q	L K S	G T A
2201	CTGTTGTGTG	CCTGCTGAAT	AACTTCTATC	CCAGAGAGGC	CAAAGTACAG
	S V V C	L L N	N F Y	P R E A	K V Q
2251	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	AACTCCCAGG	AGAGTGTAC
	W K V	D N A L	Q S G	N S Q	E S V T
2301	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	ACCCTGACGC
	E Q D	S K D	S T Y S	L S S	T L T

Fig. 26 /3

```

2351  TGAGCAAAGC AGACTACGAG AAACACAAAG TCTACGCCTG CGAAGTCACC
      L S K A   D Y E   K H K   V Y A C   E V T
2401  CATCAGGGCC TGAGCTCGCC CGTCACAAAG AGCTTCAACA GGGGAGAGTG
      H Q G   L S S P   V T K   S F N   R G E C
2451  TTAGAGGGAG AAGTGCCCCC ACCTGCTCCT CAGTTCCAGC CTGACCCCCT
      *
2501  CCCATCCTTT GGCCTCTGAC CCTTTTTCCT CAGGGGACCT ACCCCTATTG
2551  CGGTCCTCCA GCTCATCTTT CACCTCACCC CCCTCCTCCT CCTTGGCTTT
2601  AATTATGCTA ATGTTGGAGG AGAATGAATA AATAAAGTGA ATCTTTGCAC
2651  CTGTGGTGGA TCTAATAAAA GATATTTATT TTCATTAGAT ATGTGTGTTG
2701  GTTTTTTGTG TGCAGTGCCT CTATCTGGAG GCCAGGTAGG GCTGGCCTTG
2751  GGGGAGGGGG AGGCCAGAAT GACTCCAAGA GCTACAGGAA GGCAGGTCAG
2801  AGACCCCACT GGACAAACAG TGGCTGGACT CTGCACCATA ACACACAATC
2851  AACAGGGGAG TGAGCTGGAA ATTTGCTAGC GAATTCTTGA AGACGAAAGG
2901  GCCTCGTGAT ACGCCTATTT TTATAGGTTA ATGTCATGAT AATAATGGTT
2951  TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT
3001  TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
3051  AACCCTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
3101  CAACATTTCC GTGTCGCCCT TATTCCTTTT TTTGCGGCAT TTTGCCTTCC
3151  TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
3201  AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG
3251  ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
3301  TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
3351  AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
3401  TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
3451  ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
      Pvu I
3501  TGACAACCGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG
3551  GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC
3601  CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA ATGGCAACAA

```

Fig. 26 /4

3651 CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA
3701 CAATTAATAG ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG
3751 CTCGGCCCTT CCGGCTGGCT GGTATTATTGC TGATAAATCT GGAGCCGGTG
3801 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
3851 TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
3901 ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTGGT
3951 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAACTT
4001 CATTTTTAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT
4051 GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
4101 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
4151 TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
4201 GGATCAAGAG CTACCAACTC TTTTCCGAA GGTAAGTGGC TTCAGCAGAG
4251 CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC
4301 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT
4351 ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT
4401 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
4451 TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
4501 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG
4551 CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
4601 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG
4651 CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
4701 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT
4751 TGCTGGCCTT TTGCTCACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
4801 GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATAACCGT CGCCGCAGCC
4851 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
4901 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATG

Fig. 26 /5

Bst1107I

4951 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGTATA

5001 CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC CGACACCCGC

5051 CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT

5101 TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTTC

5151 ACCGTCATCA CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG

5201 GGTGTGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG

5251 CATCTCAATT AGTCAGCAAC CAGGCTCCCC AGCAGGCAGA AGTATGCAAA

5301 GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC

5351 ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG

Sfi I

5401 ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC

Stu I/Avr II

5451 TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG CTTTTGCAAA

5501 AAGCTAGCTT CACGCTGCCG CAAGCACTCA GGGCGCAAGG GCTGCTAAAG

5551 GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA GAAACGGTGC TGACCCCGGA

5601 TGAATGTCAG CTA CTGCTGGGCT ATCTGGACAA GGGAAAACGC AAGCGCAAAG

5651 AGAAAGCAGG TAGCTTGCAG TGGGCTTACA TGGCGATAGC TAGACTGGGC

5701 GGTTTTATGG ACAGCAAGCG AACCGBAATT GCCAGCTGGG GCGCCCTCTG

5751 GTAAGGTTGG GAAGCCCTGC AAAGTAACT GGATGGCTTT CTTGCCGCCA

Bgl II/Bcl I

5801 AGGATCTGAT GGCGCAGGGG ATCAAGATCT GATCAAGAGA CAGGATGAGG

5851 ATCGTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG

5901 CTTGGGTGGA GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC

5951 TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT

6001 TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAGGACGAGG

Msc I

6051 CAGCGCGGCT ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG

6101 CTCGACGTTG TCACTGAAGC GGAAGGGAC TGGCTGCTAT TGGGCGAAGT

6151 GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT

6201 CCATCATGGC TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC

Fig. 26 /6

6251 TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG
 6301 GATGGAAGCC GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG
 6351 GGCTCGCGCC AGCCGAACTG TTCGCCAGGC TCAAGGCGCG CATGCCCCGAC
 6401 GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT
 6451 GGTGGA^{Rsr II}AAAT GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG
 6501 TGGCGGACCG CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA
 6551 GAGCTTGGCG GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC
 6601 CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT
 6651 TCTGAGCGGG ACTCTGGGGT ^{Nsp V}TCGAAATGAC CGACCAAGCG ACGCCCAACC
 6701 TGCCATCACG AGATTTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC
 6751 TTCGGAATCG TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA
 6801 TCTCATGCTG GAGTTCTTCG CCCAC^{Sma I}CCCGG ^{Nru I}GCTCGATCCC CTCGCGAGTT
 6851 GGTTCAGCTG CTGCCTGAGG CTGGACGACC TCGCGGAGTT CTACCGGCAG
 6901 TGCAAATCCG TCGGCATCCA GGAAACCAGC AGCGGCTATC CGCGCATCCA
 6951 TGCCCCCGAA CTGCAGGAGT GGGGAGGCAC GATGGCCGCT TTGGTCCCGG
 7001 ATCTTTGTGA AGGAACCTTA CTTCTGTGGT GTGACATAAT TGGACAAACT
 7051 ACCTACAGAG ATTTAAAGCT CTAAGGTAAA TATAAAATTT TTAAGTGTAT
 7101 AATGTGTTAA ACTACTGATT CTAATTGTTT GTGTATTTTA GATTCCAACC
 7151 TATGGA^{Sma I}ACTG ATGAATGGGA GCAGTGGTGG AATGCCTTTA ATGAGGAAAA
 7201 CCTGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG GCTACTGCTG
 7251 ACTCTCAACA TTCTACTCCT CCAAAAAAGA AGAGAAAGGT AGAAGACCCC
 7301 AAGGACTTTC CTTCAGAATT GCTAAGTTTT TTGAGTCATG CTGTGTTTAG
 7351 TAATAGA^{Nru I}ACT CTTGCTTGCT TTGCTATTTA CACCACAAAG GAAAAAGCTG
 7401 CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAC CTTTATAAGT
 7451 AGGCATAACA GTTATAATCA TAACATACTG TTTTTTCTTA CTCCACACAG
 7501 GCATAGAGTG TCTGCTATTA ATA^{Sma I}ACTATGC TCAAAAATTG TGTACCTTTA

Fig. 26 /7

7551 GCTTTTAAAT TTGTAAAGGG GTTAATAAGG AATATTTGAT GTATAGTGCC
7601 TTGACTAGAG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
7651 CTTTAAAAAA CCTCCCACAC CTCCCCCTGA ACCTGAAACA TAAAATGAAT
Mun I
7701 GCAATTGTTG TTGTTAAGTT GTTTATTGCA GCTTATAATG GTTACAAATA
7751 AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCCTGCATT
7801 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC
7851 TAATAAAAGA TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG
7901 CAGTGCCTCT ATCTGGAGGC CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG
7951 GCCAGAATGA CTCCAAGAGC TACAGGAAGG CAGGTCAGAG ACCCCACTGG
8001 ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA CAGGGGAGTG
8051 AGCTGGAAAT TTGCTAGC

Fig. 27/1

1 TTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAAT
61 GGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTT
121 ATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCT
181 TCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTGCGCCCTTATTCC
241 CTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA
301 AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGG
361 TAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTCCAATGATGAGCACTTTTAAAGT
421 TCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTGCGCG
481 CATACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
541 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGC
Pvu I
601 GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGACAA
661 CATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATACC
Fsp I
721 AAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTTGCGCAAACTATT
781 AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGA

Fig. 27 /2

841 TAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA
901 ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
961 GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
1021 TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGT
1081 TTAATCATATATACTTTAGATTGATTTAAACTTCATTTTTTAATTTAAAGGATCTAGGT
1141 GAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTTCGTTCCACTG
1201 AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
1261 AATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTGTTGCCGGATCA
1321 AGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATAC
1381 TGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
1441 ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
1501 TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTGGGGCTGAACGGG
1561 GGGTTCGTGCACACAGCCCAGCTTGAGCGAACGACCTACACCGAACTGAGATACCTACA
1621 GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
1681 AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
1741 TCTTTATAGTCCTGTGCGGTTTTCGCCACCTCTGACTTGAGCGTCGATTTTTTGTGATGCTC
1801 GTCAGGGGGGCGGAGCCTATGGAAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
1861 CTTTTGCTGGCCTTTTGCTCCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAA
1921 CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG
1981 CGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTCTCCTTACGCATCT
2041 GTGCGGTATTTACACCGCATATGGTGCACTCTCAGTACAATCTGCTCTGATGCCGCATA
2101 GTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTATGGCTGCGCCCCGACAC
2161 CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
2221 CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGTCAGAGGTTTTACCGTTCATACCGAAA
2281 CGCGCGAGGCAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCCTAACTCCGCC
2341 CATCCCGCCCCTAACTCCGCCCAGTTCCGCCCATTCTCCGCCCCATGGCTGACTAATTTT
2401 TTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGG
2461 AGGCTTTTTTGAGGCCTAGGCTTTTGCAAAAAGCTAGCTTACAGCTCAGGGCTGCGATT

Fig. 27 /3

2521 TCGCGCCAACTTGACGGCAATCCTAGCGTGAAGGCTGGTAGGATTTTATCCCCGCTGCC
2581 ATCATGGTTCGACCATTGAACTGCATCGTCGCCGTGTCCCAAAATATGGGGATTGGCAAG
2641 AACGGAGACCTACCCTGGCCTCCGCTCAGGAACGAGTTCAAGTACTTCCAAAGAATGACC
2701 ACAACCTCTTCAGTGAAGGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTC
2761 TCCATTCCCTGAGAAGAATCGACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAA
2821 CTCAAAGAACCACCACGAGGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGA
2881 CTTATTGAACAACCGGAATTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGT
2941 TCTGTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTGTGACAAGGATC
3001 ATGCAGGAATTTGAAAGTGACACGTTTTTCCCAGAAATTGATTTGGGGAAATATAAACTT
3061 CTCCCAGAATACCCAGGCGTCCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAG
3121 TTTGAAGTCTACGAGAAGAAAGACTAACAGGAAGATGCTTTCAGTTCTCTGCTCCCCTC
3181 CTAAAGCTATGCATTTTTTATAAGACCATGGGACTTTTGCTGGCTTTAGATCTTTGTGAAG
Bgl II
3241 GAACCTTACTTCTGTGGTGTGACATAATTGGACAAACTACCTACAGAGATTTAAAGCTCT
3301 AAGGTAAATATAAAATTTTTAAGTGTATAATGTGTAAACTACTGATTCTAATTGTTTGT
3361 GTATTTTAGATTCCAACCTATGGAAGTATGAATGGGAGCAGTGGTGAATGCCTTTAAT
3421 GAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGAC
3481 TCTCAACATTCTACTCCTCCAAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCT
3541 TCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTT
3601 GCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATAT
3661 TCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTCTTACT
3721 CCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGC
3781 TTTTAAATTTGTAAAGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGAT
BsaB I
3841 CATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCCACACCT
Mun I
3901 CCCCCTGAACCTGAAACATAAAATGAATGCAATTGTGTGTTAACTTGTTTATTGCAGC
3961 TTATAATGGTTACAAATAAAGCAATAGCATCACAATTTACAAATAAAGCATTTTTTTC
4021 ACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGGATCTA
4081 ATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCCTCTAT
4141 CTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGAGGCCAGAATGACTCCAAGAGCTA
4201 CAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCATAACAC

Fig. 27 /4

EcoR I
4261 ACAATCAACAGGGGAGTGAGCTGGAAATTTGCTAGCGAATTCcagcacactggcggccgt
Spe I
4321 tACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTT
4381 CCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCC
4441 ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACG
4501 TCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT
4561 GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA
SnaB I
4621 GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT
4681 TACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
4741 GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAATCA
4801 ACGGGACTTTCCAAATGTCGTAACAACCTCCGCCCCATTGACGCAAATGGGCGGTAGGCG
4861 TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAG
4921 ACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGG
4981 CCGGGAACGGTGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTA
5041 TAGAGTCTATAGGCCACCCCCTTGGCTTCTTATGCATGCTATACTGTTTTTGGCTTGGG
Bpu1102I
5101 GTCTATACACCCCCGCTTCCTCATGTTATAGGTGATGGTATAGCTTAGCCTATAGGTGTG
Xcm I
5161 GGTTATTGACCATTATTGCCACTCCCCTATTGGTGACGATACTTTCATTACTAATCCA
5221 TAACATGGCTCTTTGCCACAACCTCTCTTTATTGGCTATATGCCAATACTGTCTTTCAG
5281 AGACTGACACGGACTCTGTATTTTTACAGGATGGGGTCTCATTTATTATTTACAAATTCA
5341 CATATACAACACCACCGTCCCCAGTGCCCGCAGTTTTTTATTAAACATAACGTGGGATCTC
BspE I
5401 CACGCGAATCTCGGGTACGTGTTCCGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTC
5461 TACATCCGAGCCCTGCTCCCATGCCTCCAGCGACTCATGGTCGCTCGGCAGCTCCTTGCT
5521 CCTAACAGTGGAGGCCAGACTTAGGCACAGCACGATGCCACCACCACAGTGTGCCGCA
.....
5581 CAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCggggagcgggcttgcaccgc
(Pvu II)
5641 tgacgcatttgaagacttaaggcagcggcagaagaagatgcaggcagctgagttgtgtgt
5701 gttctgataagagtcagaggtaactcccgttgcggtgctgttaacggtggagggcagtg
5761 agtctgagcagtgactcggtgctgcccgcgcgcgccaccagacataatagctgacagactaa
Mlu I
5821 cagactgttcctttccatgggtcttttctgcagtcaccgtccttgacACGCGTCTCGGGA
Hind III
5881 AGCTTGCCGCCACCATGGGATGGAGCTGGGTCTTTCTTTCTCTCTGTCAGGAAGTGCAG
M G W S W V F L F L L S G T A

Fig. 27 /5

(Pvu II)

5941 GTGTCCTCTCTGAGGTCCAGCTGCAACAGTCTGGACCTGAGCTGGTGAAGCCTGGGGCTT
 G V L S E V Q L Q Q S G P E L V K P G A

Xba I Dra III

6001 CAGTAAAGATGTCCTGCAAGACTTCTAGATACACATTCAGTGAATACACCATACTGGG
 S V . K M S C K T S R Y T F T E Y T I H W

CDR 1

6061 TGAGACAGAGCCATGGAAAGAGCCTTGAGTGGATTGGAGGTATTAATCCTAACAAATGGTA
 V R Q S H G K S L E W I G G I N P N N G

6121 TTCCTAACTACAACCAGAAGTTCAAGGGCAGGGCCACATTGACTGTAGGCAAGTCCTCCA
 I P N Y N Q K F K G R A T L T V G K S S

CDR 2

6181 GCACCGCTACATGGAGCTCCGCAGCCTGACATCTGAGGATTCTGCGGTCTATTTCTGTG
 S T A Y M E L R S L T S E D S A V Y F C

6241 CAAGAAGAAGAATCGCCTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAAG
 A R R R I A Y G Y D E G H A M D Y W G Q

CDR 3 BamH I

6301 GAACCTCAGTCACCGTCTCCTCAGGTGAGTGGATCCTCTGCGCCTGGGCCCAGCTCTGTC
 G T S V T V S S

6361 CCACACGCGGTACATGGCACCACCTCTCTTGACGCCTCCACCAAGGGCCCATCGGTCT
 S T K G P S V

6421 TCCCCCTGGCACCCCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGG
 F P L A P S S K S T S G G T A A L G C L

Age I

6481 TCAAGGACTACTTCCCCGAACCGGTGACGGTGTCGTGGAAGTCAAGCGCCCTGACCAGCG
 V K D Y F P E P V T V S W N S G A L T S

6541 GCGTGACACCTTCCCGGTGTCCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGG
 G V H T F P A V L Q S S G L Y S L S S V

BstE II

6601 TGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGC
 V T V P S S S L G T Q T Y I C N V N H K

6661 CCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGTGACAAAACCTCACACAT
 P S N T K V D K K V E P K S C D K T H T

6721 GCCCACCCTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTCCCCCAA
 C P P C P A P E L L G G P S V F L F P P

6781 AACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACG
 K P K D T L M I S R T P E V T C V V V D

6841 TGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA
 V S H E D P E V K F N W Y V D G V E V H

6901 ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCC
 N A K T K P R E E Q Y N S T Y R V V S V

6961 TCACCGTCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACA
 L T V L H Q D W L N G K E Y K C K V S N

Fig. 27 /6

7021 AAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
K A L P A P I E K T I S K A K G Q P R E

7081 CACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTCAGCCTGA
P Q V Y T L P P S R E E M T K N Q V S L

7141 CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
T C L V K G F Y P S D I A V E W E S N G

7201 AGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCC
Q P E N N Y K T T P P V L D S D G S F F

7261 TCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCT
L Y S K L T V D K S R W Q Q G N V F S C

7321 CCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
S V M H E A L H N H Y T Q K S L S L S P
NgoM I

7381 GTAAATGAGTGCGACGGCCGGCAAGCCCCGCTCCCCGGGCTCTCGCGGTGCGACGAGGAT
G K *

7441 GCTTGGCACGTACCCCCTGTACATACTTCCCGGGCGCCCAGCATGGAAATAAAGCACCGG

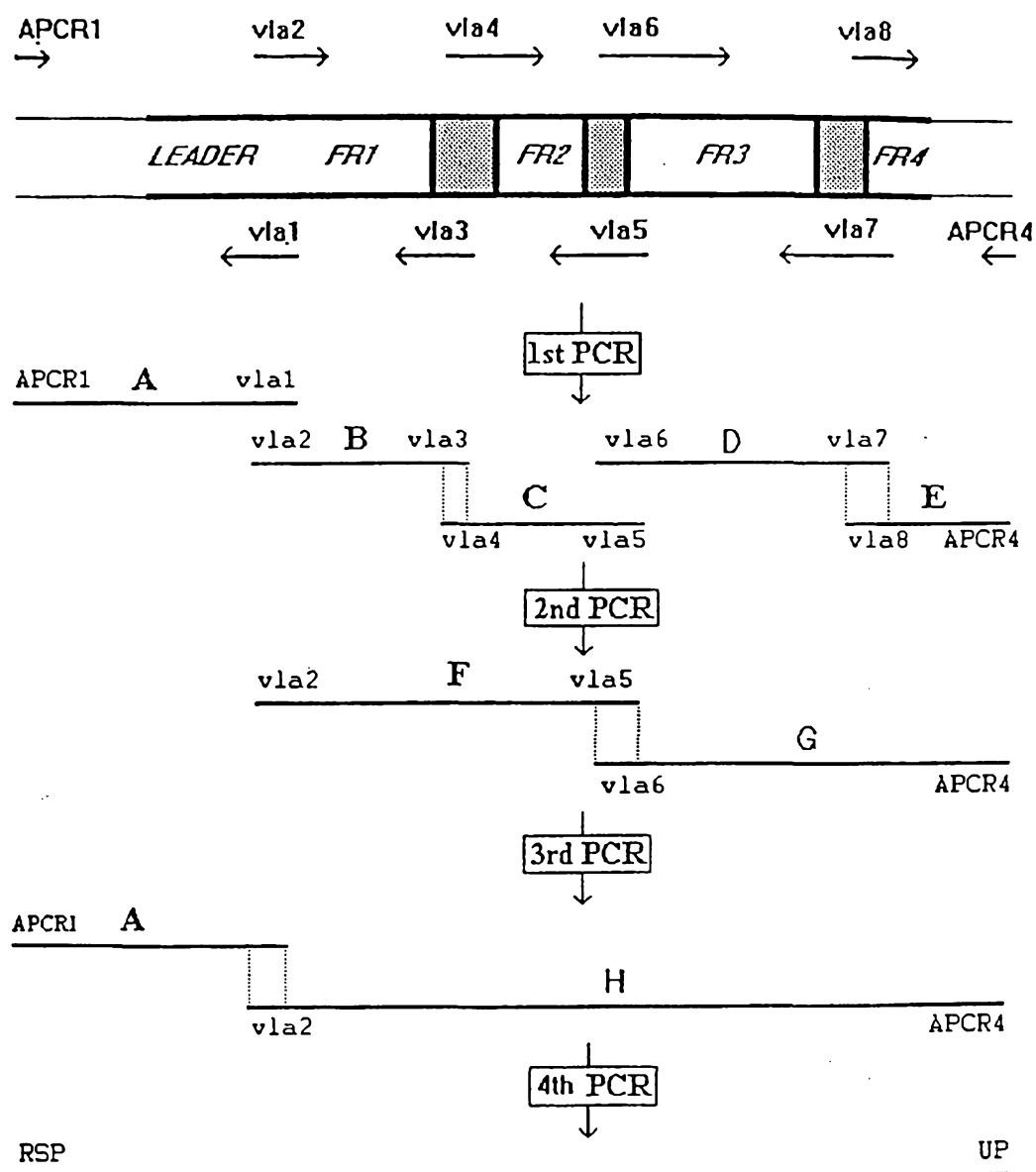
7501 ATCTAATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCC

7561 TCTATCTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGGAGGCCAGAATGACTCCAAG

7621 AGCTACAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCAT

7681 AACACACAATCAACAGGGGAGTGAGCTGGaaatttgctagcgaattaattc 7731

Fig. 28:



[illegible]

33																	51		
	L	A	W	Y	Q	Q	K	P	G	Q	P	P	K	L	L	I	F	W	A
A	TTG	GCC	TGG	TAT	CAG	CAG	AAA	CCA	GGA	CAG	CCA	CCC	AAA	CTC	CTC	ATC	TTT	TGG	GCT
	.	.	.	F	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
B	---	---	---	TC	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---
	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	Y	.	.
C	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	A-	---	---

[illegible][illegible][illegible]

27/43

Fig. 30 /1

Spe I

1 gaattccagc acactggcgg ccggttACTAG TTATTAATAG TAATCAATTA

51 CGGGGTCATT AGTTCATAGC CCATATATGG AGTTCCGCGT TACATAACTT

101 ACGGTAAATG GCGCGCCTGG CTGACCGCCC AACGACCCCC GCCCATTGAC

151 GTCAATAATG ACGTATGTTC CCATAGTAAC GCCAATAGGG ACTTTCCATT

201 GACGTCAATG GGTGGAGTAT TTACGGTAAA CTGCCCCTT GGCAGTACAT

251 CAAGTGTATC ATATGCCAAG TACGCCCCCT ATTGACGTCA ATGACGGTAA

301 ATGGCCCGCC TGGCATTATG CCCAGTACAT GACCTTATGG GACTTTCCTA

SnaB I

351 CTTGGCAGTA CATTACGTA TTAGTCATCG CTATTACCAT GGTGATGCGG

401 TTTTGGCAGT ACATCAATGG GCGTGGATAG CGGTTTGA CTACGGGGATT

451 TCCAAGTCTC CACCCCAT TG ACGTCAATGG GAGTTTGTTT TGGCACCAAA

501 ATCAACGGGA CTTTCCAAAA TGTCGTAACA ACTCCGCCCC ATTGACGCAA

551 ATGGGCGGTA GGCGTGACG GTGGGAGGTC TATATAAGCA GAGCTCGTTT

601 AGTGAACCGT CAGATCGCCT GGAGACGCCA TCCACGCTGT TTTGACCTCC

Sac II

651 ATAGAAGACA CCGGGACCGA TCCAGCCTCC GCGGCCGGGA ACGGTGCATT

701 GGAACGCGGA TTCCCCGTGC CAAGAGTGAC GTAAGTACCG CCTATAGAGT

751 CTATAGGCCC ACCCCCTTGG CTTCTTATGC ATGCTATACT GTTTTTGGCT

801 TGGGGTCTAT ACACCCCCGC TTCCTCATGT TATAGGTGAT GGTATAGCTT

851 AGCCTATAGG TGTGGGTTAT TGACCATTAT TGACCACTCC CCTATTGGTG

901 ACGATACTTT CCATTACTAA TCCATAACAT GGCTCTTTGC CACAACCTCTC

951 TTTATTGGCT ATATGCCAAT AACTGTCCT TCAGAGACTG ACACGGACTC

1001 TGTATTTTTA CAGGATGGGG TCTCATTTAT TATTTACAAA TTCACATATA

1051 CAACACCACC GTCCCCAGTG CCCGCAGTTT TTATTAAACA TAACGTGGGA

(BspE I)

1101 TCTCCACGCG AATCTCGGGT ACGTGTTCCG GACATGGGCT CTTCTCCGGT

1151 AGCGGCGGAG CTTCTACATC CGAGCCCTGC TCCCATGCCT CCAGCGACTC

1201 ATGGTCGCTC GGCAGCTCCT TGCTCCTAAC AGTGGAGGCC AGACTTAGGC

1251	ACAGCACGAT	GCCCACCACC	ACCAGTGTGC	CGCACAAGGC	CGTGGCGGTA
1301	GGGTATGTGT	CTGAAAATGA	GCTCggggag	cgggcttgca	ccgctgacgc
		Afl II			
1351	atttggaaga	<u>cttaaggcag</u>	cggcagaaga	agatgcaggc	agctgagttg
1401	ttgtgttctg	ataagagtca	gaggtaactc	ccgttgcggt	gctgttaacg
1451	gtggagggca	gtgtagtctg	agcagtactc	gttgctgccg	cgcgcgccac
1501	cagacataat	agctgacaga	ctaacagact	gttcctttcc	atgggtcttt
			Mlu I	Hind III	
1551	tctgcagtca	ccgtccttga	<u>cacgcgt</u> ctc	ggga <u>agctt</u> G	CCGCCACCAT
					M
1601	GGAGACAGAC	ACACTCCTGC	TATGGGTGCT	GCTGCTCTGG	GTTCCAGGTT
	E T D	T L L	L W V L	L L W	V P G
	(BspE I)				
1651	<u>CCTCCGGAGA</u>	CATTGTGATG	ACCCAATCTC	CAGACTCTTT	GGCTGTGTCT
	S S G D	I V M	T Q S	P D S L	A V S
1701	CTAGGGGAGA	GGGCCACCAT	CAACTGCAAG	TCCAGTCAGA	GCCTTTTATA
	L G E	R A T I	N C K	S S Q	S L L Y
	XbaI				CDR 1
1751	<u>TTCTAGAAAT</u>	CAAAAGAACT	ACTTGGCCTG	GTATCAGCAG	AAACCAGGAC
	<u>S R N</u>	<u>Q K N</u>	<u>Y L A W</u>	Y Q Q	K P G
					KpnI
1801	AGCCACCCAA	ACTCCTCATC	TTTTGGGCTA	GCACTAGGGA	ATCTGGGGTA
	Q P P K	L L I	F <u>W A</u>	S T R E	S G V
				CDR 2	
1851	<u>CCTGATAGGT</u>	TCAGTGGCAG	TGGGTTTGGG	ACAGACTTCA	CCCTCACCAT
	P D R	F S G S	G F G	T D F	T L T I
1901	TAGCAGCCTG	CAGGCTGAAG	ATGTGGCAGT	TTATTACTGT	CAGCAATATT
	S S L	Q A E	D V A V	Y Y C	<u>Q Q Y</u>
1951	TTAGCTATCC	GCTCACGTTT	GGACAAGGGA	CCAAGGTGGA	AATAAAACGT
	<u>F S Y P</u>	<u>L T F</u>	G Q G	T K V E	I K R
	CDR 3				
	BamH I				
2001	<u>GAGTggatcc</u>	ATCTGGGATA	AGCATGCTGT	TTTCTGTCTG	TCCCTAACAT
2051	GCCCTGTGAT	TATGCGCAA	CAACACACCC	AAGGGCAGAA	CTTTGTTACT
2101	TAAACACCAT	CCTGTTTGCT	TCTTTCCTCA	<u>GGA</u> ACTGTGG	CTGCACCATC
				T V	A A P S
2151	TGTCTTCATC	TTCCCGCCAT	CTGATGAGCA	GTTGAAATCT	GGAAGTGCCT
	V F I	F P P	S D E Q	L K S	G T A
2201	CTGTTGTGTG	CCTGCTGAAT	AACTTCTATC	CCAGAGAGGC	CAAAGTACAG
	S V V C	L L N	N F Y	P R E A	K V Q
2251	TGGAAGGTGG	ATAACGCCCT	CCAATCGGGT	AACTCCCAGG	AGAGTGTAC
	W K V	D N A L	Q S G	N S Q	E S V T
2301	AGAGCAGGAC	AGCAAGGACA	GCACCTACAG	CCTCAGCAGC	ACCCTGACGC
	E Q D	S K D	S T Y S	L S S	T L T

Fig. 30 /3

```

2351  TGAGCAAAGC AGACTACGAG AAACACAAAG TCTACGCCTG CGAAGTCACC
      L S K A   D Y E   K H K   V Y A C   E V T
2401  CATCAGGGCC TGAGCTCGCC CGTCACAAAG AGCTTCAACA GGGGAGAGTG
      H Q G   L S S P   V T K   S F N   R G E C
2451  TTAGAGGGAG AAGTGCCCCC ACCTGCTCCT CAGTTCCAGC CTGACCCCTT
      *                               Psp5 II
2501  CCCATCCTTT GGCTCTGAC  CCTTTTTCCT CAGGGGACCT ACCCCTATTG
2551  CGGTCCTCCA GCTCATCTTT CACCTCACCC CCCTCCTCCT CCTTGGCTTT
2601  AATTATGCTA ATGTTGGAGG AGAATGAATA AATAAAGTGA ATCTTTGCAC
2651  CTGTGGTGGA TCTAATAAAA GATATTTATT TTCATTAGAT ATGTGTGTTG
2701  GTTTTTTGTG TGCAGTGCCT CTATCTGGAG GCCAGGTAGG GCTGGCCTTG
2751  GGGGAGGGGG AGGCCAGAAT GACTCCAAGA GCTACAGGAA GGCAGGTCAG
2801  AGACCCCACT GGACAAACAG TGGCTGGACT CTGCACCATA ACACACAATC
2851  AACAGGGGAG TGAGCTGGAA ATTTGCTAGC GAATTCTTGA AGACGAAAGG
2901  GCCTCGTGAT ACGCCTATTT TTATAGGTTA ATGTCATGAT AATAATGGTT
2951  TCTTAGACGT CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT
3001  TTGTTTATTT TTCTAAATAC ATTCAAATAT GTATCCGCTC ATGAGACAAT
3051  AACCTTGATA AATGCTTCAA TAATATTGAA AAAGGAAGAG TATGAGTATT
3101  CAACATTTCCT GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC
3151  TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAAGATC
3201  AGTTGGGTGC ACGAGTGGGT TACATCGAAC TGGATCTCAA CAGCGGTAAG
3251  ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA TGAGCACTTT
3301  TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTGTTGAC GCCGGGCAAG
3351  AGCAACTCGG TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC
3401  TCACCAGTCA CAGAAAAGCA TCTTACGGAT GGCATGACAG TAAGAGAATT
3451  ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
      Pvu I
3501  TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG
3551  GGGGATCATG TAACTCGCCT TGATCGTTGG GAACCGGAGC TGAATGAAGC

```

Fig. 30 /4

3601 CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGCAGCA ATGGCAACAA
3651 CGTTGCGCAA ACTATTA ACT GGC GA ACTAC TTACTCTAGC TTCCCGGCAA
3701 CAATTAATAG ACTGGATGGA GGC GGATAAA GTTGCAGGAC CACTTCTGCG
3751 CTCGGCCCTT CCGGCTGGCT GGT TTTATTGC TGATAAATCT GGAGCCGGTG
3801 AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
3851 TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA
3901 ACGAAATAGA CAGATCGCTG AGATAGGTGC CTCACTGATT AAGCATTTGGT
3951 AACTGTCAGA CCAAGTTTAC TCATATATAC TTTAGATTGA TTTAAAACTT
4001 CATTTTAAAT TTAAAAGGAT CTAGGTGAAG ATCCTTTTTG ATAATCTCAT
4051 GACCAAAATC CCTTAACGTG AGTTTTCGTT CCACTGAGCG TCAGACCCCG
4101 TAGAAAAGAT CAAAGGATCT TCTTGAGATC CTTTTTTTCT GCGCGTAATC
4151 TGCTGCTTGC AAACAAAAAA ACCACCGCTA CCAGCGGTGG TTTGTTTGCC
4201 GGATCAAGAG CTACCAACTC TTTTCCGAA GGTA ACTGGC TTCAGCAGAG
4251 CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT AGGCCACCAC
4301 TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT
4351 ACCAGTGGCT GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGA CT
4401 CAAGACGATA GTTACCGGAT AAGGCGCAGC GGTCGGGCTG AACGGGGGGT
4451 TCGTGACAC AGCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
4501 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG
4551 CGGACAGGTA TCCGGTAAGC GGCAGGGTCG GAACAGGAGA GCGCACGAGG
4601 GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG TCGGGTTTCG
4651 CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA
4701 GCCTATGGAA AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT
4751 TGCTGGCCTT TTGCT CACAT GTTCTTTCCT GCGTTATCCC CTGATTCTGT
4801 GGATAACCGT ATTACCGCCT TTGAGTGAGC TGATAACCGCT CGCCGCAGCC

BspLU111

Fig. 30 /5

4851 GAACGACCGA GCGCAGCGAG TCAGTGAGCG AGGAAGCGGA AGAGCGCCTG
4901 ATGCGGTATT TTCTCCTTAC GCATCTGTGC GGTATTTTAC ACCGCATATG
4951 GTGCACTCTC AGTACAATCT GCTCTGATGC CGCATAGTTA AGCCAGTATA Bst1107I
5001 CACTCCGCTA TCGCTACGTG ACTGGGTCAT GGCTGCGCCC CGACACCCGC
5051 CAACACCCGC TGACGCGCCC TGACGGGCTT GTCTGCTCCC GGCATCCGCT
5101 TACAGACAAG CTGTGACCGT CTCCGGGAGC TGCATGTGTC AGAGGTTTTTC
5151 ACCGTCATCA CCGAAACGCG CGAGGCAGCT GTGGAATGTG TGTCAGTTAG
5201 GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG
5251 CATCTCAATT AGTCAGCAAC CAGGCTCCCC AGCAGGCAGA AGTATGCAAA
5301 GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT AACTCCGCCC
5351 ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC CCCATGGCTG
5401 ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG GCCTCTGAGC Sfi I
5451 TATTCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG Stu I/Avr II
5501 AAGCTAGCTT CACGCTGCCG CAAGCACTCA GGGCGCAAGG GCTGCTAAAG
5551 GAAGCGGAAC ACGTAGAAAG CCAGTCCGCA GAAACGGTGC TGACCCCGGA
5601 TGAATGTCAG CTA CTGGGCT ATCTGGACAA GGGAAAACGC AAGCGCAAAG
5651 AGAAAGCAGG TAGCTTGCAG TGGGCTTACA TGGCGATAGC TAGACTGGGC
5701 GGTTTTATGG ACAGCAAGCG AACCAGGAATT GCCAGCTGGG GCGCCCTCTG
5751 GTAAGGTTGG GAAGCCCTGC AAAGTAAACT GGATGGCTTT CTTGCCGCCA
5801 AGGATCTGAT GCGCAGGGG ATCAAGATCT GATCAAGAGA CAGGATGAGG Bgl II/Bcl I
5851 ATCGTTTCGC ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG
5901 CTTGGGTGGA GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC
5951 TGCTCTGATG CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT
6001 TTTTGTCAAG ACCGACCTGT CCGGTGCCCT GAATGAACTG CAGGACGAGG
6051 CAGCGCGGCT ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG Msc I

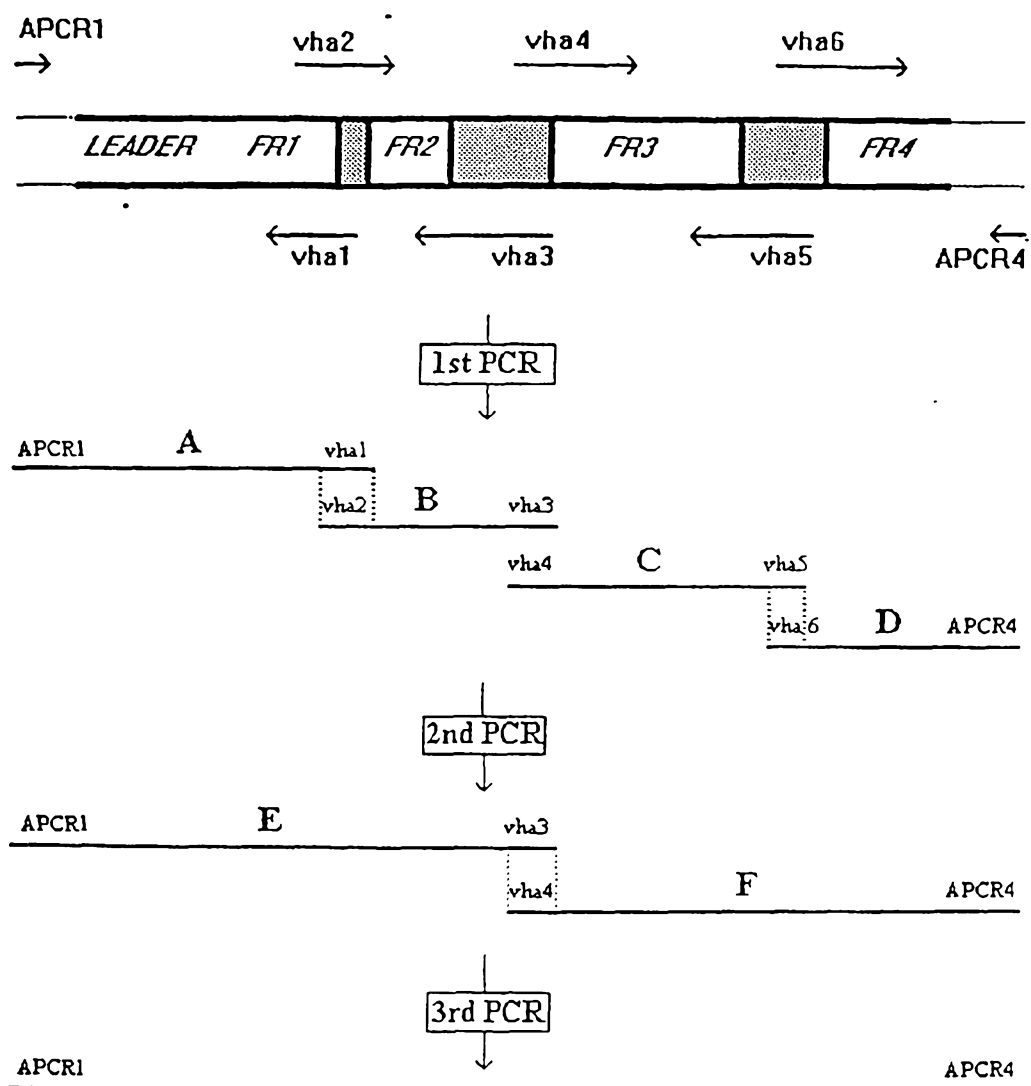
Fig. 30 /6

6101 CTCGACGTTG TCACTGAAGC GGAAGGGAC TGGCTGCTAT TGGGCGAAGT
6151 GCCGGGGCAG GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT
6201 CCATCATGGC TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC
6251 TGCCCATTCG ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG
6301 GATGGAAGCC GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG
6351 GGCTCGCGCC AGCCGAACTG TTCGCCAGGC TCAAGGCGCG CATGCCCGAC
6401 GGCGAGGATC TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT
6451 GGTGGAAAAT GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG
Rsr II
6501 TGGCGGACCG CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA
6551 GAGCTTGGCG GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC
6601 CGCTCCCGAT TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT
Nsp V
6651 TCTGAGCGGG ACTCTGGGGT TCGAAATGAC CGACCAAGCG ACGCCCAACC
6701 TGCCATCACG AGATTTCGAT TCCACCGCCG CCTTCTATGA AAGGTTGGGC
6751 TTCGGAATCG TTTTCCGGGA CGCCGGCTGG ATGATCCTCC AGCGCGGGGA
Sma I Nru I
6801 TCTCATGCTG GAGTTCTTCG CCCACCCCCG GCTCGATCCC CTCGCGAGTT
6851 GGTTCACTG CTGCCTGAGG CTGGACGACC TCGCGGAGTT CTACCGGCAG
6901 TGCAAATCCG TCGGCATCCA GGAAACCAGC AGCGGCTATC CGCGCATCCA
6951 TGCCCCCGAA CTGCAGGAGT GGGGAGGCAC GATGGCCGCT TTGGTCCCGG
7001 ATCTTTGTGA AGGAACCTTA CTTCTGTGGT GTGACATAAT TGGACAAACT
7051 ACCTACAGAG ATTTAAAGCT CTAAGGTAAA TATAAAATTT TTAAGTGTAT
7101 AATGTGTTAA ACTACTGATT CTAATTGTTT GTGTATTTTA GATTCCAACC
7151 TATGGAAGT ATGAATGGGA GCAGTGGTGG AATGCCTTTA ATGAGGAAAA
7201 CCTGTTTTGC TCAGAAGAAA TGCCATCTAG TGATGATGAG GCTACTGCTG
7251 ACTCTCAACA TTCTACTCCT CCAAAAAAGA AGAGAAAGGT AGAAGACCCC
7301 AAGGACTTTC CTTCAGAATT GCTAAGTTTT TTGAGTCATG CTGTGTTTAG

Fig. 30 /7

7351 TAATAGAACT CTTGCTTGCT TTGCTATTTA CACCACAAAG GAAAAAGCTG
7401 CACTGCTATA CAAGAAAATT ATGGAAAAAT ATTCTGTAAC CTTTATAAGT
7451 AGGCATAACA GTTATAATCA TAACATACTG TTTTTTCTTA CTCCACACAG
7501 GCATAGAGTG TCTGCTATTA ATAACTATGC TCAAAAATTG TGTACCTTTA
7551 GCTTTTTAAT TTGTAAAGGG GTTAATAAGG AATATTTGAT GTATAGTGCC
7601 TTGACTAGAG ATCATAATCA GCCATACCAC ATTTGTAGAG GTTTTACTTG
7651 CTTTAAAAAA CCTCCCACAC CTCCCCCTGA ACCTGAAACA TAAATGAAT
Mun I
7701 GCAATTGTTG TTGTAACTT GTTTATTGCA GCTTATAATG GTTACAAATA
7751 AAGCAATAGC ATCACAAATT TCACAAATAA AGCATTTTTT TCACTGCATT
7801 CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA TGTCTGGATC
7851 TAATAAAAGA TATTTATTTT CATTAGATAT GTGTGTTGGT TTTTGTGTG
7901 CAGTGCCTCT ATCTGGAGGC CAGGTAGGGC TGGCCTTGGG GGAGGGGGAG
7951 GCCAGAATGA CTCCAAGAGC TACAGGAAGG CAGGTCAGAG ACCCCACTGG
8001 ACAAACAGTG GCTGGACTCT GCACCATAAC ACACAATCAA CAGGGGAGTG
8051 AGCTGGAAAT TTGCTAGC

Fig. 31



[illegible]

Fig. 32 /2

	92				CDR3					100	A	B	C	D	I	J	K	101		103
	C	A	R		R	R	I	A	Y	G	Y	D	E	G	H	A	M	D	Y	W
A	TGC	GCC	AGA		AGA	AGA	ATC	GCC	TAT	GGT	TAC	GAC	GAG	GGC	CAT	GCT	ATG	GAC	TAC	TGG
B	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
C	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
D	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
E	.	.	.		.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

	104										113
	G	Q	G	T	L	V	T	V	S	S	
A	GGT	CAA	GGA	ACC	CTT	GTC	ACC	GTC	TCC	TCA	
B	.	.	.	.	.	.	.	.	.	.	
C	.	.	.	.	.	.	.	.	.	.	
D	.	.	.	.	.	.	.	.	.	.	
E	.	.	.	.	.	.	.	.	.	.	

Fig. 33 /1

1 TTGAAGACGAAAGGGCCTCGTGATACGCCTATTTTTATAGGTTAATGTCATGATAATAAT
 61 GGTTTCTTAGACGTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTT
 121 ATTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCT
 181 TCAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCC
 241 CTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAA
 301 AGATGCTGAAGATCAGTTGGGTGCACGAGTGGGTACATCGAACTGGATCTCAACAGCGG
 361 TAAGATCCTTGAGAGTTTTCGCCCCGAAGAACGTTTTTCCAATGATGAGCACTTTTAAAGT
 421 TCTGCTATGTGGCGCGGTATTATCCCGTGTTGACGCCGGGCAAGAGCAACTCGGTCGCCG
 481 CATACTATTCTCAGAATGACTTGGTTGAGTACTCACCAGTCACAGAAAAGCATCTTAC
 541 GGATGGCATGACAGTAAGAGAATTATGCAGTGCTGCCATAACCATGAGTGATAAACTGC
 Pvu I
 601 GGCCAACTTACTTCTGACAACGATCGGAGGACCGAAGGAGCTAACCGCTTTTTTGCACAA
 661 CATGGGGGATCATGTAACCTCGCCTTGATCGTTGGGAACCGGAGCTGAATGAAGCCATAAC
 Fsp I
 721 AAACGACGAGCGTGACACCACGATGCCTGCAGCAATGGCAACAACGTTGCGCAAACTATT
 781 AACTGGCGAACTACTTACTCTAGCTTCCCGGCAACAATTAATAGACTGGATGGAGGCGGA

Fig. 33 /2

841 TAAAGTTGCAGGACCACTTCTGCGCTCGGCCCTTCCGGCTGGCTGGTTTATTGCTGATAA
901 ATCTGGAGCCGGTGAGCGTGGGTCTCGCGGTATCATTGCAGCACTGGGGCCAGATGGTAA
961 GCCCTCCCGTATCGTAGTTATCTACACGACGGGGAGTCAGGCAACTATGGATGAACGAAA
1021 TAGACAGATCGCTGAGATAGGTGCCTCACTGATTAAGCATTGGTAACTGTCAGACCAAGT
1081 TTACTCATATATACTTTAGATTGATTTAAACTTCATTTTTAATTTAAAAGGATCTAGGT
1141 GAAGATCCTTTTTGATAATCTCATGACCAAAATCCCTTAACGTGAGTTTTCGTTCCACTG
1201 AGCGTCAGACCCCGTAGAAAAGATCAAAGGATCTTCTTGAGATCCTTTTTTTCTGCGCGT
1261 AATCTGCTGCTTGCAAACAAAAAACCACCGCTACCAGCGGTGGTTTGTTTGCCGGATCA
1321 AGAGCTACCAACTCTTTTTCCGAAGGTAAGTGGCTTCAGCAGAGCGCAGATACCAAATAC
1381 TGTCTTCTAGTGTAGCCGTAGTTAGGCCACCACTTCAAGAACTCTGTAGCACCGCCTAC
1441 ATACCTCGCTCTGCTAATCCTGTTACCAGTGGCTGCTGCCAGTGGCGATAAGTCGTGTCT
1501 TACCGGGTTGGACTCAAGACGATAGTTACCGGATAAGGCGCAGCGGTCCGGCTGAACGGG
1561 GGGTTCGTGCACACAGCCCAGCTTGGAGCGAACGACCTACACCGAACTGAGATACCTACA
1621 GCGTGAGCTATGAGAAAGCGCCACGCTTCCCGAAGGGAGAAAGGCGGACAGGTATCCGGT
1681 AAGCGGCAGGGTCGGAACAGGAGAGCGCACGAGGGAGCTTCCAGGGGGAAACGCCTGGTA
1741 TCTTTATAGTCCTGTGCGGGTTTCGCCACCTCTGACTTGAGCGTCGATTTTTGTGATGCTC
1801 GTCAGGGGGGCGGAGCCTATGGAAAACGCCAGCAACGCGGCCTTTTTACGGTTCCTGGC
1861 CTTTGTGCTGGCCTTTTGCTCACATGTTCTTTCCTGCGTTATCCCCTGATTCTGTGGATAA
1921 CCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCGCAGCCGAACGACCGAGCGCAG
1981 CGAGTCAGTGAGCGAGGAAGCGGAAGAGCGCCTGATGCGGTATTTTCTCCTTACGCATCT
2041 GTGCGGTATTTACACCGCATATGGTGCCTCTCAGTACAATCTGCTCTGATGCCGCATA
2101 GTTAAGCCAGTATACACTCCGCTATCGCTACGTGACTGGGTCATGGCTGCGCCCCGACAC
2161 CCGCCAACACCCGCTGACGCGCCCTGACGGGCTTGTCTGCTCCCGGCATCCGCTTACAGA
2221 CAAGCTGTGACCGTCTCCGGGAGCTGCATGTGTGTCAGAGGTTTTACCGTCATACCGAAA
2281 CGCGCGAGGCAGCATGCATCTCAATTAGTCAGCAACCATAGTCCCGCCCCTAACTCCGCC
2341 CATCCCGCCCCCTAACTCCGCCAGTTCCGCCATTCTCCGCCCATGGCTGACTAATTTT
2401 TTTTATTTATGCAGAGGCCGAGGCCGCTCGGCCTCTGAGCTATTCCAGAAGTAGTGAGG
2461 AGGCTTTTTTGGAGGCCTAGGCTTTTGC AAAAAGCTAGCTTACAGCTCAGGGCTGCGATT

Fig. 33 /3

2521 TCGCGCCAAACTTGACGGCAATCCTAGCGTGAAGGCTGGTAGGATTTTATCCCCGCTGCC
2581 ATCATGGTTCGACCATTGAACTGCATCGTCGCCGTGTCCCAAATATGGGGATTGGCAAG
2641 AACGGAGACCTACCCTGGCCTCCGCTCAGGAACGAGTTCAAGTACTTCCAAAGAATGACC
2701 ACAACCTCTTCAGTGGAAGGTAAACAGAATCTGGTGATTATGGGTAGGAAAACCTGGTTC
2761 TCCATTCTGAGAAGAATCGACCTTTAAAGGACAGAATTAATATAGTTCTCAGTAGAGAA
2821 CTCAAAGAACCACCACGAGGAGCTCATTTTCTTGCCAAAAGTTTGGATGATGCCTTAAGA
2881 CTTATTGAACAACCGGAATTGGCAAGTAAAGTAGACATGGTTTGGATAGTCGGAGGCAGT
2941 TCTGTTTACCAGGAAGCCATGAATCAACCAGGCCACCTCAGACTCTTTGTGACAAGGATC
3001 ATGCAGGAATTTGAAAGTGACACGTTTTTCCCAGAAATTGATTTGGGGAAATATAAACTT
3061 CTCCCAGAATACCCAGGCGTCCTCTCTGAGGTCCAGGAGGAAAAAGGCATCAAGTATAAG
3121 TTTGAAGTCTACGAGAAGAAAGACTAACAGGAAGATGCTTCAAGTTCTCTGCTCCCCTC
3181 CTAAAGCTATGCATTTTTATAAGACCATGGGACTTTTGCTGGCTTTAGATCTTTGTGAAG
3241 GAACCTTACTTCTGTGGTGTGACATAATTGGACAACTACCTACAGAGATTTAAAGCTCT
3301 AAGGTAAATATAAAATTTTTAAGTGTATAATGTGTTAACTACTGATTCTAATTGTTTGT
3361 GTATTTTAGATTCCAACCTATGGAAGTGAATGGGAGCAGTGGTGAATGCCTTTAAT
3421 GAGGAAAACCTGTTTTGCTCAGAAGAAATGCCATCTAGTGATGATGAGGCTACTGCTGAC
3481 TCTCAACATTCTACTCCTCCAAAAAGAAGAGAAAGGTAGAAGACCCCAAGGACTTTCCT
3541 TCAGAATTGCTAAGTTTTTTGAGTCATGCTGTGTTTAGTAATAGAACTCTTGCTTGCTTT
3601 GCTATTTACACCACAAAGGAAAAAGCTGCACTGCTATACAAGAAAATTATGGAAAAATAT
3661 TCTGTAACCTTTATAAGTAGGCATAACAGTTATAATCATAACATACTGTTTTTTCTTACT
3721 CCACACAGGCATAGAGTGTCTGCTATTAATAACTATGCTCAAAAATTGTGTACCTTTAGC
3781 TTTTAAATTTGTAAAGGGGTTAATAAGGAATATTTGATGTATAGTGCCTTGACTAGAGAT
3841 CATAATCAGCCATACCACATTTGTAGAGGTTTTACTTGCTTTAAAAAACCTCCACACCT
3901 CCCCCTGAACCTGAAACATAAAATGAATGCAATTGTTGTTGTTAACTTGTTTATTGCAGC
3961 TTATAATGGTTACAAATAAAGCAATAGCATCACAATTTACAAATAAAGCATTTTTTTTC
4021 ACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTGGATCTA
4081 ATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCCTCTAT
4141 CTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGAGGCCAGAATGACTCCAAGAGCTA

Fig. 33 /4

4201 CAGGAAGGCAGGTCAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCATAACAC
EcoR I
4261 ACAATCAACAGGGGAGTGAGCTGGAAATTTGCTAGCGAATTCcagcacactggcgccgt
(Spe I)
4321 tACTAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTT
4381 CCGCGTTACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCC
4441 ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCATTGACG
4501 TCAATGGGTGGAGTATTTACGGTAAACTGCCCACCTGGCAGTACATCAAGTGTATCATAT
4561 GCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA
SnaB I
4621 GTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT
4681 TACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGACTCACG
4741 GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAATCA
4801 ACGGGACTTTCAAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCG
4861 TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGATCGCCTGGAG
4921 ACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCCTCCGCGG
4981 CCGGGAACGGTGCATTGGAACGCGGATTCCCCGTGCCAAGAGTGACGTAAGTACCGCCTA
5041 TAGAGTCTATAGGCCACCCCCTTGGCTTCTTATGCATGCTATACTGTTTTTGGCTTGGG
Bpu1102I
5101 GTCTATACACCCCGCTTCCTCATGTTATAGGTGATGGTATAGCTTAGCCTATAGGTGTG
Xcm I
5161 GGTTATTGACCATTATTGACCACTCCCCTATTGGTGACGATACTTTCATTACTAATCCA
5221 TAACATGGCTCTTTGCCACAACCTCTCTTTATTGGCTATATGCCAATACACTGTCCTTCAG
5281 AGACTGACACGGACTCTGTATTTTTACAGGATGGGGTCTCATTTATTATTTACAAATTCA
5341 CATATACAACACCACCGTCCCCAGTGCCCGCAGTTTTTTATTAAACATAACGTGGGATCTC
BspE I
5401 CACGCGAATCTCGGGTACGTGTTTCCGGACATGGGCTCTTCTCCGGTAGCGGCGGAGCTTC
5461 TACATCCGAGCCCTGCTCCCATGCCTCCAGCGACTCATGGTCGCTCGGCAGCTCCTTGCT
5521 CCTAACAGTGGAGGCCAGACTTAGGCACAGCACGATGCCACCACCACCAGTGTGCCGCA
5581 CAAGGCCGTGGCGGTAGGGTATGTGTCTGAAAATGAGCTCggggagcgggcttgcaccgc
(Pvu II)
5641 tgacgcatttgaagacttaaggcagcggcagaagaagatgcaggcagctgagttgttgt
5701 gttctgataagagtcagaggtaactcccgttgcggtgctgttaacggtggagggcagtg
5761 agtctgagcagtagtactcgttgctgccgcgcgcgccaccagacataatagctgacagactaa
Mlu I
5821 cagactgttcctttccatgggtcctttctgcagtcaccgtccttgacACGCGTCTCGGGA

Fig. 33 /5

Hind III
5881 AGCTTGCCGCCACCATGGACTGGACCTGGCGCGTGTTTTGCCTGCTCGCCGTGGCTCCTG
M D W T W R V F C L L A V A P

5941 GGGCCACAGCCAGGTGCAACTGGTGCAGTCCGGCGCCGAAGTGAAGAAACCCGGTGCTT
G A H S Q V Q L V Q S G A E V K K P G A
(Pvu II) (Spe I)

6001 CCGTGAAAGTCAGCTGTAAAACTAGTAGATACACCTTCACTGAATACACCATACTAGGG
S V K V S C K T S R Y T F T E Y T I H W
Msc I CDR 1

6061 TTAGACAGGCCCTTGGCCAAAGGCTGGAGTGGATAGGAGGTATTAATCCTAACAATGGTA
V R Q A P G Q R L E W I G G I N P N N G

6121 TTCCTAACTACAACCAGAAGTTCAAGGGCCGGGCCACCTTGACCGTAGGCAAGTCTGCCA
I P N Y N Q K F K G R A T L T V G K S A
CDR 2

6181 GCACCGCTACATGGAAGTGTCCAGCCTGCGCTCCGAGGACACTGCAGTCTACTACTGCG
S T A Y M E L S S L R S E D T A V Y Y C

6241 CCAGAAGAAGAATCGCCTATGGTTACGACGAGGGCCATGCTATGGACTACTGGGGTCAAG
A R R R I A Y G Y D E G H A M D Y W G Q
CDR 3 BamH I

6301 GAACCTTGTACCGTCTCCTCAGGTGAGTGGATCCTCTGCGCCTGGGCCCAGCTCTGTC
G T L V T V S S

6361 CCACACCGCGGTACATGGCACCACCTCTCTTGAGCCTCCACCAAGGGCCCATCGGTCT
S T K G P S V

6421 TCCCCCTGGCACCTCCTCCAAGAGCACCTCTGGGGGCACAGCGGCCCTGGGCTGCCTGG
F P L A P S S K S T S G G T A A L G C L
Age I

6481 TCAAGGACTACTTCCCCGAACCGGTGACGGTGTCTGTGGAAGTCAAGCGCCCTGACCAGCG
V K D Y F P E P V T V S W N S G A L T S

6541 GCGTGCACACCTTCCCGGCTGTCTACAGTCCTCAGGACTCTACTCCCTCAGCAGCGTGG
G V H T F P A V L Q S S G L Y S L S S V
BstE II

6601 TGACCGTGCCCTCCAGCAGCTTGGGCACCCAGACCTACATCTGCAACGTGAATCACAAGC
V T V P S S S L G T Q T Y I C N V N H K

6661 CCAGCAACACCAAGGTGGACAAGAAAGTTGAGCCCAAATCTTGAGACAAAACCTCACACAT
P S N T K V D K K V E P K S C D K T H T

6721 GCCCACCCTGCCCAGCACCTGAACTCCTGGGGGGACCGTCAGTCTTCTCTTCCCCCAA
C P P C P A P E L L G G P S V F L F P P

6781 AACCCAAGGACACCCTCATGATCTCCCGGACCCCTGAGGTCACATGCGTGGTGGTGGACG
K P K D T L M I S R T P E V T C V V V D

6841 TGAGCCACGAAGACCCTGAGGTCAAGTTCAACTGGTACGTGGACGGCGTGGAGGTGCATA
V S H E D P E V K F N W Y V D G V E V H

6901 ATGCCAAGACAAAGCCGCGGGAGGAGCAGTACAACAGCACGTACCGGGTGGTCAGCGTCC
N A K T K P R E E Q Y N S T Y R V V S V

Fig. 33 /6

6961 TCACCGTCCTGCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGCAAGGTCTCCAACA
L T V L H Q D W L N G K E Y K C K V S N

7021 AAGCCCTCCCAGCCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAGCCCCGAGAAC
K A L P A P I E K T I S K A K G Q P R E

7081 CACAGGTGTACACCCTGCCCCCATCCCGGGAGGAGATGACCAAGAACCAGGTGAGCCTGA
P Q V Y T L P P S R E E M T K N Q V S L

7141 CCTGCCTGGTCAAAGGCTTCTATCCCAGCGACATCGCCGTGGAGTGGGAGAGCAATGGGC
T C L V K G F Y P S D I A V E W E S N G

7201 AGCCGGAGAACAACACTACAAGACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCC
Q P E N N Y K T T P P V L D S D G S F F

7261 TCTACAGCAAGCTCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTCTCATGCT
L Y S K L T V D K S R W Q Q G N V F S C

7321 CCGTGATGCATGAGGCTCTGCACAACCACTACACGCAGAAGAGCCTCTCCCTGTCTCCGG
S V M H E A L H N H Y T Q K S L S L S P

NgoM I

7381 GTAAATGAGTGCACGCGGCCGGCAAGCCCCGCTCCCCGGGCTCTCGCGGTGCGACGAGGAT
G K *

7441 GCTTGGCACGTACCCCTGTACATACTTCCCGGGCGCCAGCATGGAAATAAAGCACCGG

7501 ATCTAATAAAAGATATTTATTTTCATTAGATATGTGTGTTGGTTTTTTGTGTGCAGTGCC

7561 TCTATCTGGAGGCCAGGTAGGGCTGGCCTTGGGGGAGGGGGAGGCCAGAATGACTCCAAG

7621 AGCTACAGGAAGGCAGGTGAGAGACCCCACTGGACAAACAGTGGCTGGACTCTGCACCAT

7681 AACACACAATCAACAGGGGAGTGAGCTGGaaatttgctagcgaattaattc 7731

Fig. 34 A

INTRON

3' end V gene ----- 5' end of CH1

ACC GTC TCC TCA G::GTGAGTGGATCC-(N)₄₈-CCTCTCTTGACAG::CC-

T V S S *splice donor site* BamHI *splice acceptor site*

-TCC ACC AAG GGC

S T K G

⇓

ACC GTC TCC TCA G:::CC TCC ACC AAG GGC

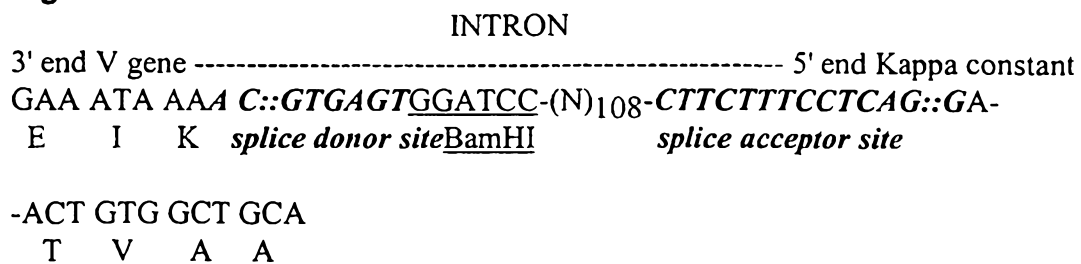
T V S S S T K G

⇓

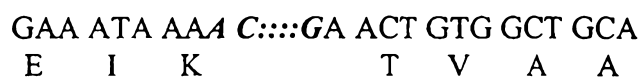
ACC GTC TCC TCA GCC TCC ACC AAG GGC

T V S S A S T K G

Fig. 34 B



⇓



⇓



Fig. 35

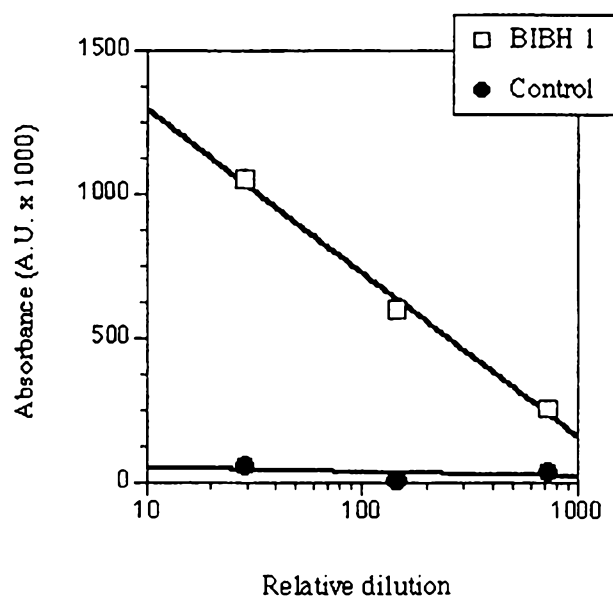


Fig. 36

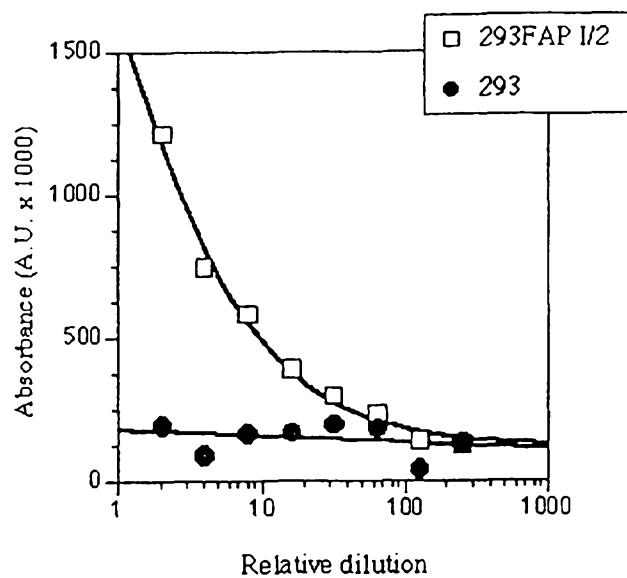


Fig. 37

