(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) **EP 1 143 006 B1**

(12) **EUROPEAN PATENT SPECIFICATION**

(45) Date of publication and mention of the grant of the patent:
02.04.2008 Bulletin 2008/14

(51) Int Cl.:
C12N 15/73 *(2006.01)* ***C12N 15/63*** *(2006.01)*

(21) Application number: **01110609.3**

(22) Date of filing: **19.08.1996**

(54) **Vectors/DNA-sequences from human combinatorial antibody libraries**

Vektoren/DNA-Sequenzen aus humanen kombinatorischen Antikörper-Bibliotheken

Vecteurs/DNA-séquences provenant de banques d'anticorps humaines

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC NL PT SE

(43) Date of publication of application:
10.10.2001 Bulletin 2001/41

(62) Document number(s) of the earlier application(s) in accordance with Art. 76 EPC:
96929278.8 / 0 859 841

(73) Proprietor: **MorphoSys IP GmbH**
82152 Martinsried (DE)

(72) Inventors:
- **Knappik, Achim**
  82166 Gräfelfing (DE)
- **Pack, Peter**
  68526 Ladenburg (DE)
- **Ge, Liming**
  81545 München (DE)

(74) Representative: **Vossius & Partner**
Siebertstrasse 4
81675 München (DE)

(56) References cited:
WO-A-91/13160

- **KRUIF DE J ET AL: "SELECTION AND APPLICATION OF HUMAN SINGLE CHAIN FV ANTIBODY FRAGMENTS FROM A SEMI-SYNTHETIC PHAGE ANTIBODY DISPLAY LIBRARY WITHDESIGNED CDR3 REGIONS" JOURNAL OF MOLECULAR BIOLOGY, LONDON, GB, vol. 248, 1995, pages 97-105, XP000984404 ISSN: 0022-2836**
- **WINTER G ET AL: "MAKING ANTIBODIES BY PHAGE DISPLAY TECHNOLOGY" ANNUAL REVIEW OF IMMUNOLOGY, ANNUAL REVIEWS INC, US, vol. 12, 1994, pages 433-455, XP000564245 ISSN: 0732-0582**
- **BARBAS C F ET AL: "SEMISYNTHETIC COMBINATORIAL ANTIBODY LIBRARIES: A CHEMICAL SOLUTION TO THE DIVERSITY PROBLEM" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 89, 1 May 1992 (1992-05-01), pages 4457-4461, XP002024223 ISSN: 0027-8424**
- **COLLET T A ET AL: "A BINARY PLASMID SYSTEM FOR SHUFFLING COMBINATORIAL ANTIBODY LIBRARIES" PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, NATIONAL ACADEMY OF SCIENCE. WASHINGTON, US, vol. 89, no. 21, 1 November 1992 (1992-11-01), pages 10026-10030, XP000322464 ISSN: 0027-8424**
- **VIRNEKAES BERNHARD ET AL: "Trinucleotide phosphoramidites: Ideal reagents for the synthesis of mixed oligonucleotides for random mutagenesis" NUCLEIC ACIDS RESEARCH, vol. 22, no. 25, 1994, pages 5600-5607, XP002901701 ISSN: 0305-1048**

EP 1 143 006 B1

Description

Field of the Invention

[0001] The present invention relates to synthetic DNA sequences which encode one or more collections of homologous proteins/(poly)peptides, and methods for generating and applying libraries of these DNA sequences. In particular, the invention relates to the preparation of a library of human-derived antibody genes by the use of synthetic consensus sequences which cover the structural repertoire of antibodies encoded in the human genome. Furthermore, the invention relates to the use of a single consensus antibody gene as a universal framework for highly diverse antibody libraries.

Background to the Invention

[0002] All current recombinant methods which use libraries of proteins/(poly)peptides, e.g. antibodies, to screen for members with desired properties, e.g. binding a given ligand, do not provide the possibility to improve the desired properties of the members in an easy and rapid manner. Usually a library is created either by inserting a random oligonucleotide sequence into one or more DNA sequences cloned from an organism, or a family of DNA sequences is cloned and used as the library. The library is then screened, e.g. using phage display, for members which show the desired property. The sequences of one or more of these resulting molecules are then determined. There is no general procedure available to improve these molecules further on.

[0003] Winter (EP 0 368 684 B1) has provided a method for amplifying (by PCR), cloning, and expressing antibody variable region genes. Starting with these genes he was able to create libraries of functional antibody fragments by randomizing the CDR3 of the heavy and/or the light chain. This process is functionally equivalent to the natural process of VJ and VDJ recombination which occurs during the development of B-cells in the immune system.

[0004] However the Winter invention does not provide a method for optimizing the binding affinities of antibody fragments further on, a process which would be functionally equivalent to the naturally occurring phenomenon of "affinity maturation", which is provided by the present invention. Furthermore, the Winter invention does not provide for artificial variable region genes, which represent a whole family of structurally similar natural genes, and which can be assembled from synthetic DNA oligonucleotides. Additionally, Winter does not enable the combinatorial assembly of portions of antibody variable regions, a feature which is provided by the present invention. Furthermore, this approach has the disadvantage that the genes of all antibodies obtained in the screening procedure have to be completely sequenced, since, except for the PCR priming regions, no additional sequence information about the library members is available. This is time and labor intensive and potentially leads to sequencing errors.

[0005] The teaching of Winter as well as other approaches have tried to create large antibody libraries having high diversity in the complementarity determining regions (CDRs) as well as in the frameworks to be able to find antibodies against as many different antigens as possible. It has been suggested that a single universal framework may be useful to build antibody libraries, but no approach has yet been successful.

[0006] John de Kruif et al., J. Mol. Biol. (April 1995), vol. 248, 97-105, disclose a semi-synthetic phage antibody display library with designed CDR3 regions.

[0007] Another problem lies in the production of reagents derived from antibodies. Small antibody fragments show exciting promise for use as therapeutic agents, diagnostic reagents, and for biochemical research. Thus, they are needed in large amounts, and the expression of antibody fragments, e.g. Fv, single-chain Fv (scFv), or Fab in the periplasm of E. coli (Skerra & Plückthun, 1988; Better et al., 1988) is now used routinely in many laboratories. Expression yields vary widely, however. While some fragments yield up to several mg of functional, soluble protein per liter and OD of culture broth in shake flask culture (Carter et al., 1992, Ptückthun et al. 1996), other fragments may almost exclusively lead to insoluble material, often found in so-called inclusion bodies. Functional protein may be obtained from the latter in modest yields by a laborious and time-consuming refolding process. The factors influencing antibody expression levels are still only poorty understood. Folding efficiency and stability of the antibody fragments, protease lability and toxicity of the expressed proteins to the host cells often severely limit actual production levels, and several attempts have been tried to increase expression yields. For example, Knappik & Plückthun (1995) could show that expression yield depends on the antibody sequence. They identified key residues in the antibody framework which influence expression yields dramatically. Similarly, Ullrich et al. (1995) found that point mutations in the CDRs can increase the yields in periplasmic antibody fragment expression. Nevertheless, these strategies are only applicable to a few antibodies. Since the Winter invention uses existing repertoires of antibodies, no influence on expressibility of the genes is possible.

[0008] Furthermore, the findings of Knappik & Plückthun and Ullrich demonstrate that the knowledge about antibodies, especially about folding and expression is still increasing. The Winter invention does not allow to incorporate such improvements into the library design.

[0009] The expressibility of the genes is important for the library quality as well, since the screening procedure relies in most cases on the display of the gene product on a phage surface, and efficient display relies on at least moderate

expression of the gene.

[0010] These disadvantages of the existing methodologies are overcome by the present invention.

[0011] The invention is a modular vector as defined in the claims, and the specification.

[0012] Artificial antibodies and fragments thereof can be constructed based on known antibody sequences, which reflect the structural properties of a whole group of homologous antibody genes. Therefore it is possible to reduce the number of different genes without any loss in the structural repertoire. This approach leads to a limited set of artificial genes, which can be synthesized de novo, thereby allowing introduction of cleavage sites and removing unwanted cleavages sites. Furthermore, this approach enables (i), adapting the codon usage of the genes to that of highly expressed genes in any desired host cell and (ii); analyzing all possible pairs of antibody light (L) and heavy (H) chains in terms of interaction preference, antigen preference or recombinant expression titer, which is virtually impossible using the complete collection of antibody genes of an organism and all combinations thereof.

[0013] The use of a limited set of completely synthetic genes makes it possible to create cleavage sites at the boundaries of encoded structural sub-elements. Therefore, each gene is built up from modules which represent structural sub-elements on the protein/(poly)peptide level. In the case of antibodies, the modules consist of "framework" and "CDR" modules. By creating separate framework and CDR modules, different combinatorial assembly possibilities are enabled. Moreover, if two or more artificial genes carry identical pairs of cleavage sites at the boundaries of each of the genetic sub-elements, pre-built libraries of sub-elements can be inserted in these genes simultaneously, without any additional information related to any particular gene sequence. This strategy enables rapid optimization of, for example, antibody affinity, since DNA cassettes encoding libraries of genetic sub-elements can be (i), pre-built, stored and reused and (ii), inserted in any of these sequences at the right position without knowing the actual sequence or having to determine the sequence of the individual library member.

[0014] Additionally, new information about amino acid residues important for binding, stability, or solubility and expression could be integrated into the library design by replacing existing modules with modules modified according to the new observations.

[0015] The limited number of consensus sequences used for creating the library allows to speed up the identification of binding antibodies after screening. After having identified the underlying consensus gene sequence, which could be done by sequencing or by using fingerprint restriction sites, just those part(s) comprising the random sequence(s) have to be determined. This reduces the probability of sequencing errors and of false-positive results.

[0016] The above mentioned cleavage sites can be used only if they are unique in the vector system where the artificial genes have been inserted. As a result, the vector has to be modified to contain none of these cleavage sites. The construction of a vector consisting of basic elements like resistance gene and origin of replication, where cleavage sites have been removed, is of general interest for many cloning attempts. Additionally, these vector(s) could be part of a kit comprising the above mentioned artificial genes and pre-built libraries.

[0017] The collection of artificial genes can be used for a rapid humanization procedure of non-human antibodies, preferably of rodent antibodies. First, the amino acid sequence of the non-human, preferably rodent antibody is compared with the amino acid sequences encoded by the collection of artificial genes to determine the most homologous light and heavy framework regions. These genes are then used for insertion of the genetic sub-elements encoding the CDRs of the non-human, preferably rodent antibody.

[0018] Surprisingly, it has been found that with a combination of only one consensus sequence for each of the light and heavy chains of a scFv fragment an antibody repertoire could be created yielding antibodies against virtually every antigen. The use of a single consensus sequence as a universal framework for the creation of useful (poly)peptide libraries and antibody consensus sequences useful therefor is disclosed.

Detailed Description of the Invention

[0019] The creation of useful libraries of (poly)peptides is disclosed. A method of setting up nucleic acid sequences suitable for the creation of said libraries is provided. In a first step, a collection of at least three homologous proteins is identified and then analyzed. Therefore, a database of the protein sequences is established where the protein sequences are aligned to each other. The database is used to define subgroups of protein sequences which show a high degree of similarity in both the sequence and, if information is available, in the structural arrangement. For each of the subgroups a (poly)peptide sequence comprising at least one consensus sequence is deduced which represents the members of this subgroup; the complete collection of (poly)peptide sequences represent therefore the complete structural repertoire of the collection of homologous proteins. These artificial (poly)peptide sequences are then analyzed, if possible, according to their structural properties to identify unfavorable interactions between amino acids within said (poly)peptide sequences or between said or other (poly)peptide sequences, for example, in multimeric proteins. Such interactions are then removed by changing the consensus sequence accordingly. The (poly)peptide sequences are then analyzed to identify sub-elements such as domains, loops, helices or CDRs. The amino acid sequence is backtranslated into a corresponding coding nucleic acid sequence which is adapted to the codon usage of the host planned for expressing said nucleic acid

sequences. A set of cleavage sites is set up in a way that each of the sub-sequences encoding the sub-elements identified as described above, is flanked by two sites which do not occur a second time within the nucleic acid sequence. This can be achieved by either identifying a cleavage site already flanking a sub-sequence of by changing one or more nucleotides to create the cleavage site, and by removing that site from the remaining part of the gene. The cleavage sites should be common to all corresponding sub-elements or sub-sequences, thus creating a fully modular arrangement of the sub-sequences in the nucleic acid sequence and of the sub-elements in the corresponding (poly)peptide.

[0020] Provided is a method of setting up one or more nucleic acid sequences encoding one or more (poly)peptide sequences suitable for the creation of libraries of (poly)peptides said (poly)peptide sequences comprising amino acid consensus sequences, said method comprising the following steps:

(a) deducing from a collection of at least three homologous proteins one or more (poly)peptide sequences comprising at least one amino acid consensus sequence;
(b) optionally, identifying amino acids in said (poly)peptide sequences to be modified so as to remove unfavorable interactions between amino acids within or between said or other (poly)peptide sequences;
(c) identifying at least one structural sub-element within each of said (poly)peptide sequences;
(d) backtranslating each of said (poly)peptide sequences into a corresponding coding nucleic acid sequence;
(e) setting up cleavage sites in regions adjacent to or between the ends of sub-sequences encoding said sub-elements, each of said cleavage sites:

(ea) being unique within each of said coding nucleic acid sequences;
(eb) being common to the corresponding sub-sequences of any said coding nucleic acids.

[0021] Further disclosed is a method which sets up two or more sets of (poly)peptides, where for each set the method as described above is performed, and where the cleavage sites are not only unique within each set but also between any two sets. This method can be applied for the creation of (poly)peptide libraries comprising for example two α-helical domains from two different proteins, where said library is screened for novel hetero-association domains.

[0022] Particularly preferred is a method of setting up two or more sets of one or more nucleic acid sequences comprising executing the steps (a) to (e) described above for each of said sets with the additional provision that said cleavage sites are unique between said sets.

[0023] A least two of the sets as described above, are derived from the same collection of proteins or at least a part of it. This describes libraries comprising for example, but not limited to, two domains from antibodies such as VH and VL, or two extracellular loops of transmembrane receptors.

[0024] The nucleic acid sequences set up as described above, are synthesized. This can be achieved by any one of several methods well known to the practitioner skilled in the art, for example, by total gene synthesis or by PCR-based approaches.

[0025] The nucleic acid sequences are cloned into a vector. The vector could be a sequencing vector, an expression vector or a display (e.g. phage display) vector, which are well known to those skilled in the art. Any vector could comprise one nucleic acid sequence, or two or more nucleic sequences, either in different or the same operon. In the last case, they could either be cloned separately or as contiguous sequences.

[0026] The removal of unfavorable interactions as described above, leads to enhanced expression of the modified (poly)peptides.

[0027] One or more sub-sequences of the nucleic acid sequences are replaced by different sequences. This can be achieved by excising the sub-sequences using the conditions suitable for cleaving the cleavage sites adjacent to or at the end of the sub-sequence, for example, by using a restriction enzyme at the corresponding restriction site under the conditions well known to those skilled in the art, and replacing the sub-sequence by a different sequence compatible with the cleaved nucleic acid sequence.

[0028] Said different sequences are selected from the group of different sub-sequences encoding the same or different sub-elements derived from the same or different (poly)peptides:

[0029] The different sequences replacing the initial sub-sequence(s) are genomic or rearranged genomic sequences, for example in grafting CDRs from non-human antibodies onto consensus antibody sequences for rapid humanization of non-human antibodies. In the most preferred embodiment, the different sequences are random sequences, thus replacing the sub-sequence by a collection of sequences to introduce variability and to create a library. The random sequences can be assembled in various ways, for example by using a mixture of mononucleotides or preferably a mixture of trinucleotides (Virnekäs et al., 1994) during automated oligonucleotide synthesis, by error-prone PCR or by other methods well known to the practitioner in the art. The random sequences may be completely randomized or biased towards or against certain codons according to the amino acid distribution at certain positions in known protein sequences. Additionally, the collection of random sub-sequences may comprise different numbers of codons, giving rise to a collection of sub-elements having different lengths.

[0030] The nucleic acid sequences can be expressed from a suitable vector and under suitable conditions well known to those skilled in the art.

[0031] The (poly)peptides expressed from said nucleic acid sequences are screened and, optionally, optimized. Screening may be performed by using one of the methods well known to the practitioner in the art, such as phage-display, selectively infective phage, polysome technology to screen for binding, assay systems for enzymatic activity or protein stability. (Poly)peptides having the desired property can be identified by sequencing of the corresponding nucleic acid sequence or by amino acid sequencing or mass spectrometry. In the case of subsequent optimization, the nucleic acid sequences encoding the initially selected (poly)peptides can optionally be used without sequencing. Optimization is performed by repeating the replacement of sub-sequences by different sequences, preferably by random sequences, and the screening step one or more times.

[0032] The desired property the (poly)peptides are screened for is preferably, but not exclusively, selected from the group of optimized affinity or specificity for a target molecule, optimized enzymatic activity, optimized expression yields, optimized stability and optimized solubility.

[0033] The cleavage sites flanking the sub-sequences are sites recognized and cleaved by restriction enzymes, with recognition and cleavage sequences being either identical or different, the restricted sites either having blunt or sticky ends.

[0034] The length of the sub-elements is preferably, but not exclusively ranging between 1 amino acid, such as one residue in the active site of an enzyme or a structure-determining residue, and 150 amino acids, as for whole protein domains. Most preferably, the length ranges between 3 and 25 amino acids, such as most commonly found in COR loops of antibodies.

[0035] The nucleic acid sequences could be RNA or, preferably, DNA.

[0036] The (poly)peptides have an amino acid pattern characteristic of a particular species. This can for example be achieved by deducing the consensus sequences from a collection of homologous proteins of just one species, most preferably from a collection of human proteins. Since the (poly)peptides comprising consensus sequences are artificial, they have to be compared to the protein sequence(s) having the closest similarity to ensure the presence of said characteristic amino acid pattern.

[0037] The creation of libraries of (poly)peptides comprising at least part of members or derivatives of the immunoglobulin superfamily is disclosed preferably of member or derivatives of the immunoglobulins. Most preferably, the invention provides for the creation of libraries of human antibodies, wherein said (poly)peptides are or are derived from heavy or light chain variable regions wherein said structural sub-elements are framework regions (FR) 1, 2, 3, or 4 or complementary determining regions (CDR) 1, 2, or 3. In a first step, a database of published antibody sequences of human origin is established where the antibody sequences are aligned to each other. The database is used to define subgroups of antibody sequences which show a high degree of similarity in both the sequence and the canonical fold of CDR loops (as determined by analysis of antibody structures). For each of the subgroups a consensus sequence is deduced which represents the members of this subgroup; the complete collection of consensus sequences represent therefore the complete structural repertoire of human antibodies.

These artificial genes are then constructed e.g. by total gene synthesis or by the use of synthetic genetic subunits. These genetic subunits correspond to structural sub-elements on the (poly)peptide level. On the DNA level, these genetic subunits are defined by cleavage sites at the start and the end of each of the sub-elements, which are unique in the vector system. All genes which are members of the collection of consensus sequences are constructed such that they contain a similar pattern of corresponding genetic sub-sequences. Most preferably, said (poly)peptides are or are derived from the HuCAL consensus genes: Vκ1, Vκ2, Vκ3, Vκ4, Vλ1, Vλ2, Vλ3, VH1A, VH1B, VH2, VH3, VH4, VH5, VH6, Cκ, Cλ, CH1 or any combination of said HuCAL consensus genes.

This collection of DNA molecules can then be used to create libraries of antibodies or antibody fragments, preferably Fv, disulphide-linked Fv, single-chain Fv (scFv), or Fab fragments, which may be used as sources of specificities against new target antigens. Moreover, the affinity of the antibodies can be optimized using pre-built library cassettes and a general procedure. A method for identifying one or more genes encoding one or more antibody fragments which binds to a target, is disclosed comprising the steps of expressing the antibody fragments, and then screening them to isolate one or more antibody fragments which bind to a given target molecule. Preferably, an scFv fragment library comprising the combination of HuCAL VH3 and HuCAL Vλ2 consensus genes and at least a random sub-sequence encoding the heavy chain CDR3 sub-element is screened for binding antibodies. If necessary, the modular design of the genes can then be used to excise from the genes encoding the antibody fragments one or more genetic sub-sequences encoding structural sub-elements, and replacing them by one or more second sub-sequences encoding structural sub-elements. The expression and screening steps can then be repeated until an antibody having the desired affinity is generated. Particularly preferred is a method in which one or more of the genetic subunits (e.g. the CDRs) are replaced by a random collection of sequences (the library) using the said cleavage sites. Since these cleavage sites are (i) unique in the vector system and (ii) common to all consensus genes, the same (pre-built) library can be inserted into all artificial antibody genes. The resulting library is then screened against any chosen antigen. Binding antibodies are selected, collected

and used as starting material for the next library. Here, one or more of the remaining genetic subunits are randomized as described above.

[0038] A further embodiment relates to fusion proteins by providing for a DNA sequence which encodes both the (poly) peptide, as described above, as well as an additional moiety. Particularly preferred are moieties which have a useful therapeutic function. For example, the additional moiety may be a toxin molecule which is able to kill cells (Vitetta et al., 1993). There are numerous examples of such toxins, well known to the one skilled in the art, such as the bacterial toxins Pseudomonas exotoxin A, and diphtheria toxin, as well as the plant toxins ricin, abrin, modeccin, saporin, and gelonin. By fusing such a toxin for example to an antibody fragment, the toxin can be targeted to, for example, diseased cells, and thereby have a beneficial therapeutic effect. Alternatively, the additional moiety may be a cytokine, such as IL-2 (Rosenberg & Lotze, 1986), which has a particular effect (in this case a T-cell proliferative effect) on a family of cells. In a further embodiment, the additional moiety may confer on its (poly)peptide partner a means of detection and/or purification. For example, the fusion protein could comprise the modified antibody fragment and an enzyme commonly used for detection purposes, such as alkaline phosphatase (Blake et al., 1984). There are numerous other moieties which can be used as detection or purification tags, which are well known to the practitioner skilled in the art. Particularly preferred are peptides comprising at least five histidine residues (Hochuli et al., 1988), which are able to bind to metal ions, and can therefore be used for the purification of the protein to which they are fused (Lindner et al., 1992). Also provided for are additional moieties such as the commonly used C-myc and FLAG tags (Hopp et al., 1988; Knappik & Plückthun, 1994).

[0039] Disclosed is a method, wherein at least part of said (poly)peptide sequences or (poly)peptides is connected to a sequence encoding at least one additional moiety or to at least one additional moiety, respectively.

[0040] Particularly preferred is a method, wherein said connection is formed via a contiguous nucleic acid sequence or amino acid sequence, respectively.

[0041] Most preferably, said additional moiety is a toxin, a cytokine, a reporter enzyme, a moiety being capable of binding a metal ion, a peptide, a tag suitable for detection and/or purification, or a homo- or hetero-association domain.

[0042] Particularly preferred is a method wherein the expression of said nucleic acid sequences results in the generation of a repertoire of biological activities and/or specificities, preferably in the generation of a repertoire based on a universal framework.

[0043] By engineering one or more fused additional domains, antibody fragments or any other (poly)peptide can be assembled into larger molecules which also fall under the scope of the present invention. For example, mini-antibodies (Pack, 1994) are dimers comprising two antibody fragments, each fused to a self-associating dimerization domain. Dimerization domains which are particularly preferred include those derived from a leucine zipper (Pack & Plückthun, 1992) or helix-turn-helix motif (Pack et al., 1993).

[0044] All of the above embodiments of the present invention can be effected using standard techniques of molecular biology known to anyone skilled in the art.

[0045] In a further embodiment, the random collection of sub-sequences (the library) is inserted into a singular nucleic acid sequence encoding one (poly)peptide, thus creating a (poly)peptide library based on one universal framework. Preferably a random collection of CDR sub-sequences is inserted into a universal antibody framework, for example into the HuCAL H3κ2 single-chain Fv fragment described above.

[0046] Disclosed are nucleic acid sequence(s), vector(s) containing the nucleic acid sequence(s), host cell(s) containing the vector(s), and (poly)peptides, obtainable according to the methods described above.

[0047] A method of producing a (poly)peptide or a collection of (poly)peptides as defined above comprising culturing a host cell of the present invention or a collection of host cells according to the present invention under suitable conditions and isolating said (poly)peptide or said collection of (poly)peptides is disclosed.

[0048] The present disclosure relates to a (poly)peptide devisable by the method according to the present invention, encoded by a nucleic acid sequence according to the present invention or obtainable by a method according to the present invention.

[0049] The disclosure relates to a collection of (poly)peptides devisable by a method according to the present invention, encoded by a collection of nucleic acid sequences according to the present invention or obtainable by a method according to the present invention.

[0050] The invention provides for modular vector systems, as disclosed in the claims, being compatible with the modular nucleic acid sequences encoding the (poly)peptides. The modules of the vectors are flanked by restriction sites unique within the vector system and essentially unique with respect to the restriction sites incorporated into the nucleic acid sequences encoding the (poly)peptides, except for example the restriction sites necessary for cloning the nucleic acid sequences into the vector. The list of vector modules comprises origins of single-stranded replication, origins of double-stranded replication for high- and low copy number plasmids, promoter/operator, repressor or terminator elements, resistance genes, potential recombination sites, gene III for display on filamentous phages, signal sequences, purification and detection tags, and sequences of additional moieties. The vectors are preferably, but not exclusively, expression vectors or vectors suitable for expression and screening of libraries.

[0051] Disclosed is a kit, comprising one or more of the list of nucleic acid sequence(s), recombinant vector(s), (pbly) peptide(s), and vector(s) according to the methods described above, and suitable host cell(s) for producing the (poly) peptide(s).

[0052] Disclosed is the creation of libraries of human antibodies. In a first step, a database of published antibody sequences of human origin is established. The database is used to define subgroups of antibody sequences which show a high degree of similarity in both the sequence and the canonical fold (as determined by analysis of antibody structures). For each of the subgroups a consensus sequence is deduced which represents the members of this subgroup; the complete collection of consensus sequences represent therefore the complete structural repertoire of human antibodies.

[0053] These artificial genes are then constructed by the use of synthetic genetic subunits. These genetic subunits correspond to structural sub-elements on the protein level. On the DNA level, these genetic subunits are defined by cleavage sites at the start and the end of each of the sub-elements, which are unique in the vector system. All genes which are members of the collection of consensus sequences are constructed such that they contain a similar pattern of said genetic subunits.

[0054] This collection of DNA molecules can then be used to create libraries of antibodies which may be used as sources of specificities against new target antigens. Moreover, the affinity of the antibodies can be optimised using pre-built library cassettes and a general procedure. The invention provides a method for identifying one or more genes encoding one or more antibody fragments which binds to a target, comprising the steps of expressing the antibody fragments, and then screening them to isolate one or more antibody fragments which bind to a given target molecule. If necessary, the modular design of the genes can then be used to excise from the genes encoding the antibody fragments one or more genetic sub-sequences encoding structural sub-elements, and replacing them by one or more second sub-sequences encoding structural sub-elements. The expression and screening steps can then be repeated until an antibody having the desired affinity is generated.

[0055] Particularly preferred is a method in which one or more of the genetic subunits (e.g. the CDR's) are replaced by a random collection of sequences (the library) using the said cleavage sites. Since these cleavage sites are (i) unique in the vector system and (ii) common to all consensus genes, the same (pre-built) library can be inserted into all artificial antibody genes. The resulting library is then screened against any chosen antigen. Binding antibodies are eluted, collected and used as starting material for the next library. Here, one or more of the remaining genetic subunits are randomised as described above.

[0056] Disclosed is a method of designing two or more genes encoding a collection of two or more proteins, comprising the steps of:

(a) either

(aa) identifying two or more homologous gene sequences, or
(ab) analyzing at least three homologous genes, and

deducing two or more consensus gene sequences therefrom,
(b) optionally, modifying codons in said consensus gene sequences to remove unfavourable interactions between amino acids in the resulting proteins,
(c) identifying sub-sequences which encode structural sub-elements in said consensus gene sequences
(d) modifying one or more bases in regions adjacent to or between the ends of said sub-sequences to define one or more cleavage sites, each of which:

(da) are unique within each consensus gene sequence,
(db) do not form compatible sites with respect to any single sub-sequence,
(dc) are common to all homologous sub-sequences.

[0057] Further disclosed is a method of preparing two or more genes encoding a collection of two or more proteins, comprising the steps of :

(a) designing said genes according to the present invention, and
(b) synthesizing said genes.

[0058] Furthermore, disclosed is a collection of genes prepared according to the method of the present invention.

[0059] Furthermore disclosed is a collection of two or more genes derived from gene sequences which:

(a) are either homologous, or represent consensus gene sequences derived from at least three homologous genes, and

(b) carry cleavage sites, each of which:

(ba) lie at or adjacent to the ends of genetic sub-sequences which encode structural sub-elements,
(bb) are unique within each gene sequence,
(bc) do not form compatible sites with respect to any single sub-sequence, and
(bd) are common to all homologous sub-sequences.

[0060] Also disclosed is a collection of genes, in which each of said gene sequences has a nucleotide composition characteristic of a particular species.

[0061] Said species is preferably human.

[0062] Further disclosed is a collection of genes according to the present invention, in which one or more of said gene sequences encodes at least part of a member of the immunoglobulin superfamily, preferably of the immunoglobulin family.

[0063] Said structural sub-elements correspond preferably to any combination of framework regions 1, 2, 3, and 4, and/or CDR regions 1, 2, and 3 of antibody heavy chains.

[0064] Disclosed is also that said structural sub-elements correspond to any combination of framework regions 1,2,3, and 4, and/or CDR regions 1, 2, and 3 of antibody light chains.

[0065] Furthermore, disclosed is a collection of vectors comprising a collection of gene sequences according to the present invention.

[0066] The collection of vectors comprises the additional feature that the vector does not comprise any cleavage site that is contained in the collection of genes according to the present invention.

[0067] Disclosed is a method for identifying one or more genes encoding one or more proteins having a desirable property, comprising the steps of:

(a) expressing from a collection of vectors according to the present invention a collection of proteins.
(b) screening said collection to isolate one or more proteins having a desired property,
(c) identifying the genes encoding the proteins isolated in step (b),
(d) optionally, excising from the genes encoding the proteins isolated in step (b) one or more genetic sub-sequences encoding structural sub-elements, and replacing said sub-sequence(s) by one or more second sub-sequences encoding structural sub-elements, to generate new
vectors according to the present invention,
(e) optionally, repeating steps (a) to (c).

[0068] Provided is a method for identifying one or more genes encoding one or more antibody fragments which binds to a target, comprising the steps of:

(a) expressing from the collection of vectors according to the present invention a collection of proteins,
(b) screening said collection to isolate one or more antibody fragments which bind to said target,
(c) identifying the genes encoding the proteins isolated in step (b),
(d) optionally, excising from the genes encoding the antibody fragments isolated in step (b) one or more genetic sub-sequences encoding structural sub-elements, and replacing said sub-sequence(s) by one or more second sub-sequences encoding structural sub-generate new vectors according to according to the present invention,
(e) optionally, repeating steps (a) to (c).

[0069] The present specification also discloses a kit comprising two or more genes derived from gene sequences which:

(a) are either homologous, or represent consensus gene sequences derived from at least three homologous genes, and
(b) carry cleavage sites, each of which:

(ba) lie at or adjacent to the ends of genetic sub-sequences which encode structural sub-elements,
(bb) are unique within each gene sequence,
(bc) do not form compatible sites with respect to any single sub-sequence, and
(bd) are common to all homologous sub-sequences.

[0070] The specification also discloses a kit comprising two or more genetic sub-sequences which encode structural sub-elements, which can be assembled to form genes, and which carry cleavage sites, each of which:

(a) lie at or adjacent to the ends of said genetic sub-sequences,

(b) do not form compatible sites with respect to any single sub-sequence, and
(d) are common to all homologous sub-sequences.

Definitions

Protein:

[0071] The term protein comprises monomeric polypeptide chains as well as homo- or heteromultimeric complexes of two or more polypeptide chains connected either by covalent interactions (such as disulphide bonds) or by non-covalent interactions (such as hydrophobic or electrostatic interactions).

Analysis of homologous proteins:

[0072] The amino acid sequences of three or more proteins are aligned to each other (allowing for introduction of gaps) in a way which maximizes the correspondence between identical or similar amino acid residues at all positions. These aligned sequences are termed homologous if the percentage of the sum of identical and/or similar residues exceeds a defined threshold. This threshold is commonly regarded by those skilled in the art as being exceeded when at least 15% of the amino acids in the aligned genes are identical, and at least 30% are similar. Examples for families of homologous proteins are: immunoglobulin superfamily, scavenger receptor superfamily, fibronectin superfamilies (e.g. type II and III), complement control protein superfamily, cytokine receptor superfamily, cystine knot proteins, tyrosine kinases, and numerous other examples well known to one of ordinary skill in the art.

Consensus sequence:

[0073] Using a matrix of at least three aligned amino acid sequences, and allowing for gaps in the alignment, it is possible to determine the most frequent amino acid residue at each position. The consensus sequence is that sequence which comprises the amino acids which are most frequently represented at each position. In the event that two or more amino acids are equally represented at a single position, the consensus sequence includes both or all of those amino acids.

Removing unfavorable interactions:

[0074] The consensus sequence is per se in most cases artificial and has to be analyzed in order to change amino acid residues which, for example, would prevent the resulting molecule to adapt a functional tertiary structure or which would block the interaction with other (poly)peptide chains in multimeric complexes. This can be done either by (i) building a three-dimensional model of the consensus sequence using known related structures as a template, and identifying amino acid residues within the model which may interact unfavorably with each other, or (ii) analyzing the matrix of aligned amino acid sequences in order to detect combinations of amino acid residues within the sequences which frequently occur together in one sequence and are therefore likely to interact with each other. These probable interaction-pairs are then tabulated and the consensus is compared with these "interaction maps". Missing or wrong interactions in the consensus are repaired accordingly by introducing appropriate changes in amino acids which minimize unfavorable interactions.

Identification of structural sub-elements:

[0075] Structural sub-elements are stretches of amino acid residues within a protein/(poly)peptide which correspond to a defined structural or functional part of the molecule. These can be loops (e.g. CDR loops of an antibody) or any other secondary or functional structure within the proteinl(poly)peptide (domains, α-helices, β-sheets, framework regions of antibodies, etc.). A structural sub-element can be identified using known structures of similar or homologous (poly) peptides, or by using the above mentioned matrices of aligned amino acid sequences. Here the variability at each position is the basis for determining stretches of amino acid residues which belong to a structural sub-element (e.g. hypervariable regions of an antibody).

Sub-sequence:

[0076] A sub-sequence is defined as a genetic module which is flanked by unique cleavage sites and encodes at least one structural sub-element. It is not necessarily identical to a structural sub-element.

Cleavage site:

[0077] A short DNA sequence which is used as a specific target for a reagent which cleaves DNA in a sequence-specific manner (e.g. restriction endonucleases).

Compatible cleavage sites:

[0078] Cleavage sites are compatible with each other, if they can be efficiently ligated without modification and, preferably, also without adding an adapter molecule.

Unique cleavage sites:

[0079] A cleavage site is defined as unique if it occurs only once in a vector containing at least one of the genes of interest, or if a vector containing at least one of the genes of interest could be treated in a way that only one of the cleavage sites could be used by the cleaving agent.

Corresponding (poly)peptide sequences:

[0080] Sequences deduced from the same part of one group of homologous proteins are called corresponding (poly) peptide sequences.

Common cleavage sites:

[0081] A cleavage site in at least two corresponding sequences, which occurs at the same functional position (i.e. which flanks a defined sub-sequence), which can be hydrolyzed by the same cleavage tool and which yields identical compatible ends is termed a common cleavage site.

Excising genetic sub-sequences:

[0082] A method which uses the unique cleavage sites and the corresponding cleavage reagents to cleave the target DNA at the specified positions in order to isolate, remove or replace the genetic sub-sequence flanked by these unique cleavage sites.

Exchanging genetic sub-sequences:

[0083] A method by which an existing sub-sequence is removed using the flanking cleavage sites of this sub-sequence, and a new sub-sequence or a collection of sub-sequences, which contain ends compatible with the cleavage sites thus created, is inserted.

Expression of genes:

[0084] The term expression refers to in vivo or in vitro processes, by which the information of a gene is transcribed into mRNA and then translated into a protein/(poly)peptide. Thus, the term expression refers to a process which occurs inside cells, by which the information of a gene is transcribed into mRNA and then into a protein. The term expression also includes all events of post-translational modification and transport, which are necessary for the (poly)peptide to be functional.

Screening of protein/(poly)peptide libraries:

[0085] Any method which allows isolation of one or more proteins/(poly)peptides having a desired property from other proteinsl(poly)peptides within a library.

Amino acid pattern characteristic for a species:

[0086] A (poly)peptide sequence is assumed to exhibit an amino acid pattern characteristic for a species if it is deduced from a collection of homologous proteins from just this species.

Immunoglobulin superfamily (IgSF):

[0087] The IgSF is a family of proteins comprising domains being characterized by the immunoglobulin fold. The IgSF comprises for example T-cell receptors and the immunoglobulins (antibodies).

Antibody framework:

[0088] A framework of an antibody variable domain is defined by Kabat et al. (1991) as the part of the variable domain which serves as a scaffold for the antigen binding loops of this variable domain.

Antibody CDR:

[0089] The CDRs (complementarity determining regions) of an antibody consist of the antigen binding loops, as defined by Kabat et al. (1991). Each of the two variable domains of an antibody Fv fragment contain three CDRs.

HuCAL:

[0090] Acronym for Human Combinatorial Antibody Library. Antibody Library based on modular consensus genes according to the invention (see Example 1).

Antibody fragment:

[0091] Any portion of an antibody which has a particular function, e.g. binding of antigen. Usually, antibody fragments are smaller than whole antibodies. Examples are Fv, disulphide-linked Fv, single-chain Fv (scFv), or Fab fragments. Additionally, antibody fragments are often engineered to include new functions or properties.

Universal framework:

[0092] One single framework which can be used to create the full variability of functions, specificities or properties which is originally sustained by a large collection of different frameworks, is called universal framework.

Binding of an antibody to its target:

[0093] The process which leads to a tight and specific association between an antibody and a corresponding molecule or ligand is called binding. A molecule or ligand or any part of a molecukle or ligand which is recognized by an antibody is called the target.

Replacing genetic sub-sequences

[0094] A method by which an existing sub-sequence is removed using the flanking cleavage sites of this sub-sequence, and a new sub-sequence or collection of sub-sequences, which contains ends compatible with the cleavage sites thus created, is inserted.

Assembling of genetic sequences:

[0095] Any process which is used to combine synthetic or natural genetic sequences in a specific manner in order to get longer genetic sequences which contain at least parts of the used synthetic or natural genetic sequences.

Analysis of homologous genes:

[0096] The corresponding amino acid sequences of two or more genes are aligned to each other in a way which maximizes the correspondence between identical or similar amino acid residues at all positions. These aligned sequences are termed homologous if the percentage of the sum of identical and/or similar residues exceeds a defined threshold. This threshold is commonly regarded by those skilled in the art as being exceeded when at least 15 per cent of the amino acids in the aligned genes are identical, and at least 30 per cent are similar.

Legends to Figures and Tables

[0097]

Fig. 1: Flow chart outlining the process of construction of a synthetic human antibody library based on consensus sequences.

Fig. 2: Alignment of consensus sequences designed for each subgroup (amino acid residues are shown with their standard one-letter abbreviation). (**A**) kappa sequences, (**B**) lambda sequences and (**C**), heavy chain sequences. The positions are numbered according to Kabat (1991). In order to maximize homology in the alignment, gaps (-) have been introduced in the sequence at certain positions.

Fig. 3: Gene sequences of the synthetic V kappa consensus genes. The corresponding amino acid sequences (see Fig. 2) as well as the unique cleavage sites are also shown.

Fig. 4: Gene sequences of the synthetic V lambda consensus genes. The corresponding amino acid sequences (see Fig. 2) as well as the unique cleavage sites are also shown.

Fig. 5: Gene sequences of the synthetic V heavy chain consensus genes. The corresponding amino acid sequences (see Fig. 2) as well as the unique cleavage sites are also shown.

Fig. 6: Oligonucleotides used for construction of the consensus genes. The oligos are named according to the corresponding consensus gene, e.g. the gene Vκ1 was constructed using the six oligonucleotides O1K1 to O1K6. The oligonucleotides used for synthesizing the genes encoding the constant domains Cκ (OCLK1 to 8) and CH1 (OCH1 to 8) are also shown.

Fig. 7A/B: Sequences of the synthetic genes encoding the constant domains Cκ (**A**) and CH1 (**B**). The corresponding amino acid sequences as well as unique cleavage sites introduced in these genes are also shown.

Fig. 7C: Functional map and sequence of module M24 comprising the synthetic Cλ gene segment (huCL lambda).

Fig. 7D: Oligonucleotides used for synthesis of module M24.

Fig. 8: Sequence and restriction map of the synthetic gene encoding the consensus single-chain fragment VH3-Vκ2. The signal sequence (amino acids 1 to 21) was derived from the *E. coli* phoA gene (Skerra & Plückthun, 1988). Between the phoA signal sequence and the VH3 domain, a short sequence stretch encoding 4 amino acid residues (amino acid 22 to 25) has been inserted in order to allow detection of the single-chain fragment in Western blot or ELISA using the monoclonal antibody M1 (Knappik & Plückthun, 1994). The last 6 basepairs of the sequence were introduced for cloning purposes (EcoRI site).

Fig. 9: Plasmid map of the vector pIG10.3 used for phage display of the H3κ2 scFv fragment. The vector is derived from pIG10 and contains the gene for the lac operon repressor, lacI, the artificial operon encoding the H3κ2-gene3ss fusion under control of the lac promoter, the Ipp terminator of transcription, the single-strand replication origin of the *E. coli* phage f1 (F1_ORI), a gene encoding β-lactamase (bla) and the ColEI derived origin of replication.

Fig.10: Sequencing results of independent clones from the initial library, translated into the corresponding amino acid sequences. (**A**) Amino acid sequence of the VH3 consensus heavy chain CDR3 (position 93 to 102, Kabat numbering). (**B**) Amino acid sequences of 12 clones of the 10-mer library. (**C**) Amino acid sequences of 11 clones of the 15-mer library, *: single base deletion.

Fig. 11: Expression test of individual library members. (**A**) Expression of 9 independent clones of the 10-mer library. (**B**) Expression of 9 independent clones of the 15-mer library. The lane designated with M contains the size marker. Both the gp3-scFv fusion and the scFv monomer are indicated.

Fig. 12: Enrichment of specific phage antibodies during the panning against FITC-BSA. The initial as well as the subsequent fluorescein-specific sub-libraries were panned against the blocking buffer and the ratio of the phage eluted from the FITC-BSA coated well vs. that from the powder milk coated well from each panning round is presented as the "specificity factor".

Fig. 13: Phage ELISA of 24 independent clones after the third round of panning tested for binding on FITC-BSA.

Fig. 14: Competition ELISA of selected FITC-BSA binding clones. The ELISA signals (OD_{405nm}) of scFv binding without inhibition are taken as 100%.

Fig. 15: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against FITC-BSA, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering).

Fig. 16: Coomassie-Blue stained SDS-PAGE of the purified anti-fluorescein scFv fragments: M: molecular weight marker, A: total soluble cell extract after induction, B: fraction of the flow-through, C, D and E: purified scFv fragments 1HA-3E4, 1HA-3E5 and 1HA-3E10, respectively.

Fig. 17: Enrichment of specific phage antibodies during the panning against β-estradiol-BSA, testosterone-BSA, BSA, ESL-1, interleukin-2, lymphotoxin-β, and LeY-BSA after three rounds of panning.

Fig. 18: ELISA of selected ESL-1 and β-estradiol binding clones

Fig. 19: Selectivity and cross-reactivity of HuCAL antibodies: in the diagonal specific binding of HuCAL antibodies can be seen, off-diagonal signals show non-specific cross-reactivity.

Fig. 20: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against β-estradiol-BSA, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering). One clone is derived from the 10mer library.

Fig. 21: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against testosterone-BSA, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering).

Fig. 22: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against lymphotoxin-β, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering). One clone comprises a 14mer CDR, presumably introduced by incomplete coupling of the trinucleotide mixture during oligonucleotide synthesis.

Fig. 23: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against ESL-1, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering). Two clones are derived from the 10mer library. One clone comprises a 16mer CDR, presumably introduced by chain elongation during oligonucleotide synthesis using trinucleotides.

Fig. 24: Sequencing results of the heavy chain CDR3s of independent clones after 3 rounds of panning against BSA, translated into the corresponding amino acid sequences (position 93 to 102, Kabat numbering).

Fig. 25: Schematic representation of the modular pCAL vector system.

Fig. 25a: List of restriction sites already used in or suitable for the modular HuCAL genes and pCAL vector system.

Fig. 26: List of the modular vector elements for the pCAL vector series: shown are only those restriction sites which are part of the modular system.

Fig. 27: Functional map and sequence of the multi-cloning site module (MCS)

Fig. 28: Functional map and sequence of the pMCS cloning vector series.

Fig. 29: Functional map and sequence of the pCAL module M1 (see Fig. 26).

Fig. 30: Functional map and sequence of the pCAL module M7-III (see Fig. 26).

Fig. 31: Functional map and sequence of the pCAL module M9-II (see Fig. 26).

Fig. 32: Functional map and sequence of the pCAL module M11-II (see Fig. 26).

Fig. 33: Functional map and sequence of the pCAL module M14-Ext2 (see Fig. 26).

Fig. 34: Functional map and sequence of the pCAL module M17 (see Fig. 26).

Fig. 35: Functional map and sequence of the modular vector pCAL4.

Fig. 35a: Functional maps and sequences of additional pCAL modules (M2, M3, M7I, M7II, M8, M10II, M11II, M12, M13, M19, M20, M21, M41) and of low-copy number plasmid vectors (pCALO1 to pCALO3).

Fig. 35b: List of oligonucleotides and primers used for synthesis of pCAL vector modules.

Fig. 36: Functional map and sequence of the β-lactamase cassette for replacement of CDRs for CDR library cloning.

Fig. 37: Oligo and primer design for Vκ CDR3 libraries

Fig. 38: Oligo and primer design for Vλ CDR3 libraries

Fig. 39: Functional map of the pBS13 expression vector series.

Fig. 40: Expression of all 49 HuCAL scFvs obtained by combining each of the 7 VH genes with each of the 7 VL genes (pBS13, 30°C): Values are given for the percentage of soluble vs. insoluble material, the total and the soluble amount compared to the combination H3κ2, which was set to 100%. In addition, the corresponding values for the McPC603 scFv are given.

Table 1: Summary of human immunoglobulin germline sequences used for computing the germline membership of rearranged sequences. (**A**) kappa sequences, (**B**) lambda sequences and (**C**), heavy chain sequences. (1) The germline name used in the various calculations, (2) the references number for the corresponding sequence (see appendix for sequence related citations), (3) the family where each sequence belongs to and (4), the various names found in literature for germline genes with identical amino acid sequences.

Table 2: Rearranged human sequences used for the calculation of consensus sequences. (**A**) kappa sequences, (**B**) lambda sequences and (**C**), heavy chain sequences. The table summarized the name of the sequence (1), the length of the sequence in amino acids (2), the germline family (3) as well as the computed germline counterpart (4). The number of amino acid exchanges between the rearranged sequence and the germline sequence is tabulated in (5), and the percentage of different amino acids is given in (6). Column (7) gives the references number for the corresponding sequence (see appendix for sequence related citations).

Table 3: Assignment of rearranged V sequences to their germline counterparts. (**A**) kappa sequences, (**B**) lambda sequences and (**C**), heavy chain sequences. The germline genes are tabulated according to their family (1), and the number of rearranged genes found for every germline gene is given in (2).

Table 4: Computation of the consensus sequence of the rearranged V kappa sequences. (**A**), V kappa subgroup 1,

(**B**), V kappa subgroup 2, (**C**), V kappa subgroup 3 and (**D**), V kappa subgroup 4. The number of each amino acid found at each position is tabulated together with the statistical analysis of the data. (1) Amino acids are given with their standard one-letter abbreviations (and B means D or N, Z means E or Q and X means any amino acid). The statistical analysis summarizes the number of sequences found at each position (2), the number of occurrences of the most common amino acid (3), the amino acid residue which is most common at this position (4), the relative frequency of the occurrence of the most common amino acid (5) and the number of different amino acids found at each position (6).

Table 5: Computation of the consensus sequence of the rearranged V lambda sequences. (**A**), V lambda subgroup 1, (**B**), V lambda subgroup 2, and (**C**), V lambda subgroup 3. The number of each amino acid found at each position is tabulated together with the statistical analysis of the data. Abbreviations are the same as in Table 4.

Table 6: Computation of the consensus sequence of the rearranged V heavy chain sequences. (**A**), V heavy chain subgroup 1A, (**B**), V heavy chain subgroup 1B, (**C**), V heavy chain subgroup 2, (**D**), V heavy chain subgroup 3, (**E**), V heavy chain subgroup 4, (**F**), V heavy chain subgroup 5, and (**G**), V heavy chain subgroup 6. The number of each amino acid found at each position is tabulated together with the statistical analysis of the data. Abbreviations are the same as in Table 4.

Examples

Example 1: Design of a Synthetic Human Combinatorial Antibody Library (HuCAL)

[0098] The following example describes the design of a fully synthetic human combinatorial antibody library (HuCAL), based on consensus sequences of the human immunoglobulin repertoire, and the synthesis of the consensus genes. The general procedure is outlined in Fig. 1.

1.1 Sequence *database*

1.1.1 Collection and alignment of human immunoglobulin sequences

[0099] In a first step, sequences of variable domains of human immunoglobulins have been collected and divided into three sub bases: V heavy chain (VH), V kappa (Vκ) and V lambda (Vλ). For each sequence, the gene sequence was then translated into the corresponding amino acid sequence. Subsequently, all amino acid sequences were aligned according to Kabat et al. (1991). In the case of Vλ sequences, the numbering system of Chuchana et al. (1990) was used. Each of the three main databases was then divided into two further sub bases: the first sub base contained all sequences derived from rearranged V genes, where more than 70 positions of the sequence were known. The second sub base contained all germline gene segments (without the D- and J- minigenes; pseudogenes with internal stop codons were also removed). In all cases, where germline sequences with identical amino acid sequence but different names were found, only one sequence was used (see Table 1). The final databases of rearranged sequences contained 386, 149 and 674 entries for Vκ, Vλ and VH, respectively. The final databases of germline sequences contained 48, 26 and 141 entries for Vκ, Vλ and VH, respectively.

1.1.2 Assignment of sequences to subgroups

[0100] The sequences in the three germline databases where then grouped according to sequence homology (see also Tomlinson et al., 1992, Williams & Winter, 1993, and Cox et al., 1994). In the case of Vκ, 7 families could be established. Vλ was divided into 8 families and VH into 6 families. The VH germline genes of the VH7 family (Van Dijk et al., 1993) were grouped into the VH1 family, since the genes of the two families are highly homologous. Each family contained different numbers of germline genes, varying from 1 (for example VH6) to 47 (VH3).

1.2 Analysis of sequences

1.2.1 Computation of germline membership

[0101] For each of the 1209 amino acid sequences in the databases of rearranged genes, the nearest germline counterpart, i.e. the germline sequence with the smallest number of amino acid differences was then calculated. After the germline counterpart was found, the number of somatic mutations which occurred in the rearranged gene and which led to amino acid exchanges could be tabulated. In 140 cases, the germline counterpart could not be calculated exactly, because more than one germline gene was found with an identical number of amino acid exchanges. These rearranged sequences were removed from the database. In a few cases, the number of amino acid exchanges was found to be

unusually large (>20 for VL and >25 for VH), indicating either heavily mutated rearranged genes or derivation from germline genes not present in the database. Since it was not possible to distinguish between these two possibilities, these sequences were also removed from the database. Finally, 12 rearranged sequences were removed from the database because they were found to have very unusual CDR lengths and composition or unusual amino acids at canonical positions (see below). In summary, 1023 rearranged sequences out of 1209 (85%) could be clearly assigned to their germline counterparts (see Table 2).

[0102] After this calculation, every rearranged gene could be arranged in one of the families established for the germline genes. Now the usage of each germline gene, i.e. the number of rearranged genes which originate from each germline gene, could be calculated (see Table 2). It was found that the usage was strongly biased towards a subset of germline genes, whereas most of the germline genes were not present as rearranged genes in the database and therefore apparently not used in the immune system (Table 3). This observation had already been reported in the case of Vκ (Cox, et al., 1994). All germline gene families, where no or only very few rearranged counterparts could be assigned, were removed from the database, leaving 4 Vκ, 3 Vλ, and 6 VH families.

1.2.2 Analysis of CDR conformations

[0103] The conformation of the antigen binding loops of antibody molecules, the CDRs, is strongly dependent on both the length of the CDRs and the amino acid residues located at the so-called canonical positions (Chothia & Lesk, 1987). It has been found that only a few canonical structures exist, which determine the structural repertoire of the immunoglobulin variable domains (Chothia et al., 1989). The canonical amino acid positions can be found in CDR as well as framework regions. The 13 used germline families defined above (7 VL and 6 VH) were now analyzed for their canonical structures in order to define the structural repertoire encoded in these families.

[0104] In 3 of the 4 Vκ families (Vκ1, 2 and 4), one different type of CDR1 conformation could be defined for every family. The family Vκ3 showed two types of CDR1 conformation: one type which was identical to Vκ1 and one type only found in Vκ3. All Vκ CDR2s used the same type of canonical structure. The CDR3 conformation is not encoded in the germline gene segments. Therefore, the 4 Vκ families defined by sequence homology and usage corresponded also to 4 types of canonical structures found in Vκ germline genes.

[0105] The 3 Vλ families defined above showed 3 types of CDR1 conformation, each family with one unique type. The Vλ1 family contained 2 different CDR1 lengths (13 and 14 amino acids), but identical canonical residues, and it is thought that both lengths adopt the same canonical conformation (Chothia & Lesk, 1987). In the CDR2 of the used Vλ germlines, only one canonical conformation exists, and the CDR3 conformation is not encoded in the germline gene segments. Therefore, the 3 Vλ families defined by sequence homology and usage corresponded also to 3 types of canonical structures.

[0106] The structural repertoire of the human VH sequences was analyzed in detail by Chothia et al., 1992. In total, 3 conformations of CDR1 (H1-1, H1-2 and H1-3) and 6 conformations of CDR2 (H2-1, H2-2, H2-3, H2-4, H2-5 and H2-x) could be defined. Since the CDR3 is encoded in the D- and J-minigene segments, no particular canonical residues are defined for this CDR.

All the members of the VH1 family defined above contained the CDR1 conformation H1-1, but differed in their CDR2 conformation: the H2-2 conformation was found in 6 germline genes, whereas the conformation H2-3 was found in 8 germline genes. Since the two types of CDR2 conformations are defined by different types of amino acid at the framework position 72, the VH1 family was divided into two subfamilies: VH1A with CDR2 conformation H2-2 and VH1B with the conformation H2-3. The members of the VH2 family all had the conformations H1-3 and H2-1 in CDR1 and CDR2, respectively. The CDR1 conformation of the VH3 members was found in all cases to be H1-1, but 4 different types were found in CDR2 (H2-1, H2-3, H2-4 and H2-x). In these CDR2 conformations, the canonical framework residue 71 is always defined by an arginine. Therefore, it was not necessary to divide the VH3 family into subfamilies, since the 4 types of CDR2 conformations were defined solely by the CDR2 itself. The same was true for the VH4 family. Here, all 3 types of CDR1 conformations were found, but since the CDR1 conformation was defined by the CDR itself (the canonical framework residue 26 was found to be glycine in all cases), no subdivisions were necessary. The CDR2 conformation of the VH4 members was found to be H2-1 in all cases. All members of the VH5 family were found to have the conformation H1-1 and H2-2, respectively. The single germline gene of the VH6 family had the conformations H1-3 and H2-5 in CDR1 and CDR2, respectively.

[0107] In summary, all possible CDR conformations of the Vκ and Vλ genes were present in the 7 families defined by sequence comparison. From the 12 different CDR conformations found in the used VH germline genes, 7 could be covered by dividing the family VH1 into two subfamilies, thereby creating 7 VH families. The remaining 5 CDR conformations (3 in the VH3 and 2 in the VH4 family) were defined by the CDRs themselves and could be created during the construction of CDR libraries. Therefore, the structural repertoire of the used human V genes could be covered by 49 (7 x 7) different frameworks.

1.2.3 Computation of consensus sequences

[0108] The 14 databases of rearranged sequences (4 Vκ, 3 Vλ and 7 VH) were used to compute the HuCAL consensus sequences of each subgroup (4 HuCAL- Vκ, 3 HuCAL- Vλ, 7 HuCAL- VH, see Table 4, 5 and 6). This was done by counting the number of amino acid residues used at each position (position variability) and subsequently identifying the amino acid residue most frequently used at each position. By using the rearranged sequences instead of the used germline sequences for the calculation of the consensus, the consensus was weighted according to the frequency of usage. Additionally, frequently mutated and highly conserved positions could be identified. The consensus sequences were cross-checked with the consensus of the germline families to see whether the rearranged sequences were biased at certain positions towards amino acid residues which do not occur in the collected germline sequences, but this was found not to be the case. Subsequently, the number of differences of each of the 14 consensus sequences to each of the germline sequences found in each specific family was calculated. The overall deviation from the most homologous germline sequence was found to be 2.4 amino acid residues (s.d. = 2.7), ensuring that the "artificial" consensus sequences can still be considered as truly human sequences as far as immunogenicity is concerned.

1.3 Structural analysis

[0109] So far, only sequence information was used to design the consensus sequences. Since it was possible that during the calculation certain artificial combinations of amino acid residues have been created, which are located far away in the sequence but have contacts to each other in the three dimensional structure, leading to destabilized or even misfolded frameworks, the 14 consensus sequences were analyzed according to their structural properties.
It was rationalized that all rearranged sequences present in the database correspond to functional and therefore correctly folded antibody molecules. Hence, the most homologous rearranged sequence was calculated for each consensus sequence. The positions where the consensus differed from the rearranged sequence were identified as potential "artificial residues" and inspected.
The inspection itself was done in two directions. First, the local sequence stretch around each potentially "artificial residue" was compared with the corresponding stretch of all the rearranged sequences. If this stretch was found to be truly artificial, i.e. never occurred in any of the rearranged sequences, the critical residue was converted into the second most common amino acid found at this position and analyzed again. Second, the potentially "artificial residues" were analyzed for their long range interactions. This was done by collecting all available structures of human antibody variable domains from the corresponding PDB files and calculating for every structure the number and type of interactions each amino acid residue established to each side-chain. These "interaction maps" were used to analyze the probable side-chain/side-chain interactions of the potentially "artificial residues". As a result of this analysis, the following residues were exchanged (given is the name of the gene, the position according to Kabat's numbering scheme, the amino acid found at this position as the most abundant one and the amino acid which was used instead):

VH2: $S_{65}T$
Vκ1: $N_{34}A$,
Vκ3: G_9A, $D_{60}A$, $R_{77}S$
Vλ3: $V_{78}T$

1.4 Design of CDR sequences

[0110] The process described above provided the complete consensus sequences derived solely from the databases of rearranged sequences. It was rationalized that the CDR1 and CDR2 regions should be taken from the databases of used germline sequences, since the CDRs of rearranged and mutated sequences are biased towards their particular antigens. Moreover, the germline CDR sequences are known to allow binding to a variety of antigens in the primary immune response, where only CDR3 is varied. Therefore, the consensus CDRs obtained from the calculations described above were replaced by germline CDRs in the case of VH and Vκ. In the case of Vλ, a few amino acid exchanges were introduced in some of the chosen germline CDRs in order to avoid possible protease cleavage sites as well as possible structural constraints.

[0111] The CDRs of following germline genes have been chosen:

HuCAL gene	CDR1	CDR2
HuCAL-VH1A	VH1-12-1	VH1-12-1
HuCAL-VH1B	VH1-13-16	VH1-13-6,-7,-8,-9
HuCAL-VH2	VH2-31-10,-11,-12,-13	VH2-31-3,-4

(continued)

HuCAL gene	CDR1	CDR2
HuCAL-VH3	VH3-13-8,-9,-10	VH3-13-8,-9,-10
HuCAL-VH4	VH4-11-7 to -14	VH4-11-8,-9,-11,-12,-14,-16
		VH4-31-17,-18,-19,-20
HuCAL-VH5	VH5-12-1,-2	VH5-12-1,-2
HuCAL-VH6	VH6-35-1	VH6-35-1
HuCAL-Vκ1	Vκ1-14,-15	Vκ1-2,-3,-4,-5,-7,-8,-12,-13,-18,-19
HuCAL-Vκ2	Vκ2-6	Vκ2-6
HuCAL-Vκ3	Vκ3-1,-4	Vκ3-4
HuCAL-Vκ4	Vκ4-1	Vκ4-1
HuCAL-Vλ1	HUMLV117,DPL5	DPL5
HuCAL-Vλ2	DPL11,DPL12	DPL12
HuCAL-Vλ3	DPL23	HUMLV318

[0112] In the case of the CDR3s, any sequence could be chosen since these CDRs were planned to be the first to be replaced by oligonucleotide libraries. In order to study the expression and folding behavior of the consensus sequences in *E. coli,* it would be useful to have all sequences with the same CDR3, since the influence of the CDR3s on the folding behavior would then be identical in all cases. The dummy sequences QQHYTTPP and ARWGGDGFYAMDY were selected for the VL chains (kappa and lambda) and for the VH chains, respectively. These sequences are known to be compatible with antibody folding in *E. coli* (Carter et al., 1992).

1.5 Gene design

[0113] The final outcome of the process described above was a collection of 14 HuCAL amino acid sequences, which represent the frequently used structural antibody repertoire of the human immune system (see Figure 2). These sequences were back-translated into DNA sequences. In a first step, the back-translation was done using only codons which are known to be frequently used in *E. coli.* These gene sequences were then used for creating a database of all possible restriction endonuclease sites, which could be introduced without changing the corresponding amino acid sequences. Using this database, cleavage sites were selected which were located at the flanking regions of all sub-elements of the genes (CDRs and framework regions) and which could be introduced in all HuCAL VH, Vκ or Vλ genes simultaneously at the same position. In a few cases it was not possible to find cleavage sites for all genes of a subgroup. When this happened, the amino acid sequence was changed, if this was possible according to the available sequence and structural information. This exchange was then analyzed again as described above. In total, the following 6 amino acid residues were exchanged during this design (given is the name of the gene, the position according to Kabat's numbering scheme, the amino acid found at this position as the most abundant one and the amino acid which was used instead):

VH2: T_3Q
VH6: $S_{42}G$
Vκ3: E_1D, $I_{58}V$
Vκ4: $K_{24}R$
Vλ3: $T_{22}S$

[0114] In one case (5'-end of VH framework 3) it was not possible to identify a single cleavage site for all 7 VH genes. Two different type of cleavage sites were used instead: BstEII for HuCAL VH1A, VH1B, VH4 and VH5, and NspV for HuCAL VH2, VH3, VH4 and VH6.

[0115] Several restriction endonuclease sites were identified, which were not located at the flanking regions of the sub-elements but which could be introduced in every gene of a given group without changing the amino acid sequence. These cleavage sites were also introduced in order to make the system more flexible for further improvements. Finally, all but one remaining restriction endonuclease sites were removed in every gene sequence. The single cleavage site, which was not removed was different in all genes of a subgroup and could be therefore used as a "fingerprint" site to ease the identification of the different genes by restriction digest. The designed genes, together with the corresponding amino acid sequences and the group-specific restriction endonuclease sites are shown in Figure 3, 4 and 5, respectively.

1.6 Gene synthesis and cloning

[0116] The consensus genes were synthesized using the method described by Prodromou & Pearl, 1992, using the oligonucleotides shown in Fig. 6. Gene segments encoding the human constant domains Cκ, Cλ and CH1 were also synthesized, based on sequence information given by Kabat et al., 1991 (see Fig. 6 and Fig. 7). Since for both the CDR3 and the framework 4 gene segments identical sequences were chosen in all HuCAL Vκ, Vλ and VH genes, respectively, this part was constructed only once, together with the corresponding gene segments encoding the constant domains. The PCR products were cloned into pCR-Script KS(+) (Stratagene, Inc.) or pZErO-1 (Invitrogen, Inc.) and verified by sequencing.

Example 2: Cloning and Testing of a HuCAL-Based Antibody Library

[0117] A combination of two of the synthetic consensus genes was chosen after construction to test whether binding antibody fragments can be isolated from a library based on these two consensus frameworks. The two genes were cloned as a single-chain Fv (scFv) fragment, and a VH-CDR3 library was inserted. In order to test the library for the presence of functional antibody molecules, a selection procedure was carried out using the small hapten fluorescein bound to BSA (FITC-BSA) as antigen.

2.1 Cloning of the HuCAL VH3-Vκ2 scFv fragment.

[0118] In order to test the design of the consensus genes, one randomly chosen combination of synthetic light and heavy gene (HuCAL-Vκ2 and HuCAL-VH3) was used for the construction of a single-chain antibody (scFv) fragment. Briefly, the gene segments encoding the VH3 consensus gene and the CH1 gene segment including the CDR3 - framework 4 region, as well as the Vκ2 consensus gene and the Cκ gene segment including the CDR3 - framework 4 region were assembled yielding the gene for the VH3-CH1 Fd fragment and the gene encoding the Vκ2-Cκ light chain, respectively. The CH1 gene segment was then replaced by an oligonucleotide cassette encoding a 20-mer peptide linker with the sequence AGGGSGGGGSGGGGSGGGGS. The two oligonucleotides encoding this linker were 5'- TCAGCGGGT-GGCGGTTCTGGCGGCG GTGGGAGCGGTGGCGGTGGTTCTGGCGGTGGTGGTTCCGATATCGGTCCACG TACGG-3' and 5'-AATTCCGTACGTGGACCGATATCGGAACCACCACCGCCAGAACCACCGCCACCGCTCC CAC-CGCCGCCAGAACCGCCACCCGC-3', respectively. Finally, the HuCAL-Vκ2 gene was inserted via EcoRV and BsiWI into the plasmid encoding the HuCAL-VH3-linker fusion, leading to the final gene HuCAL-VH3-Vκ2, which encoded the two consensus sequences in the single-chain format VH-linker-VL. The complete coding sequence is shown in Fig. 8.

2.2 Construction of a monovalent phage-display phagemid vector pIG10.3

[0119] Phagemid pIG10.3 (Fig. 9) was constructed in order to create a phage-display system (Winter et al., 1994) for the H3κ2 scFv gene. Briefly, the EcoRI/HindIII restriction fragment in the phagemid vector pIG10 (Ge et al., 1995) was replaced by the c-myc followed by an amber codon (which encodes an glutamate in the amber-suppresser strain XL1 Blue and a stop codon in the non-suppresser strain JM83) and a truncated version of the gene III (fusion junction at codon 249, see Lowman et al., 1991) through PCR mutagenesis.

2.3 Construction of H-CDR3 libraries

[0120] Heavy chain CDR3 libraries of two lengths (10 and 15 amino acids) were constructed using trinucleotide codon containing oligonucleotides (Virnekäs et al., 1994) as templates and the oligonucleotides complementing the flanking regions as primers. To concentrate only on the CDR3 structures that appear most often in functional antibodies, we kept the salt-bridge of R_{H94} and D_{H101} in the CDR3 loop. For the 15-mer library, both phenylalanine and methionine were introduced at position 100 since these two residues were found to occur quite often in human CDR3s of this length (not shown). For the same reason, valine and tyrosine were introduced at position 102. All other randomized positions contained codons for all amino acids except cystein, which was not used in the trinucleotide mixture.
The CDR3 libraries of lengths 10 and 15 were generated from the PCR fragments using oligonucleotide templates O3HCDR103T (5'-GATACGGCCGTGTATTATTGCGCGCGT $(TRI)_6$G-ATTATTGGGGCCAAGGCACCCTG-3') and O3HCDR153T (5'-GATACGGCCGT GTATTATTG-CGCGCGT$(TRI)_{10}$(TTTIATG)GAT(GTTITAT)TGGGGCCAAG-GCACCCTG-3'), and primers O3HCDR35 (5'-GATACGGCCGTGTATTATTGC-3') and O3HCDR33 (5'-CAGGGT-GCCTTGGCCCC-3'), where TRI are trinucleotide mixtures representing all amino acids without cystein, (TTT/ATG) and (GTT/TAT) are trinucleotide mixtures encoding the amino acids phenylalanine/methionine and valine/tyrosine, respectively. The potential diversity of these libraries was 4.7×10^7 and 3.4×10^{10} for 10-mer and 15-mer library, respectively. The library cassettes were first synthesized from PCR amplification of the oligo templates in the presence of both primers:

25 pmol of the oligo template 03HCDR103T or 03HCDR153T, 50 pmol each of the primers 03HCDR35 and 03HCDR33, 20 nmol of dNTP, 10x buffer and 2.5 units of Pfu DNA polymerase (Stratagene) in a total volume of 100 μl for 30 cycles (1 minute at 92°C, 1 minute at 62°C and 1 minute at 72°C). A hot-start procedure was used. The resulting mixtures were phenol-extracted, ethanol-precipitated and digested overnight with EagI and StyI. The vector pIG10.3-scH3κ2cat, where the EagI-StyI fragment in the vector pIG10.3-scH3κ2 encoding the H-CDR3 was replaced by the chloramphenicol acetyl-transferase gene (cat) flanked with these two sites, was similarly digested. The digested vector (35 μg) was gel-purified and ligated with 100 μg of the library cassette overnight at 16°C. The ligation mixtures were isopropanol precipitated, air-dried and the pellets were redissolved in 100 μl of ddH2O. The ligation was mixed with 1 ml of freshly prepared electrocompetent XL1 Blue on ice. 20 rounds of electroporation were performed and the transformants were diluted in SOC medium, shaken at 37°C for 30 minutes and plated out on large LB plates (Amp/Tet/Glucose) at 37°C for 6-9 hrs. The number of transformants (library size) was $3.2x10^7$ and $2.3x10^7$ for the 10-mer and the 15-mer library, respectively. The colonies were suspended in 2xYT medium (Amp/TetIGlucose) and stored as glycerol culture.

In order to test the quality of the initial library, phagemids from 24 independent colonies (12 from the 10-mer and 12 from the 15-mer library, respectively) were isolated and analyzed by restriction digestion and sequencing. The restriction analysis of the 24 phagemids indicated the presence of intact vector in all cases. Sequence analysis of these clones (see Fig. 10) indicated that 22 out of 24 contained a functional sequence in their heavy chain CDR3 regions. 1 out of 12 clones of the 10-mer library had a CDR3 of length 9 instead of 10, and 2 out of 12 clones of the 15-mer library had no open reading frame, thereby leading to a non-functional scFv; one of these two clones contained two consecutive inserts, but out of frame (data not shown). All codons introduced were presented in an even distribution.

Expression levels of individual library members were also measured. Briefly, 9 clones from each library were grown in 2xYT medium containing Amp/Tet/0.5% glucose at 37°C overnight. Next day, the cultures were diluted into fresh medium with Amp/Tet. At an OD_{600nm} of 0.4, the cultures were induced with 1 mM of IPTG and shaken at RT overnight. Then the cell pellets were suspended in 1 ml of PBS buffer + 1 mM of EDTA. The suspensions were sonicated and the supernatants were separated on an SDS-PAGE under reducing conditions, blotted on nylon membrane and detected with anti-FLAG M1 antibody (see Fig. 11). From the nine clones of the 10-mer library, all express the scFv fragments. Moreover, the gene III / scFv fusion proteins were present in all cases. Among the nine clones from the 15-mer library analyzed, 6/9 (67%) led to the expression of both scFv and the gene III/scFv fusion proteins. More importantly, all clones expressing the scFvs and gene III/scFv fusions gave rise to about the same level of expression.

2.4 Biopanning

[0121] Phages displaying the antibody libraries were prepared using standard protocols. Phages derived from the 10-mer library were mixed with phages from the 15-mer library in a ratio of 20:1 ($1x10^{10}$ cfu/well of the 10-mer and $5x10^8$ cfu/well of the 15-mer phages, respectively). Subsequently, the phage solution was used for panning in ELISA plates (Maxisorp, Nunc) coated with FITC-BSA (Sigma) at concentration of 100 μg/ml in PBS at 4°C overnight. The antigen-coated wells were blocked with 3% powder milk in PBS and the phage solutions in 1% powder milk were added to each well and the plate was shaken at RT for 1 hr. The wells were then washed with PBST and PBS (4 times each with shaking at RT for 5 minutes). The bound phages were eluted with 0.1 M triethylamine (TEA) at RT for 10 minutes. The eluted phage solutions were immediately neutralized with 1/2 the volume of 1 M Tris·Cl, pH 7.6. Eluted phage solutions (ca. 450 μl) were used to infect 5 ml of XL1 Blue cells at 37°C for 30 min. The infected cultures were then plated out on large LB plates (Amp/Tet/Glucose) and allowed to grow at 37°C until the colonies were visible. The colonies were suspended in 2xYT medium and the glycerol cultures were made as above described. This panning round was repeated twice, and in the third round elution was carried out with addition of fluorescein in a concentration of 100 μg/ml in PBS. The enrichment of specific phage antibodies was monitored by panning the initial as well as the subsequent fluorescein-specific sub-libraries against the blocking buffer (Fig. 12). Antibodies with specificity against fluorescein were isolated after 3 rounds of panning.

2.5 ELISA measurements

[0122] One of the criteria for the successful biopanning is the isolation of individual phage clones that bind to the targeted antigen or hapten. We undertook the isolation of anti-FITC phage antibody clones and characterized them first in a phage ELISA format. After the 3rd round of biopanning (see above), 24 phagemid containing clones were used to inoculate 100 μl of 2xYT medium (Amp/Tet/Glucose) in an ELISA plate (Nunc), which was subsequently shaken at 37°C for 5 hrs. 100 μl of 2xYT medium (Amp/Tet/1 mM IPTG) were added and shaking was continued for 30 minutes. A further 100 μl of 2xYT medium (Amp/Tet) containing the helper phage ($1 x 10^9$ cfu/well) was added and shaking was done at RT for 3 hrs. After addition of kanamycin to select for successful helper phage infection, the shaking was continued overnight. The plates were then centrifuged and the supernatants were pipetted directly into ELISA wells coated with 100 μl FITC-BSA (100μg/ml) and blocked with milk powder. Washing was performed similarly as during the panning

procedure and the bound phages were detected with anti-M13 antibody-POD conjugate (Pharmacia) using soluble POD substrate (Boehringer-Mannheim). Of the 24 clones screened against FITC-BSA, 22 were active in the ELISA (Fig. 13). The initial libraries of similar titer gave rise to no detectable signal.

Specificity for fluorescein was measured in a competitive ELISA. Periplasmic fractions of five FITC specific scFvs were prepared as described above. Western blotting indicated that all clones expressed about the same amount of scFv fragment (data not shown). ELISA was performed as described above, but additionally, the periplasmic fractions were incubated 30 min at RT either with buffer (no inhibition), with 10 mg/ml BSA (inhibition with BSA) or with 10 mg/ml fluorescein (inhibition with fluorescein) before adding to the well. Binding scFv fragment was detected using the anti-FLAG antibody M1. The ELISA signal could only be inhibited, when soluble fluorescein was added, indicating binding of the scFvs was specific for fluorescein (Fig. 14).

2.6 Sequence analysis

[0123] The heavy chain CDR3 region of 20 clones were sequenced in order to estimate the sequence diversity of fluorescein binding antibodies in the library (Fig. 15). In total, 16 of 20 sequences (80%) were different, showing that the constructed library contained a highly diverse repertoire of fluorescein binders. The CDR3s showed no particular sequence homology, but contained on average 4 arginine residues. This bias towards arginine in fluorescein binding antibodies had already been described by Barbas et al., 1992.

2.7 Production

[0124] *E. coli* JM83 was transformed with phagemid DNA of 3 selected clones and cultured in 0.5 L 2xYT medium. Induction was carried out with 1 mM IPTG at OD_{600nm} = 0.4 and growth was continued with vigorous shaking at RT overnight. The cells were harvested and pellets were suspended in PBS buffer and sonicated. The supernatants were separated from the cell debris via centrifugation and purified via the BioLogic system (Bio-Rad) by with a POROS®MC 20 column (IMAC, PerSeptive Biosystems, Inc.) coupled with an ion-exchange chromatography column. The ion-exchange column was one of the POROS®HS, CM or HQ or PI 20 (PerSeptive Biosystems, Inc.) depended on the theoretical pI of the scFv being purified. The pH of all the buffers was adjusted to one unit lower or higher than the pI of the scFv being purified throughout. The sample was loaded onto the first IMAC column, washed with 7 column volumes of 20 mM sodium phosphate, 1 M NaCl and 10 mM imidazole. This washing was followed by 7 column volumes of 20 mM sodium phosphate and 10 mM imidazole. Then 3 column volumes of an imidazole gradient (10 to 250 mM) were applied and the eluent was connected directly to the ion-exchanger. Nine column volumes of isocratic washing with 250 mM imidazole was followed by 15 column volumes of 250 mM to 100 mM and 7 column volumes of an imidazole / NaCl gradient (100 to 10 mM imidazole, 0 to 1 M NaCl). The flow rate was 5 ml/min. The purity of scFv fragments was checked by SDS-PAGE Coomassie staining (Fig. 16). The concentration of the fragments was determined from the absorbance at 280 nm using the theoretically determined extinction coefficient (Gill & von Hippel, 1989). The scFv fragments could be purified to homogeneity (see Fig. 16). The yield of purified fragments ranged from 5 to 10 mg/L/OD.

Example 3: HuCAL H3k2 Library Against a Collection of Antigens

[0125] In order to test the library used in Example 2 further, a new selection procedure was carried out using a variety of antigens comprising β-estradiol, testosterone, Lewis-Y epitope (LeY), interleukin-2 (IL-2), lymphotoxin-β (LT-β), E-selectin ligand-1 (ESL-1), and BSA.

3.1 Biopanning

[0126] The library and all procedures were identical to those described in Example 2. The ELISA plates were coated with β-estradiol-BSA (100 μg/ml), testosterone-BSA (100 μg/ml), LeY-BSA (20 μg/ml) IL-2 (20 μg/ml), ESL-1 (20 μg/ml) and BSA (100 μg/ml), LT-β (denatured protein, 20 μg/ml). In the first two rounds, bound phages were eluted with 0.1 M triethylamine (TEA) at RT for 10 minutes. In the case of BSA, elution after three rounds of panning was carried out with addition of BSA in a concentration of 100 μg/ml in PBS. In the case of the other antigens, third round elution was done with 0.1 M triethylamine. In all cases except LeY, enrichment of binding phages could be seen (Figure 17). Moreover, a repetition of the biopanning experiment using only the 15-mer library resulted in the enrichment of LeY-binding phages as well (data not shown).

3.2. ELISA measurements

[0127] Clones binding to β-estradiol, testosterone, LeY, LT-β, ESL-1 and BSA were further analyzed and characterized

as described in Example 2 for FITC. ELISA data for anti-β-estradiol and anti-ESL-1 antibodies are shown in Fig. 18. In one experiment, selectivity and cross-reactivity of binding scFv fragments were tested. For this purpose, an ELISA plate was coated with FITC, testosterone, β-estradiol, BSA, and ESL-1, with 5 wells for each antigen arranged in 5 rows, and 5 antibodies, one against each of the antigens, were screened against each of the antigens. Fig. 19 shows the specific binding of the antibodies to the antigen it was selected for, and the low cross-reactivity with the other four antigens.

3.3 Sequence analysis

[0128] The sequencing data of several clones against β-estradiol (34 clones), testosterone (12 clones), LT-β (23 clones), ESL-1 (34 clones), and BSA (10 clones) are given in Figures 20 to 24.

Example 4: Vector Construction

[0129] To be able to take advantage of the modularity of the consensus gene repertoire, a vector system had to be constructed which could be used in phage display screening of HuCAL libraries and subsequent optimization procedures. Therefore, all necessary vector elements such as origins of single-stranded or double-stranded replication, promotor/operator, repressor or terminator elements, resistance genes, potential recombination sites, gene III for display on filamentous phages, signal sequences, or detection tags had to be made compatible with the restriction site pattern of the modular consensus genes. Figure 25 shows a schematic representation of the pCAL vector system and the arrangement of vector modules and restriction sites therein. Figure 25a shows a list of all restriction sites which are already incorporated into the consensus genes or the vector elements as part of the modular system or which are not yet present in the whole system. The latter could be used in a later stage for the introduction of or within new modules.

4.1 Vector modules

[0130] A series of vector modules was constructed where the restriction sites flanking the gene sub-elements of the HuCAL genes were removed, the vector modules themselves being flanked by unique restriction sites. These modules were constructed either by gene synthesis or by mutagenesis of templates. Mutagenesis was done by add-on PCR, by site-directed mutagenesis (Kunkel et al., 1991) or multisite oligonucleotide-mediated mutagenesis (Sutherland et al., 1995; Perlak, 1990) using a PCR-based assembly method.
Figure 26 contains a list of the modules constructed. Instead of the terminator module M9 (HindIII-Ipp-PacI), a larger cassette M9II was prepared to introduce FseI as additional restriction site. M9II can be cloned via HindIII/BsrGI.
All vector modules were characterized by restriction analysis and sequencing. In the case of module M11-II, sequencing of the module revealed a two-base difference in positions 164/65 compared to the sequence database of the template. These two different bases (CA → GC) created an additional BanII site. Since the same two-base difference occurs in the f1 origin of other bacteriophages, it can be assumed that the two-base difference was present in the template and not created by mutagenesis during cloning. This BanII site was removed by site-directed mutagenesis, leading to module M11-III. The BssSI site of module M14 could initially not be removed without impact on the function of the ColE1 origin, therefore M14-Ext2 was used for cloning of the first pCAL vector series. Figures 29 to 34 are showing the functional maps and sequences of the modules used for assembly of the modular vector pCAL4 (see below). The functional maps and sequences of additional modules can be found in Figure 35a. Figure 35b contains a list of oligonucleotides and primers used for the synthesis of the modules.

4.2 Cloning vector pMCS

[0131] To be able to assemble the individual vector modules, a cloning vector pMCS containing a specific multi-cloning site (MCS) was constructed. First, an MCS cassette (Fig. 27) was made by gene synthesis. This cassette contains all those restriction sites in the order necessary for the sequential introduction of all vector modules and can be cloned via the 5'-HindIII site and a four base overhang at the 3'-end compatible with an AatII site. The vector pMCS (Figure 28) was constructed by digesting pUC19 with AatII and HindIII, isolating the 2174 base pair fragment containing the bla gene and the ColE1 origin, and ligating the MCS cassette.

4.3 Cloning of modular vector pCAL4

[0132] This was cloned step by step by restriction digest of pMCS and subsequent ligation of the modules M1 (via AatII/XbaI), M7III (via EcoRI/HindIII), and M9II (via HindIII/BsrGI), and M11-II (via BsrGI/NheI). Finally, the bla gene was replaced by the cat gene module M17 (via AatII/BglII), and the wild type ColE1 origin by module M14-Ext2 (via BglII/NheI). Figure 35 is showing the functional map and the sequence of pCAL4.

4.4 Cloning of low-copy number plasmid vectors pCALO

[0133] A series of low-copy number plasmid vectors was constructed in a similar way using the p15A module M12 instead of the ColE1 module M14-Ext2. Figure 35a is showing the functional maps and sequences of the vectors pCALO1 to pCALO3.

Example 5: Construction of a HuCAL scFv Library

5.1 Cloning of all 49 HuCAL scFv fragments

[0134] All 49 combinations of the 7 HuCAL-VH and 7 HuCAL-VL consensus genes were assembled as described for the HuCAL VH3-Vκ2 scFv in Example 2 and inserted into the vector pBS12, a modified version of the pLisc series of antibody expression vectors (Skerra et *al.,* 1991).

5.2 Construction of a CDR cloning cassette

[0135] For replacement of CDRs, a universal β-lactamase cloning cassette was constructed having a multi-cloning site at the 5'-end as well as at the 3'-end. The 5'-multi-cloning site comprises all restriction sites adjacent to the 5'-end of the HuCAL VH and VL CDRs, the 3'-multi-cloning site comprises all restriction sites adjacent to the 3' end of the HuCAL VH and VL CDRs. Both 5'- and 3'-multi-cloning site were prepared as cassettes via add-on PCR using synthetic oligonucleotides as 5'- and 3'-primers using wild type β-lactamase gene as template. Figure 36 shows the functional map and the sequence of the cassette bla-MCS.

5.3. Preparation of VL-CDR3 library cassettes

[0136] The VL-CDR3 libraries comprising 7 random positions were generated from the PCR fragments using oligonucleotide templates Vx1&Vκ3, Vκ2 and Vκ4 and primers O_K3L_5 and O_K3L_3 (Fig. 37) for the Vκ genes, and Vλ and primers O_L3L_5 (5'-GCAGAAGGCGAACGTCC-3') and O_L3LA_3 (Fig. 38) for the Vλ genes. Construction of the cassettes was performed as described in Example 2.3.

5.4 Cloning of HuCAL scFv genes with VL-CDR3 libraries

[0137] Each of the 49 single-chains was subcloned into pCAL4 via XbaI/EcoRI and the VL-CDR3 replaced by the β-lactamase cloning cassette via BbsI/MscI, which was then replaced by the corresponding VL-CDR3 library cassette synthesized as described above. This CDR replacement is described in detail in Example 2.3 where the cat gene was used.

5.5 Preparation of VH-CDR3 library cassette

[0138] The VH-CDR3 libraries were designed and synthesized as described in Example 2.3.

5.6 Cloning of HuCAL scFv genes with VL- and VH-CDR3 libraries

[0139] Each of the 49 single-chain VL-CDR3 libraries was digested with BssHII/StyI to replace VH-CDR3. The "dummy" cassette digested with BssHII/StyI was inserted, and was then replaced by a corresponding VH-CDR3 library cassette synthesized as described above.

Example 6: Expression tests

[0140] Expression and toxicity studies were performed using the scFv format VH-linker-VL. All 49 combinations of the 7 HuCAL-VH and 7 HuCAL-VL consensus genes assembled as described in Example 5 were inserted into the vector pBS13, a modified version of the pLisc series of antibody expression vectors (Skerra et *al.,* 1991). A map of this vector is shown in Fig. 39.

E. coli JM83 was transformed 49 times with each of the vectors and stored as glycerol stock. Between 4 and 6 clones were tested simultaneously, always including the clone H3κ2, which was used as internal control throughout. As additional control, the McPC603 scFv fragment (Knappik & Plückthun, 1995) in pBS13 was expressed under identical conditions. Two days before the expression test was performed, the clones were cultivated on LB plates containing 30 μg/ml chloramphenicol and 60 mM glucose. Using this plates an 3 ml culture (LB medium containing 90 μg chloramphenicol

and 60 mM glucose) was inoculated overnight at 37 °C. Next day the overnight culture was used to inoculate 30 ml LB medium containing chloramphenicol (30 μg/ml). The starting OD_{600nm} was adjusted to 0.2 and a growth temperature of 30 °C was used. The physiology of the cells was monitored by measuring every 30 minutes for 8 to 9 hours the optical density at 600 nm. After the culture reached an OD_{600nm} of 0.5, antibody expression was induced by adding IPTG to a final concentration of 1 mM. A 5 ml aliquot of the culture was removed after 2 h of induction in order to analyze the antibody expression. The cells were lysed and the soluble and insoluble fractions of the crude extract were separated as described in Knappik & Plückthun, 1995. The fractions were assayed by reducing SDS-PAGE with the samples normalized to identical optical densities. After blotting and immunostaining using the α-FLAG antibody M1 as the first antibody (see Ge et *al.,* 1994) and an Fc-specific anti-mouse antiserum conjugated to alkaline phosphatase as the second antibody, the lanes were scanned and the intensities of the bands of the expected size (appr. 30 kDa) were quantified densitometrically and tabulated relative to the control antibody (see Fig. 40).

Example 7: Optimization of Fluorescein Binders

7.1. Construction of L-CDR3 and H-CDR2 library cassettes

[0141] A L-CDR3 library cassette was prepared from the oligonucleotide template CDR3L (5'-TGGAAGCTGAA-GACGTGGGCGTGTATTATTGCCAGCAG(TR5)$(TRI)_4$CCG(TRI)TT TGGCCAGGGTACGAAAGTT-3') and primer 5'-AACTTTCGTACCCTGGCC-3' for synthesis of the complementary strand, where (TRI) was a trinucleotide mixture representing all amino acids except Cys, (TR5) comprised a trinucleotide mixture representing the 5 codons for Ala, Arg, His, Ser, and Tyr.

A H-CDR2 library cassette was prepared from the oligonucleotide template CDRsH (5'-AGGGTCTCGAGTGGGTGAGC (TRI)ATT$(TRI)_2$.

$_3$ $(6)_2$ (TRI) ACC (TRI) TATGCGGATAGCGTGAAAGGCCGTTTT- ACCATTTCACGTGATAATTCGAAAAACACCA- 3'), and primer 5'-TGGTGTTTTTCGAATTATCA-3' for synthesis of the complementary strand, where (TRI) was a trinucleotide mixture representing all amino acids except Cys, (6) comprised the incorporation of (A/G) (A/C/G) T, resulting in the formation of 6 codons for Ala, Asn, Asp, Gly, Ser, and Thr, and the length distribution being obtained by performing one substoichiometric coupling of the (TRI) mixture during synthesis, omitting the capping step normally used in DNA synthesis.

[0142] DNA synthesis was performed on a 40 nmole scale, oligos were dissolved in TE buffer, purified via gel filtration using spin columns (S-200), and the DNA concentration determined by OD measurement at 260 nm (OD 1.0 = 40 μg/ml). 10 nmole of the oligonucleotide templates and 12 nmole of the corresponding primers were mixed and annealed at 80°C for 1 min, and slowly cooled down to 37°C within 20 to 30 min. The fill-in reaction was performed for 2 h at 37°C using Klenow polymerase (2.0 μl) and 250 nmole of each dNTP. The excess of dNTPs was removed by gel filtration using Nick-Spin columns (Pharmacia), and the double-stranded DNA digested with BbsI/MscI (L-CDR3), or XhoI/SfuI (H-CDR2) over night at 37°C. The cassettes were purified via Nick-Spin columns (Pharmacia), the concentration determined by OD measurement, and the cassettes aliquoted (15 pmole) for being stored at -80°C.

7.2 Library cloning:

[0143] DNA was prepared from the collection of FITC binding clones obtained in Example 2 (approx. 10^4 to clones). The collection of scFv fragments was isolated via XbaI/EcoRI digest. The vector pCAL4 (100 fmole, 10 μg) described in Example 4.3 was similarly digested with XbaI/EcoRI, gel-purified and ligated with 300 fmole of the scFv fragment collection over night at 16°C. The ligation mixture was isopropanol precipitated, air-dried, and the pellets were redissolved in 100 μl of dd H_2O. The ligation mixture was mixed with 1 ml of freshly prepared electrocompetent SCS 101 cells (for optimization of L-CDR3), or XL1 Blue cells (for optimization of H-CDR2) on ice. One round of electroporation was performed and the transformants were eluted in SOC medium, shaken at 37°C for 30 minutes, and an aliquot plated out on LB plates (Amp/Tet/Glucose) at 37°C for 6-9 hrs. The number of transformants was 5 x 10^4.

Vector DNA (100 μg) was isolated and digested (sequence and restriction map of scH3κ2 see Figure 8) with BbsI/MscI for optimization of L-CDR3, or XhoI/NspV for optimization of H-CDR2. 10 μg of purified vector fragments (5 pmole) were ligated with 15 pmole of the L-CDR3 or H-CDR2 library cassettes over night at 16°C. The ligation mixtures were isopropanol precipitated, air-dried, and the pellets were redissolved in 100 μl of dd H_2O. The ligation mixtures were mixed with 1 ml of freshly prepared electrocompetent XL1 Blue cells on ice. Electroporation was performed and the transformants were eluted in SOC medium and shaken at 37°C for 30 minutes. An aliquot was plated out on LB plates (Amp/Tet/Glucose) at 37°C for 6-9 hrs. The number of transformants (library size) was greater than 10^8 for both libraries. The libraries were stored as glycerol cultures.

7.3. Bio panning

[0144] This was performed as described for the initial H3κ2 H-CDR3 library in Example 2.1. Optimized scFvs binding to FITC could be characterized and analyzed as described in Example 2.2 and 2.3, and further rounds of optimization could be made if necessary.

References

[0145]

Barbas III, C.F., Bain, J.D., Hoekstra, D.M. & Lerner, R.A., PNAS 89, 4457-4461 (1992).
Better, M., Chang, P., Robinson, R. & Horwitz, A.H., Science 240, 1041-1043 (1988).
Blake, M.S., Johnston, K.H., Russel-Jones, G.J. & Gotschlich, E.C., Anal. Biochem. 136, 175-179 (1984).
Carter, P., Kelly, R.F., Rodrigues, M.L., Snedecor, B., Covrrubias, M., Velligan, M.D., Wong, W.L.T., Rowland, A.M., Kotts, C.E., Carver, M.E., Yang, M., Bourell, J.H., Shepard, H.M. & Henner, D., Bio/Technology 10, 163-167 (1992).
Chothia, C. & Lesk, A.M., J. Biol. Chem. 196, 910-917 (1987).
Chothia, C., Lesk, A.M., Gherardi, E., Tomlinson, I.A., Walter, G., Marks, J.D., Llewelyn, M.B. & Winter, G., J. Mol. Biol. 227. 799-817 (1992).
Chothia, C., Lesk, A.M., Tramontano, A., Levitt, M., Smith-Gill, S.J., Air, G., Sheriff, S., Padlan, E.A., Davies, D., Tulip, W.R., Colman, P.M., Spinelli, S., Alzari, P.M. & Poljak, R.J., Nature 342, 877-883 (1989).
Chuchana, P., Blancher, A., Brockly, F., Alexandre, D., Lefranc, G & Lefranc, M.-P., Eur. J. Immunol. 20, 1317-1325 (1990).
Cox, J.P.L., Tomlinson, I.M. & Winter, G., Eur. J. Immunol. 24, 827-836 (1994).
Ge, L., Knappik, A., Pack, P., Freund, C. & Plückthun, A., In: Antibody Engineering. Borrebaeck, C.A.K. (Ed.). p. 229-266 (1995), Oxford University Press, New York, Oxford.)
Gill, S.C. & von Hippel, P.H., Anal. Biochem. 182, 319.326 (1989).
Hochuli, E., Bannwarth, W., Döbeli, H., Gentz, R. & Stüber, D., Bio/Technology 6. 1321-1325 (1988).
Hopp, T.P., Prickett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L. & Conlon, P.J. Bio/Technology 6, 1204-1210 (1988).
Kabat, E.A., Wu, T.T., Perry, H.M., Gottesmann, K.S. & Foeller, C., Sequences of proteins of immunological interest, NIH publication 91-3242 (1991).
Knappik, A. & Plückthun, A., Biotechniques 17, 754-761 (1994).
Knappik, A. & Plückthun, A., Protein Engineering 8, 81-89 (1995).
Kunkel, T.A., Bebenek, K. & McClary, J., Methods in Enzymol. 204, 125-39 (1991).
Lindner, P., Guth, B., Wülfing, C., Krebber, C., Steipe, B., Müller, F. & Plückthun, A., Methods: A Companion to Methods Enzymol. 4, 41-56 (1992).
Lowman, H.B., Bass, S.H., Simpson, N. and Wells, J.A., Biochemistry 30, 10832-10838 (1991).
Pack, P. & Plückthun, A., Biochemistry 31, 1579-1584 (1992).
Pack, P., Kujau, M., Schroeckh, V., Knüpfer, U., Wenderoth, R., Riesenberg D. & Plückthun, A., Bio/Technology 11, 1271-1277 (1993).
Pack, P., Ph.D. thesis, Ludwig-Maximilians-Universitat München (1994).
Perlak, F. J., Nuc. Acids Res. 18, 7457-7458 (1990).
Plückthun, A., Krebber, A., Krebber, C., Horn, U., Knüpfer, U., Wenderoth, R., Nieba, L., Proba, K. & Riesenberg, D., A practical approach. Antibody Engineering (Ed. J. McCafferty). IRL Press, Oxford, pp. 203-252 (1996).
Prodromou, C. & Pearl, L.H., Protein Engineering 5, 827-829 (1992).
Rosenberg, S.A. & Lotze, M.T., Ann. Rev. Immunol. 4, 681-709 (1986).
Skerra, A. & Pluckthun, A., Science 240, 1038-1041 (1988).
Skerra, A., Pfitzinger, I. & Plückthun, A., Bio/Technology 9, 273-278 (1991).
Sutherland, L., Davidson, J., Glass, L.L., & Jacobs, H.T., BioTechniques 18, 458-464, 1995.
Tomlinson, I.M., Walter, G., Marks, J.D., Llewelyn, M.B. & Winter, G., J. Mol. Biol. 227, 776-798 (1992).
Ullrich, H.D., Patten, P.A., Yang, P.L., Romesberg, F.E. & Schultz, P.G., Proc. Natl. Acad. Sci. USA 92, 11907-11911 (1995).
Van Dijk, K.W., Mortari, F., Kirkham, P.M., Schroeder Jr., H.W. & Milner, E.C.B., Eur. J. Immunol. 23, 832-839 (1993).
Virnekäs, B., Ge, L., Plückthun, A., Schneider, K.C., Wellnhofer, G. & Moroney, S.E., Nucleic Acids Research 22, 5600-5607 (1994).
Vitetta, E.S., Thorpe, P.E. & Uhr, J., Immunol. Today 14, 253-259 (1993).
Williams, S.C. & Winter, G., Eur. J. Immunol. 23, 1456-1481 (1993).
Winter, G., Griffiths, A.D., Hawkins, R.E. & Hoogenboom, H.R., Ann. Rev. Immunol. 12, 433-455 (1994).

Table 1A: Human kappa germline gene segments

Used Name[1]	Reference[2]	Family[3]	Germline genes[4]
Vk1-1	9	1	O8; O18; DPK1
Vk1-2	1	1	L14; DPK2
Vk1-3	2	1	L15(1); HK101; HK146; HK189
Vk1-4	9	1	L11
Vk1-5	2	1	A30
Vk1-6	1	1	LFVKS
Vk1-7	1	1	LFVK431
Vk1-8	1	1	L1; HK137
Vk1-9	1	1	A20; DPK4
Vk1-10	1	1	L18; Va"
Vk1-11	1	1	L4; L18; Va'; V4a
Vk1-12	2	1	L5; L19(1); Vb; Vb4; DPK5; L19(2); Vb"; DPK6
Vk1-13	2	1	L15(2); HK134; HK166; DPK7
Vk1-14	8	1	L8; Vd; DPK8
Vk1-15	8	1	L9; Ve
Vk1-16	1	1	L12(1); HK102; V1
Vk1-17	2	1	L12(2)
Vk1-18	1	1	O12a (V3b)
Vk1-19	6	1	O2; O12; DPK9
Vk1-20	2	1	L24; Ve"; V13; DPK10
Vk1-21	1	1	O4; O14
Vk1-22	2	1	L22
Vk1-23	2	1	L23
Vk2-1	1	2	A2; DPK12
Vk2-2	6	2	O1; O11(1); DPK13
Vk2-3	6	2	O12(2); V3a
Vk2-4	2	2	L13
Vk2-5	1	2	DPK14
Vk2-6	4	2	A3; A19; DPK15
Vk2-7	4	2	A29; DPK27
Vk2-8	4	2	A13
Vk2-9	1	2	A23
Vk2-10	4	2	A7; DPK17
Vk2-11	4	2	A17; DPK18
Vk2-12	4	2	A1; DPK19
Vk3-1	11	3	A11; humkv305; DPK20
Vk3-2	1	3	L20; Vg"
Vk3-3	2	3	L2; L16; humkv328; humkv328h2; humkv328h5; DPK21
Vk3-4	11	3	A27; humkv325; VkRF; DPK22
Vk3-5	2	3	L25; DPK23
Vk3-6	2	3	L10(1)
Vk3-7	7	3	L10(2)
Vk3-8	7	3	L6; Vg
Vk4-1	3	4	B3; VkIV; DPK24
Vk5-1	10	5	B2; EV15
Vk6-1	12	6	A14; DPK25
Vk6-2	12	6	A10; A26; DPK26
Vk7-1	5	7	B1

Table 1B: Human lambda germline gene segments

Used Name[1]	Reference[2]	Family[3]	Germline genes[4]
DPL1	1	1	
DPL2	1	1	HUMLV1L1
DPL3	1	1	HUMLV122
DPL4	1	1	VLAMBDA 1.1
HUMLV117	2	1	
DPL5	1	1	HUMLV117D
DPL6	1	1	
DPL7	1	1	IGLV1S2
DPL8	1	1	HUMLV1042
DPL9	1	1	HUMLV101
DPL10	1	2	
VLAMBDA 2.1	3	2	
DPL11	1	2	
DPL12	1	2	
DPL13	1	2	
DPL14	1	2	
DPL16	1	3	Humlv418; IGLV3S1
DPL23	1	3	VIIIL1
Humlv318	4	3	
DPL18	1	7	4A; HUMIGLVA
DPL19	1	7	
DPL21	1	8	VL8.1
HUMLV801	5	8	
DPL22	1	9	
DPL24	1	unassigned	VLAMBDA N.2
gVLX-4.4	6	10	

Table 1C: Human heavy chain germline gene segments

Used Name[1]	Rererence[2]	Family[3]	Germline genes[4]
VH1-12-1	19	1	DP10; DA-2; DA-6
VH1-12-8	22	1	RR.VH1.2
VH1-12-2	6	1	hvl263
VH1-12.9	7	1	YAC-7; RR.VH1.1; 1-69
VH1-12-3	19	1	DP3
VHL-12-4	19	1	DP21; 4d275a; VH7a
VH1-12-5	18	1	1-4.1b; V1-4.1b
VH1-12-6	21	1	1D37;VH7b;7-81;YAC-10
VH1-12-7	19	1	DP14; VH1GRR; V1-18
VH1-13-1	10	1	71-5; DP2
VH1-13-2	10	1	E3-10
VH1-13-3	19	1	DP1
VH1-13-4	12	1	V35
VH1-13-5	8	1	V1-2b
VH1-13-6	18	1	1-2; DP75
VH1-13-7	21	1	V1-2
VH1-13-8	19	1	DP8
VH1-13-9	3	1	1-1

(continued)

Used Name[1]	Rererence[2]	Family[3]	Germline genes[4]
VH1-13-10	19	1	DP12
VH1-13-11	15	1	V13C
VH1-13-12	18	1	1-3b; DP25; V1-3b
VH1-13-13	3	1	1-92
VH1-13-14	18	1	1-3; V1-3
VH1-13-15	19	1	DP15; V1-8
VH1-13-16	3	1	21-2; 3-1; DP7; V1-46
VH1-13-17	16	1	HG3
VH1-13-18	19	1	DP4; 7-2; V1-45
VH1-13-19	27	1	COS 5
VH1-1X-1	19	1	DP5; 1-24P
VH2-21-1	18	2	II-5b
VH2-31-1	2	2	VH2S12-1
VH2-31-2	2	2	VH2S12-7
VH2-31-3	2	2	VH2S12-9; DP27
VH2-31-4	2	2	VH2S12-10
VH2-31-5	14	2	V2-26; DP26; 2-26
VH2-31-6	15	2	VF2-26
VH2-31-7	19	2	DP28; DA-7
VH2-31-14	7	2	YAC-3; 2-70
VH2-31-8	2	2	VH2S12-5
VH2-31-9	2	2	VH2S12-12
VH2-31.10	18	2	H-5: V2-5
VH2-31-11	2	2	VH2S12-2; VH2S12-8
VH2-31-12	2	2	VH2S12-4; VH2S12-6
VH2-31-13	2	2	VH2S12-14
VH3-11-1	13	3	v65-2; DP44
VH3-11-2	19	3	DP45
VH3-11-3	3	3	13-2; DP48
VH3-11-4	19	3	DP52
VH3-11-5	14	3	v3-13
VH3-11-6	19	3	DP42
VH3-11-7	3	3	8-1B; YAC-5; 3-66
VH3-11-8	14	3	V3-53
VH3-13-1	3	3	22-2B; DP35; V3-11
VH3-13-5	19	3	DP59; VH19; V3-35
VH3-13-6	25	3	fl-pl; DP61
VH3-13-7	19	3	DP46; GL-SJ2; COS 8; hv3005; hv3005f3; 3d21b; 56pl
VH3-13-8	24	3	VH26
VH3-13-9	5	3	vh26c
VH3-13-10	19	3	DP47; VH26; 3-23
VH3-13-11	3	3	1-91
VH3-13-12	19	3	DP58
VH3-13-13	3	3	1-9111; DP49; 3-30; 3d28.1
VH3-13-14	24	3	3019B9; DP50; 3-33; 3d277
VH3-13-15	27	3	COS 3
VH3-13-16	19	3	DP51
VH3-13-17	16	3	H11
VH3-13-18	19	3	DP53; COS 6; 3-74; DA-8
VH3-13-19	19	3	DP54: VH3-11; V3-7

(continued)

Used Name[1]	Rererence[2]	Family[3]	Germline genes[4]
VH3-13-20	14	3	V3-64; YAC-6
VH3-13-21	14	3	V3-48
VH3-13-22	14	3	V3-43; DP33
VH3-13-23	14	3	V3-33
VH3-13-24	14	3	V3-21; DP77
VH3-13-25	14	3	V3-20; DP32
VH3-13-26	14	3	V3-9; DP31
VH3-14-1	3	3	12-2; DP29; 3-72; DA-3
VH3-14-4	7	3	YAC-9; 3-73; MTGL
VH3-14-2	4	3	VHD26
VH3-14-3	19	3	DP30
VH3-1X-1	1	3	LSG8.1; LSG9.1; LSG10.1; HUM121GVH; HUM13lGVH
VH3-1X-2	1	3	LSG11.1; HUM41GVH
VH3-1X-3	3	3	9-1; DP38; LSG7.1; RCG1.1; LSG1.1; LSG3.1; LSG5.1; HUM15lGVH; HUM21GVH; HUM91GVH
VH3-1X-4	1	3	LSG4.1
VH3-1X-5	1	3	LSG2.1
VH3-1X-6	1	3	LSG6.1; HUM10lGVH
VH3-1X-7	18	3	3-15; V3-15
VH3-1X-8	1	3	LSG12.1; HUM5lGVH
VH3-1X-9	14	3	V3-49
VH4-11-1	22	4	Tou-VH4.21
VH4-11-2	17	4	VH4.21; DP63; VH5; 4d76; V4-34
VH4-11-3	23	4	4.44
VH4-11-4	23	4	4.44.3
VH4-11-5	23	4	4.36
VH4-11-6	23	4	4.37
VH4-11-7	18	4	IV-4; 4.35: V4-4
VH4-11-8	17	4	VH4.11; 3d197d; DP71; 58p2
VH4-11-9	20	4	H7
VH4-11-10	20	4	H8
VH4-11-11	20	4	H9
VH4-11-12	17	4	VH4.16
VH4-11-13	23	4	4.38
VH4-11-14	17	4	VH4.15
VH4-11-15	11	4	58
VH4-11-16	10	4	71-4; V4-59
VH4-21-1	11	4	11
VH4-21-2	17	4	VH4.17; VH4.23; 4d255; 4.40; DP69
VH4-21-3	17	4	VH4.19; 79; V4-4b
VH4-21-4	19	4	DP70; 4d68; 4.41
VH4-21-5	19	4	DP67; VH4-4B
VH4-21-6	17	4	VH4.22; VHSP; VH-JA
VH4-21-7	17	4	VH4.13; 1-911; 12G-1; 3d28d; 4.42; DP68; 4-28
VH4-21-8	26	4	hv4005; 3d24d
VH4-21-9	17	4	VH4.14
VH4-31-1	23	4	4.34; 3d230d; DP78
VH4-31-2	23	4	4.34.2
VH4-31-3	19	4	DP64; 3d216d
VH4-31-4	19	4	DP65; 4-31; 3d277d

(continued)

Used Name[1]	Rererence[2]	Family[3]	Germline genes[4]
VH4-31-5	23	4	4.33; 3d75d
VH4-31-6	20	4	H10
VH4-31-7	20	4	H11
VH4-31-8	23	4	4.31
VH4-31-9	23	4	4.32
VH4-31-10	20	4	3d277d
VH4-31-11	20	4	3d216d
VH4-31-12	20	4	3d279d
VH4-31-13	17	4	VH4.18; 4d154; DP79
VH4-31-14	8	4	V4-39
VH4-31-15	11	4	2-1; DP79
VH4-31-16	23	4	4.30
VH4-31-17	17	4	VH4.12
VH4-31-18	10	4	71-2; DP66
VH4-31-19	23	4	4.39
VH4-31-20	8	4	V4-61
VH5-12-1	9	5	VH251; DP73; VHVCW; 51-R1; VHVLB; VHVCH; VHVTT; VHVAU; VHVBLK; VhAU; V5-51
VH5-12-2	17	5	VHVJB
VH5-12-3	3	5	1-v; DP80; 5-78
VH5-12-4	9	5	VH32; VHVRG; VHVMW; 5-2R1
VH6-35-1	4	6	VHVI; VH6; VHVIIS; VHVITE; VHVIJB; VHVICH; VHVICW; VHVIBLK; VHVIMW; DP74; 6-1G1; V6-1

Table 2A: rearranged human kappa sequences

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
III-3R	108	1	O8	1	1,1%	70
No.86	109	1	O8	3	3,2%	80
AU	108	1	O8	6	6,3%	103
ROY	108	1	O8	6	6.3%	43
IC4	108	1	O8	6	6,3%	70
HIV-B26	106	1	O8	3	3,2%	8
GRI	108	1	O8	8	8,4%	30
AG	106	1	O8	8	8,6%	116
REI	108	1	O8	9	9,5%	86
CLL PATIENT 16	88	1	O8	2	2,3%	122
CLL PATIENT 14	87	1	O8	2	2,3%	122
CLL PATIENT 15	88	1	O8	2	2,3%	122
GM4672	108	1	O8	11	11,6%	24
HUM. YFC51.1	108	1	O8	12	12,6%	110
LAY	108	1	O8	12	12,6%	48
HIV-b13	106	1	O8	9	9,7%	8
MAL-NaCl	108	1	O8	13	13,7%	102
STRAb SA-1A	108	1	O2	0	0,0%	120

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
HuVHCAMP	108	1	O8	13	13,7%	100
CRO	108	1	O2	10	10,5%	30
Am107	108	1	O2	12	12,6%	108
WALKER	107	1	O2	4	4,2%	57
III-2R	109	1	A20	0	0,0%	70
FOG1-A4	107	1	A20	4	4,2%	41
HK137	95	1	L1	0	0,0%	10
CEA4-8A	107	1	O2	7	7,4%	41
Va'	95	1	L4	0	0,0%	90
TRI.21	108	1	O2	4	4,2%	92
HAU	108	1	O2	6	6,3%	123
HKI02	95	1	L12(1)	0	0,0%	9
H20C3K	108	1	L12(2)	3	3,2%	125
CHEB	108	1	O2	7	7,4%	5
HK134	95	1	L15(2)	0	0,0%	10
TEL9	108	1	O2	9	9,5%	73
TR 1.32	103	1	O2	3	3,2%	92
RF-KES	97	1	A20	4	4,2%	121
WES	108	1	L5	10	10,5%	61
DILpl	95	1	O4	1	1,1%	70
SA-4B	107	1	L12(2)	8	8,4%	120
HK101	95	1	L15(1)	0	0,0%	9
TR1.23	108	1	O2	5	5,3%	92
HF2-1/17	108	1	A30	0	0,0%	4
2E7	108	1	A30	1	1,1%	62
33.C9	107	1	L12(2)	7	7,4%	126
3D6	105	1	L12(2)	2	2,1%	34
1-2a	108	1	L8	8	8,4%	70
RF-KL1	97	1	L8	4	4,2%	121
TNF-E7	108	1	A30	9	9.5%	41
TRI.22	108	1	O2	7	7,4%	92
HIV-B35	106	1	O2	2	2,2%	8
HIV-b22	106	1	O2	2	2,2%	8
HIV-b27	106	1	O2	2	2,2%	8
HIV-B8	107	1	O2	10	10,8%	8
HIV-b8	107	1	O2	10	10,8%	8
RF-SJ5	95	1	A30	5	5,3%	113
GAL(I)	108	1	A30	6	6,3%	64
R3.SHSG	108	1	O2	6	6,3%	70
HIV-b14	106	1	A20	2	2,2%	8
TNF-E1	105	1	L5	8	8,4%	41
WEA	108	1	A30	8	8,4%	37
EU	108	1	L12(2)	5	5,3%	40
FOG1-G8	108	1	L8	11	11,6%	41
1X7RG1	108	1	L1	8	8,4%	70
BLI	108	1	L8	3	3,2%	72
KUE	108	1	L12(2)	11	11,6%	32
LUNm01	108	1	L12(2)	10	10,5%	6
HIV-b1	106	1	A20	4	4,3%	8

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
HIV-s4	103	1	O2	2	2,2%	8
CAR	107	1	L12(2)	11	11,7%	79
BR	107	1	L12(2)	11	11,6%	50
CLL PATIENT 10	88	1	O2	0	0,0%	122
CLL PATIENT 12	88	1	O2	0	0,0%	122
KING	108	1	L12(2)	12	12,6%	30
V13	95	1	L24	0	0,0%	46
CLL PATIENT 11	87	1	O2	0	0,0%	122
CLL PATIENT 13	87	1	O2	0	0,0%	122
CLLPATIENT9	88	1	O12	1	1,1%	122
HIV-B2	106	1	A20	9	9,7%	8
HIV-b2	106	1	A20	9	9,7%	8
CLL PATIENT 5	88	1	A20	1	1,1%	122
CLL PATIENT 1	88	1	L8	2	2,3%	122
CLL PATIENT 2	88	1	L8	0	0,0%	122
CLL PATIENT 7	88	1	L5	0	0,0%	122
CLL PATIENT 8	88	1	L5	0	0,0%	122
HIV-b5	105	1	L5	11	12,0%	8
CLL PATIENT 3	87	1	L8	1	1,1%	122
CLL PATIENT 4	88	1	L9	0	0,0%	122
CLL PATIENT 18	85	1	L9	6	7,1%	122
CLL PATIENT 17	86	1	L12(2)	7	8,1%	122
HIV-b20	107	3	A27	11	11,7%	8
2C12	108	1	L12(2)	20	21,1%	68
IB11	108	1	L12(2)	20	21,1%	68
1H1	108	1	L12(2)	21	22,1%	68
2A12	108	1	L12(2)	21	22,1%	68
CUR	109	3	A27	0	0,0%	66
GLO	109	3	A27	0	0,0%	16
RF-TS1	96	3	A27	0	0,0%	121
GAR’	109	3	A27	0	0,0%	67
FLO	109	3	A27	0	0,0%	66
PIE	109	3	A27	0	0,0%	91
HAH 14.1	109	3	A27	1	1,0%	51
HAH 14.2	109	3	A27	1	1,0%	51
HAH 16.1	109	3	A27	1	1,0%	51
NOV	109	3	A27	1	1,0%	52

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
33.F12	108	3	A27	1	1,0%	126
8E10	110	3	A27	1	1,0%	25
TH3	109	3	A27	1	1,0%	25
HIC (R)	108	3	A27	0	0,0%	51
SON	110	3	A27	1	1,0%	67
PAY	109	3	A27	1	1,0%	66
GOT	109	3	A27	1	1,0%	67
mAbA6H4C5	109	3	A27	1	1,0%	12
BOR'	109	3	A27	2	2,1%	84
RF-SJ3	96	3	A27	2	2,1%	121
SIE	109	3	A27	2	2,1%	15
ESC	109	3	A27	2	2,1%	98
HEW'	110	3	A27	2	2,1%	98
YES8c	109	3	A27	3	3,1%	33
TI	109	3	A27	3	3,1%	114
mAb113	109	3	A27	3	3,1%	71
HEW	107	3	A27	0	0,0%	94
BRO	106	3	A27	0	0,0%	94
ROB	106	3	A27	0	0,0%	94
NG9	96	3	A27	4	4,2%	11
NEU	109	3	A27	4	4,2%	66
WOL	109	3	A27	4	4,2%	2
35G6	109	3	A27	4	4,2%	59
RF-SJ4	109	3	A11	0	0,0%	88
KAS	109	3	A27	4	4,2%	84
BRA	106	3	A27	1	1,1%	94
HAH	106	3	A27	1	1,1%	94
HIC	105	3	A27	0	0,0%	94
FS-2	109	3	A27	6	6,3%	87
JH'	107	3	A27	6	6,3%	38
EV1-15	109	3	A27	6	6,3%	83
SCA	108	3	A27	6	6,3%	65
mAb112	109	3	A27	6	6,3%	71
SIC	103	3	A27	3	3,3%	94
SA-4A	109	3	A27	6	6,3%	120
SER	108	3	A27	6	6,3%	98
GOL'	109	3	A27	7	7,3%	82
B5G10K	105	3	A27	9	9,7%	125
HG2B10K	110	3	A27	9	9,4%	125
Taykv322	105	3	A27	5	5,4%	52
CLL PATIENT 24	89	3	A27	1	1,1%	122
HIV-b24	107	3	A27	7	7,4%	8
HIV-b6	107	3	A27	7	7,4%	8
Taykv310	99	3	A27	1	1,1%	52
KA3D1	108	3	L6	0	0,0%	85
19.E7	107	3	L6	0	0,0%	126
rsv6L	109	3	A27	12	12,5%	7
Taykv320	98	3	A27	1	1,2%	52

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
Vh	96	3	L10(2)	0	0,0%	89
LS8	108	3	L6	1	1,1%	109
LS1	108	3	L6	1	1,1%	109
LS2S3-3	107	3	L6	2	2,1%	99
LS2	108	3	L6	1	1,1%	109
LS7	108	3	L6	1	1,1%	109
LS2S3-4d	107	3	L6	2	2,1%	99
LS2S3-4a	107	3	L6	2	2,1%	99
LS4	108	3	L6	1	1,1%	109
LS6	108	3	L6	1	1,1%	109
LS2S3-10a	107	3	L6	2	2,1%	99
LS2S3-8c	107	3	L6	2	2,1%	99
LS5	108	3	L6	1	1,1%	109
LS2S3-5	107	3	L6	3	3,2%	99
LUNm03	109	3	A27	13	13,5%	6
IARC/BL41	108	3	A27	13	13,7%	55
slkv22	99	3	A27	3	3,5%	13
POP	108	3	L6	4	4,2%	111
LS2S3-10b	107	3	L6	3	3,2%	99
LS2S3-8f	107	3	L6	3	3,2%	99
LS2S3-12	107	3	L6	3	3,2%	99
HIV-B30	107	3	A27	11	11,7%	8
HIV-B20	107	3	A27	11	11,7%	8
HIV-b3	108	3	A27	11	11,7%	8
HIV-s6	104	3	A27	9	9,9%	8
YSE	107	3	L2/L16	1	1,1%	72
POM	109	3	L2/L16	9	9,4%	53
Humkv328	95	3	L2/L16	1	1,1%	19
CLL	109	3	L2/L16	3	3,2%	47
LES	96	3	L2/L16	3	3,2%	38
HIV-s5	104	3	A27	11	12,1%.	8
HIV-s7	104	3	A27	11	12,1%	8
slkvl	99	3	A27	7	8,1%	13
Humka3 les	95	3	L2/L16	4	4,2%	18
slkv12	101	3	A27	8	9,2%	13
RF-TS2	95	3	L2/L16	3	3,2%	121
11-1	109	3	L2/L16	4	4,2%	70
HIV-s3	105	3	A27	13	14,3%	8
RF-TMC1	96	3	L6	10	10,5%	121
GER	109	3	L2/L16	7	7,4%	75
GF4/1.1	109	3	L2/L16	8	8,4%	36
mAb114	109	3	L2/L16	6	6,3%	71
HIV-loop13	109	3	L2/L16	7	7,4%	8
bkv16	86	3	L6	1	1,2%	13
CLL PATIENT 29	86	3	L6	1	1,2%	122
slkv9	98	3	L6	3	3,5%	13
bkv17	99	3	L6	1	1,2%	13
slkv14	99	3	L6	1	1,2%	13

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
slkv16	101	3	L6	2	2,3%	13
bkv33	101	3	L6	4	4,7%	13
slkv15	99	3	L6	2	2,3%	13
bkv6	100	3	L6	3	3,5%	13
R6B8K	108	3	L2/L16	12	12,6%	125
AL 700	107	3	L2/L16	9	9,5%	117
slkv11	100	3	L2/L16	3	3,5%	13
slkv4	97	3	L6	4	4,8%	13
CLL PATIENT 26	87	3	L2/L16	1	1,1%	122
AL Sel24	103	3	L2/L16	9	9,5%	117
slkv13	100	3	L2/L16	6	7,0%	13
bkv7	100	3	L2/L16	5	5,8%	13
bkv22	100	3	L2/L16	6	7,0%	13
CLL PATIENT 27	84	3	L2/L 1 6	0	0,0%	122
bkv35	100	3	L6	8	9,3%	13
CLL PAT1ENT 25	87	3	L2/L16	4	4,6%	122
slkv3	86	3	L2/L16	7	8,1%	13
slkv7	99	1	O2	7	8,1%	13
HuFd79	111	3	L2/L16	24	24,2%	21
RAD	99	3	A27	9	10,3%	78
CLL PATIENT 28	83	3	L2/L16	4	4,8%	122
REE	104	3	L2/L16	25	27,2%	95
FR4	99	3	A27	8	9,2%	77
MD3.3	92	3	L6	1	1,3%	54
MD3.1	92	3	L6	0	0.0%	54
GA3.6	92	3	L6	2	2,6%	54
M3.5N	92	3	L6	3	3,8%	54
WEI’	82	3	A27	0	0,0%	65
MD3.4	92	3	L2/L16	1	1,3%	54
MD3.2	91	3	L6	3	3,8%	54
VER	97	3	A27	19	22,4%	20
CLL PATIENT 30	78	3	L6	3	3,8%	122
M3.1N	92	3	L2/L16	1	1,3%	54
MD3.6	91	3	L2/L16	0	0,0%	54
MD3.8	91	3	L2/L16	0	0,0%	54
GA3.4	92	3	L6	7	9,0%	54
M3.6N	92	3	A27	0	0,0%	54
MD3.10	92	3	A27	0	0,0%	54
MD3.13	91	3	A27	0	0,0%	54
MD3.7	93	3	A27	0	0,0%	54
MD3.9	93	3	A27	0	0,0%	54
GA3.1	93	3	A27	6	7,6%	54
bkv32	101	3	A27	5	5,7%	13
GA3.5	93	3	A27	5	6,3%	54

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
GA3.7	92	3	A27	7	8,9%	54
MD3.12	92	3	A27	2	2,5%	54
M3.2N	90	3	L6		7,8%	54
MD3.5	92	3	A27	1	1,3%	54
M3.4N	91	3	L2/L16	8	10.3%	54
M3.8N	91	3	L2/L16	7	9,0%	54
M3.7N	92	3	A27	3	3,8%	54
GA3.2	92	3	A27	9	11,4%	54
GA3.8	93	3	A27	4	5,1%	54
GA3.3	92	3	A27	8	10,1%	54
M3.3N	92	3	A27	5	6,3%	54
B6	83	3	A27	8	11,3%	78
E29.1 KAPPA	78	3	L2/L16	0	0,0%	22
SCW	108	1	O8	12	12,6%	31
RE1-based CAMPATH-9	107	1	O8	14	14,7%	39
RZ	107	1	O8	14	14,7%	50
BI	108	1	O8	14	14,7%	14
AND	107	1	O2	13	13,7%	69
2A4	109	1	O2	12	12,6%	23
KA	108	1	O8	19	20,0%	107
MEV	109	1	O2	14	14,7%	29
DEE	106	1	O2	13	14,0%	76
OU(IOC)	108	1	O2	18	18,9%	60
HuRSV19VK	111	1	08	21	21,0%	115
SP2	108	1	O2	17	17,9%	93
BJ26	99	1	O8	21	24,1%	1
NI	112	1	O8	24	24,2%	106
BMA 0310EUCIV2	106	1	L12(1)	21	22,3%	105
CLL PATIENT 6	71	1	A20	0	0,0%	122
BJ19	85	1	O8	16	21,9%	1
GM 607	113	2	A3	0	0,0%	58
R5A3K	114	2	A3	1	1,0%	125
R1C8K	114	2	A3	1	1,0%	125
VK2.R149	113	2	A3	2	2,0%	118
TR1.6	109	2	A3	4	4,0%	92
TR1.37	104	2	A3	5	5,0%	92
FS-1	113	2	A3	6	6,0%	87
TR1.8	110	2	A3	6	6,0%	92
NIM	113	2	A3	8	8,0%	28
Inc	112	2	A3	11	11,0%	35
TEW	107	2	A3	6	6,4%	96
CUM	114	2	O1	7	6,9%	44
HRF1	71	2	A3	4	5,6%	124
CLL PATIENT 19	87	2	A3	0	0,0%	122

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
CLL PATIENT 20	87	2	A3	0	0,0%	122
MIL	112	2	A3	16	16,2%	26
FR	113	2	A3	20	20,0%	101
MAL-Urine	83	1	O2	6	8,6%	102
Taykv306	73	3	A27	1	1,6%	52
Taykv312	75	3	A27	1	1,6%	52
HIV-b29	93	3	A27	14	17,5%	8
1-185-37	110	3	A27	0	0,0%	119
1-187-29	110	3	A27	0	0,0%	119
TT117	110	3	A27	9	9,4%	63
HIV-loop8	108	3	A27	16	16,8%	8
rsv23L	108	3	A27	16	16,8%	7
HIV-b7	107	3	A27	14	14,9%	8
HIV-b11	107	3	A27	15	16,0%	8
HIV-LC1	107	3	A27	19	20,2%	8
HIV-LC7	107	3	A27	20	21,3%	8
HIV-LC22	107	3	A27	21	22,3%	8
HlV-LC13	107	3	A27	21	22,3%	8
HIV-LC3	107	3	A27	21	22,3%	8
HIV-LC5	107	3	A27	21	22,3%	8
HIV-LC28	107	3	A27	21	22,3%	8
HIV-b4	107	3	A27	22	23,4%	8
CLL PAT1ENT 31	87	3	A27	15	17,2%	122
HIV-loop2	108	3	L2/L16	17	17,9%	8
HIV-loop35	108	3	L2/L16	17	17,9%	8
HIV-LCI1	107	3	A27	23	24,5%	8
HIV-LC24	107	3	A27	23	24,5%	8
HIV-b12	107	3	A27	24	25,5%	8
HIV-LC25	107	3	A27	24	25,5%	8
HIV-b21	107	3	A27	24	25,5%	8
HIV-LC26	107	3	A27	26	27,7%	8
G3D10K	108	1	L12(2)	12	12,6%	125
TT125	108	1	L5	8	8,4%	63
HIV-s2	103	3	A27	28	31,1%	8
265-695	108	1	L5	7	7,4%	3
2-115-19	108	1	A30	2	2,1%	119
rsv13L	107	1	O2	20	21,1%	7
HIV-b18	106	1	O2	14	15,1%	8
RF-KL5	98	3	L6	36	36,7%	97
ZM1-1	113	2	A17	7	7,0%	3
HIV-s8	103	1	O8	16	17,8%	8
K-EV15	95	5	B2	0	0,0%	112
RF-TS3	100	2	A23	0	0,0%	121
HF-21/28	111	2	A17	1	1,0%	17
RPMI6410	113	2	A17	1	1,0%	42
JC11	113	2	A17	1	1,0%	49
O-81	114	2	A17	5	5,0%	45

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
FK-001	113	4	B3	0	0,0%	81
CD5+.28	101	4	B3	1	1,0%	27
LEN	114	4	B3	1	1,0%	104
UC	114	4	B3	1	1,0%	111
CDS+.5	101	4	B3	1	1,0%	27
CDS+.26	101	4	B3	1	1,0%	27
CDS+.12	101	4	B3	2	2.0%	27
CDS+.23	101	4	B3	2	2,0%	27
CDS+.7	101	4	B3	2	2,0%	27
VJI	113	4	B3	3	3,0%	56
LOC	113	4	B3	3	3,0%	72
MAL	113	4	B3	3	3.0%	72
CD5+.6	101	4	B3	3	3,0%	27
H2F	113	4	B3	3	3,0%	70
PB17IV	114	4	B3	4	4,0%	74
CD5+.27	101	4	B3	4	4,0%	27
CD5+.9	101	4	B3	4	4,0%	27
CDS-.28	101	4	B3	5	5,0%	27
CD5-.26	101	4	B3	6	5,9%	27
CD5+.24	101	4	B3	6	5,9%	27
CDS+.10	101	4	B3	6	5,9%	27
CDS-.19	101	4	B3	6	5,9%	27
CD5-.18	101	4	B3	7	6,9%	27
CDS-.16	101	4	B3	8	7,9%	27
CD5-.24	101	4	B3	8	7,9%	27
CD5-.17	101	4	B3	10	9,9%	27
MD4.1	92	4	B3	0	0.0%	54
MD4.4	92	4	B3	0	0,0%	54
MD4.5	92	4	B3	0	0,0%	54
MD4.6	92	4	B3	0	0,0%	54
MD4.7	92	4	B3	0	0,0%	54
MD4.2	92	4	B3	1	1,3%	54
MD4.3	92	4	B3	5	6,3%	54
CLL PATIENT 22	87	2	A17	2	2,3%	122
CLL PATIENT 23	84	2	A17	2	2,4%	122

Table 2B: rearranged human lambda sequences

Name'	aa[2]	Computed family[3]	Germline gene[4]	Diff. to gerinline[5]	% diff. to germline[6]	Reference[7]
WAH	110	1	DPL3	7	7%	68
1B9/F2	112	1	DPL3	7	7%	9
DIA	112	1	DPL2	7	7%	36
mAb67	89	1	DPL3	0	0%	29
HiH2	110	1	DPL3	12	11%	3
NIG-77	112	1	DPL2	9	9%	72

(continued)

Name'	aa[2]	Computed family[3]	Germline gene[4]	Diff. to gerinline[5]	% diff. to germline[6]	Reference[7]
OKA	112	1	DPL2	7	7%	84
KOL	112	1	DPL2	12	11%	40
T2:C5	111	1	DPL5	0	0%	6
T2:C14	110	1	DPL5	0	0%	6
PR-TS1	110	1	DPL5	0	0%	55
4G12	111	1	DPL5	1	1%	35
KIM46L	112	1	HUMLV117	0	0%	8
Fog-B	111	1	DPL5	3	3%	31
9F2L	111	1	DPL5	3	3%	79
mAb111	110	1	DPL5	3	3%	48
PHOX15	111	1	DPL5	4	4%	49
BL2	111	1	DPL5	4	4%	74
NIG-64	111	1	DPL5	4	4%	72
RF-SJ2	100	1	DPL5	6	6%	78
AL EZ1	112	1	DPL5	7	7%	41
ZIM	112	1	HUMLV117	7	7%	18
RF-SJ1	100	1	DPL5	9	9%	78
IGLV1.1	98	1	DPL4	0	0%	1
NEW	112	1	HUMLV117	11	10%	42
CB-201	87	1	DPL2	1	1%	62
MEM	109	1	DPL2	6	6%	50
H210	111	2	DPL10	4	4%	45
NOV	110	2	DPL10	8	8%	25
NEI	111	2	DPL10	8	8%	24
AL MC	110	2	DPL11	6	6%	28
MES	112	2	DPL11	8	8%	84
FOG1-A3	111	2	DPL11	9	9%	27
AL NOV	112	2	DPL11	7	7%	28
HMST-1	110	2	DPL11	4	4%	82
HBW4-1	108	2	DPL12	9	9%	52
WH	110	2	DPL11	11	11%	34
11-50	110	2	DPL11	7	7%	82
HBp2	110	2	DPL12	8	8%	3
NIG-84	113	2	DPL11	12	11%	73
VIL	112	2	DPL11	9	9%	58
TRO	111	2	DPL12	10	10%	61
ES492	108	2	DPL11	15	15%	76
mAb216	89	2	DPL12	1	1%	7
BSA3	109	3	DPL16	0	0%	49
THY-29	110	3	DPL16	0	0%	27
PR-TS2	108	3	DPL16	0	0%	55
E29.1 LAMBDA	107	3	DPL16	1	1%	13
mAb63	109	3	DPL16	2	2%	29
TEL14	110	3	DPL16	6	6%	49
6H-3C4	108	3	DPL16	7	7%	39
SH	109	3	DPL16	7	7%	70
AL GIL	109	3	DPL16	8	8%	23
H6-3C4	108	3	DPL16	8	8%	83
V-lambda-2.DS	111	2	DPL11	3	3%	15

(continued)

Name'	aa[2]	Computed family[3]	Germline gene[4]	Diff. to gerinline[5]	% diff. to germline[6]	Reference[7]
8.12 ID	110	2	DPL 11	3	3%	81
DSC	111	2	DPL11	3	3%	56
PV11	110	2	DPL11	1	1%	56
33.H11	110	2	DPL11	4	4%	81
AS17	111	2	DPL11	7	7%	56
SD6	110	2	DPL 11	7	7%	56
KS3	110	2	DPL11	9	9%	56
PV6	110	2	DPL12	5	5%	56
NGD9	110	2	DPL11	7	7%	56
MUC1-1	111	2	DPL11	11	10%	27
A30c	111	2	DPL10	6	6%	56
KS6	110	2	DPL12	6	6%	56
TEL13	111	2	DPL11	11	10%	49
AS7	110	2	DPL12	6	6%	56
MCG	112	2	DPL12	12	11%	20
U266L	110	2	DPL12	13	12%	77
PR-SJ2	110	2	DPL12	14	13%	55
BOH	112	2	DPL12	11	10%	37
TOG	111	2	DPL11	19	18%	53
TEL16	111	2	DPL11	19	18%	49
No.13	110	2	DPL10	14	13%	52
BO	112	2	DPL12	18	17%	80
WIN	112	2	DPL12	17	16%	11
BUR	104	2	DPL12	15	15%	46
NIG-58	110	2	DPL12	20	19%	69
WEIR	112	2	DPL11	26	25%	21
THY-32	111	1	DPL8	8	8%	27
TNF-H9G1	111	1	DPL8	9	9%	27
mAb61	111	1	DPL3	1	1%	29
LV1L1	98	1	DPL2	0	0%	54
HA	113	1	DPL3	14	13%	63
LA1L1	111	1	DPL2	3	3%	54
RHE	112	1	DPL1	17	16%	22
KIB12L	113	1	DPL8	17	16%	79
LOC	113	1	DPL2	15	14%	84
NIG-51	112	1	DPL2	12	11%	67
NEWM	104	1	DPL8	23	22%	10
MD3-4	106	3	DPL23	14	13%	4
COX	112	1	DPL2	13	12%	84
HiH10	106	3	DPL23	13	12%	3
VOR	112	1	DPL2	16	15%	16
AL POL	113	1	DPL2	16	15%	57
CD4-74	111	1	DPL2	19	18%	27
AMYLOID MOL	102	3	DPL23	15	15%	30
OST577	108	3	Humlv318	10	10%	4
NIG-48	113	1	DPL3	42	40%	66
CARR	108	3	DPL23	18	17%	19
mAb60	108	3	DPL23	14	13%	29
NIG-68	99	3	DPL23	25	26%	32

(continued)

Name'	aa[2]	Computed family[3]	Germline gene[4]	Diff. to gerinline[5]	% diff. to germline[6]	Reference[7]
KERN	107	3	DPL23	26	25%	59
ANT	106	3	DPL23	17	16%	19
LEE	110	3	DPL23	18	17%	85
CLE	94	3	DPL23	17	17%	19
VL8	98	8	DPL21	0	0%	81
MOT	110	3	Humlv318	23	22%	38
GAR	108	3	DPL23	26	25%	33
32.B9	98	8	DPL21	5	5%	81
PUG	108	3	Humlv318	24	23%	19
T1	115	8	HUMLV801	52	50%	6
RF-TS7	96	7	DPL18	4	4%	60
YM-1	116	8	HUMLV801	51	49%	75
K6H6	112	8	HUMLV801	20	19%	44
K5C7	112	8	HUMLV801	20	19%	44
K5B8	112	8	HUMLV801	20	19%	44
K5G5	112	8	HUMLV801	20	19%	44
K4B8	112	8	HUMLV801	19	18%	44
K6F5	112	8	HUMLV801	17	16%	44
HIL	108	3	DPL23	22	21%	47
KIR	109	3	DPL23	20	19%	19
CAP	109	3	DPL23	19	18%	84
1B8	110	3	DPL23	22	21%	43
SHO	108	3	DPL23	19	18%	19
HAN	108	3	DPL23	20	19%	19
cML23	96	3	DPL23	3	3%	12
PR-SJI	96	3	DPL23	7	7%	55
BAU	107	3	DPL23	9	9%	5
TEX	99	3	DPL23	8	8%	19
X(PET)	107	3	DPL23	9	9%	51
DOY	106	3	DPL23	9	9%	19
COT	106	3	DPL23	13	12%	19
Pag-1	111	3	Humlv318	5	5%	31
DIS	107	3	Humlv318	2	2%	19
WIT	108	3	Humlv318	7	7%	19
1.RH	108	3	Humlv318	12	11%	19
S1-1	108	3	Humlv318	12	11%	52
DEL	108	3	Humlv318	14	13%	17
TYR	108	3	Humlv318	11	10%	19
J.RH	109	3	Humlv318	13	12%	19
THO	112	2	DPL13	38	36%	26
LBV	113	1	DPL3	38	36%	2
WLT	112	1	DPL3	33	31%	14
SUT	112	2	DPL12	37	35%	65

Table 2C: rearranged human heavy chain sequences

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
21/28	119	1	VH1-13-12	0	0,0%	31
8E10	123	1	VH1-13-12	0	0,0%	31
MUC1-1	118	1	VH1-13-6	4	4,1%	42
gF1	98	1	VH1-13-12	10	10,2%	75
VHGL 1.2	98	1	VH1-13-6	2	2,0%	26
HVIL1	98	1	VH1-13-6	0	0,0%	81
RF-TS7	104	1	VH1-13-6	3	3,1%	96
E55 1.A15	106	1	VH1-13-15	1	1,0%	26
HA1L1	126	1	VH1-13-6	7	7,1%	81
UC	123	1	VH1-13-6	5	5,1%	115
WIL2	123	1	VH1-13-6	6	6,1%	55
R3.5H5G	122	1	VH1-13-6	10	10,2%	70
N89P2	123	1	VH1-13-16	11	11,2%	77
mAb113	126	1	VH1-13-6	10	10,2%	71
LS2S3-3	125	1	VH1-12-7	5	5,1%	98
LS2S3-12a	125	1	VH1-12-7	5	5,1%	98
LS2S3-5	125	1	VH1-12-7	5	5,1%	98
LS2S3-12e	125	1	VH1-12-7	5	5,1%	98
LS2S3-4	125	1	VH1-12-7	5	5,1%	98
LS2S3-10	125	1	VH1-12-7	5	5,1%	98
LS2S3-12d	125	1	VH1-12-7	6	6,1%	98
LS2S3-8	125	1	VH1-12-7	5	5,1%	98
LS2	125	1	VH1-12-7	6	6,1%	113
LS4	105	1	VH1-12-7	6	6,1%	113
LS5	125	1	VH1-12-7	6	6,1%	113
LS1	125	1	VH1-12-7	6	6,1%	113
LS6	125	1	VH1-12-7	6	6,1%	113
LS8	125	1	VH1-12-7	7	7,1%	113
THY-29	122	1	VHI-12-7	0	0,0%	42
1B9/F2	122	1	VH1-12-7	10	10,2%	21
51P1	122	1	VH1-12-1	0	0,0%	105
NEI	127	1	VH1-12-1	0	0,0%	55
AND	127	1	VH1-12-1	0	0,0%	55
L7	127	1	VH1-12-1	0	0,0%	54
L22	124	1	VH1-12-1	0	0.0%	54
L24	127	1	VH1-12-1	0	0,0%	54
L26	116	1	VH1-12-1	0	0,0%	54
L33	119	1	VH1-12-1	0	0,0%	54
L34	117	1	VH1-12-1	0	0,0%	54
L36	118	1	VH1-12-1	0	0,0%	54
L39	120	1	VH1-12-1	0	0,0%	54
L41	120	1	VH1-12-1	0	0,0%	54
L42	125	1	VH1-12-1	0	0.0%	54
VHGL 1.8	101	1	VH1-12-1	0	0,0%	26
783c	127	1	VH1-12-1	0	0,0%	22
X17115	127	1	VH1-12-1	0	0,0%	37
L25	124	1	VH1-12-1	0	0,0%	54
L17	120	1	VH1-12-1	1	1,0%	54

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
L30	127	1	VH1-12-1	1	1,0%	54
L37	120	1	VHI-12-1	1	1.0%	54
TNF-E7	116	1	VHI-12-1	2	2,0%	42
mAb111	122	1	VHI-12-1	7	7,1%	71
III-2R	122	1	VHI-12-9	3	3,1%	70
KAS	121	1	VHI-12-1	7	7,1%	79
YES8c	122	1	VH1-12-1	8	8,2%	34
RF-TS1	123	1	VHI-12-1	8	8,2%	82
BOR'	121	1	VH1-12-8	7	7,1%	79
VHGL 1.9	101	1	VH1-12-1	8	8,2%	26
mAb410.30F305	117	1	VH1-12-9	5	5,1%	52
EV1-15	127	1	VH1-12-8	10	10,2%	78
mAb112	122	1	VH1-12-1	11	11,2%	71
EU	117	1	VH1-12-1	11	11,2%	28
H210	127	1	VH1-12-1	12	12,2%	66
TRANSGENE	104	1	VH1-12-1	0	0,0%	111
CLL2-1	93	1	VH1-12-1	0	0,0%	30
CLL10 13-3	97	1	VH1-12-1	0	0,0%	29
LS7	99	1	VH1-12-7	4	4,1%	113
ALL7-1	87	1	VH1-12-7	0	0,0%	30
CUL3-1	91	1	VHI-12-7	1	1,0%	30
ALL56-1 1	85	1	VHI-13-8	0	0,0%	30
ALL1-1	87	1	VHI-13-6	1	1,0%	30
ALL4-1	94	1	VHI-13-8	0	0,0%	30
ALL56 15-4	85	1	VHI-13-8	5	5,1%	29
CLL4-1	88	1	VHI-13-1	1	1,0%	30
Au92.1	98	1	VHI-12-5	0	0,0%	49
RF-TS3	120	1	VHI-12-5	1	1,0%	82
Au4.1	98	1	VHI-12-5	1	1,0%	49
HP1	121	1	VHI-13-6	13	13,3%	110
BLI	127	1	VH1-13-15	5	5,1%	72
No.13	127	1	VHI-12-2	19	19,4%	76
TR1.23	122	1	VH1-13-2	23	23,5%	88
S1-1	125	1	VH1-12-2	18	18,4%	76
TR1.10	119	1	VH1-13-12	14	14,3%	88
E55 1.A2	102	1	VH1-13-15	3	3,1%	26
SP2	119	1	VH1-13-6	15	15.3%	89
TNF-H9G1	111	1	VH1-13-18	2	2,0%	42
G3D10H	127	1	VH1-13-16	19	19,4%	127
TR1.9	118	1	VH1-13-12	14	14,3%	88
TR1.8	121	1	VH1-12-1	24	24,5%	88
LUNm01	127	1	VH1-13-6	22	22,4%	9
K1B12H	127	1	VH1-12-7	23	23,5%	127
L3B2	99	1	VH1-13-6	2	2,0%	46
ss2	100	1	VH1-13-6	2	2,0%	46
No.86	124	1	VH1-12-1	20	20,4%	76
TR1.6	124	1	VH1-12-1	19	19,4%	88
ss7	99	1	VH1-12-7	3	3,1%	46
s5B7	102	1	VH1-12-1	0	0,0%	46

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
s6A3	97	1	VHI-12-1	0	0,0%	46
ss6	99	1	VH1-12-1	0	0,0%	46
L2H7	103	1	VH1-13-12	0	0,0%	46
s6BG8	93	1	VHI-13-12	0	0,0%	46
s6C9	107	1	VH1-13-12	0	0,0%	46
HIV-b4	124	1	VH1-13-12	21	21,4°/a	12
HIV-b12	124	1	VH1-13-12	21	21,4%	12
L3G5	98	1	VH1-13-6	1	1,0%	46
22	115	1	VH1-13-6	11	11,2%	118
L2A12	99	1	VH1-13-15	3	3,1%	46
PHOX15	124	1	VH1-12-7	20	20,4%	73
LUNm03	127	1	VH1-IX-1	18	18,4%	9
CEA4-8A	129	1	VH1-12-7	1	1,0%	42
M60	121	2	VH2-31-3	3	3,0%	103
HiH10	127	2	VH2-31-5	9	9,0%	4
COR	119	2	VH2-31-2	11	11,0%	91
2-115-19	124	2	VH2-31-11	8	8,1%	124
OU	125	2	VH2-31-14	20	25,6%	92
HE	120	2	VH2-31-13	19	19,0%	27
CLL33 40-1	78	2	VH2-31-5	2	2,0%	29
E55 3.9	88	3	VH3-11-5	7	7,2%	26
MTFC3	125	3	VH3-14-4	21	21,0%	131
MTFC11	125	3	VH3-14-4	21	21,0%	131
MTFJ1	114	3	VH3-14-4	21	21,0%	131
MTFJ2	114	3	VH3-14-4	21	21,0%	131
MTFUJ4	100	3	VH3-14-4	21	21,0%	131
MTFUJ5	100	3	VH3-14-4	21	21,0%	131
MTFUJ2	100	3	VH3-14-4	22	22,0%	131
MTFC8	125	3	VH3-14-4	23	23,0%	131
TD e Vq	113	3	VH3-14-4	0	0,0%	16
rMTF	114	3	VH3-14-4	5	5,0%	131
MTFUJ6	100	3	VH3-14-4	10	10,0%	131
RF-KES	107	3	VH3-14-4	9	9,0%	85
N51P8	126	3	VH3-14-1	9	9,0%	77
TEI	119	3	VH3-13-8	21	21,4%	20
33.H11	115	3	VH3-13-19	10	10,2%	129
SB1/D8	101	3	VH3-1X-8	14	14,0%	2
38P1	119	3	VH3-11-3	0	0,0%	104
BRO'IGM	119	3	VH3-11-3	13	13,4%	19
NIE	119	3	VH3-13-7	15	15,3%	87
3D6	126	3	VH3-13-26	5	5,1%	35
ZM1-1	112	3	VH3-11-3	8	8,2%	5
E553.15	110	3	VH3-13-26	0	0,0%	26
gF9	108	3	VH3-13-8	15	15.3%	75
THY-32	120	3	VH3-13-26	3	3,1%	42
RF-KL5	100	3	VH3-13-26	5	5,1%	96
OST577	122	3	VH3-13-13	6	6,1%	5
BO	113	3	VH3-13-19	15	15,3%	10
TT125	121	3	VH3-13-10	15	15,3%	64

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
2-115-58	127	3	VH3-13-10	11	11,2%	124
KOL	126	3	VH3-13-14	16	16,3%	102
mAb60	118	3	VH3-13-17	14	14,3%	45
RF-AN	106	3	VH3-13-26	8	8,2%	85
BUT	115	3	VH3-11-6	13	13,4%	119
KOL-based						
CAMPATH-9	118	3	VH3-13-13	16	16,3%	41
B1	119	3	VH3-13-19	13	13,3%	53
N98P1	127	3	VH3-13-1	13	13,3%	77
TT117	107	3	VH3-13-10	12	12,2%	64
WEA	114	3	VH3-13-12	15	15,3%	40
HIL	120	3	VH3-13-14	14	14,3%	23
s5A10	97	3	VH3-13-14	0	0,0%	46
s5D11	98	3	VH3-13-7	0	0,0%	46
s6C8	100	3	VH3-13-7	0	0,0%	46
s6H12	98	3	VH3-13-7	0	0,0%	46
VH10.7	119	3	VH3-13-14	16	16,3%	128
HIV-loop2	126	3	VH3-13-7	16	16,3%	12
HIV-loop35	126	3	VH3-13-7	16	16,3%	12
TRO	122	3	VH3-13-1	13	13,3%	61
SA-4B	123	3	VH3-13-1	15	15,3%	125
L2B5	98	3	VH3-13-13	0	0,0%	46
s6E11	95	3	VH3-13-13	0	0,0%	46
s6H7	100	3	VH3-13-13	0	0,0%	46
ss1	102	3	VH3-13-13	0	0,0%	46
ss8	94	3	VH3-13-13	0	0,0%	46
DOB	120	3	VH3-13-26	21	21,4%	116
THY-33	115	3	VH3-13-15	20	20,4%	42
NOV	118	3	VH3-13-19	14	14,3%	38
rsv13H	120	3	VH3-13-24	20	20,4%	11
L3G11	98	3	VH3-13-20	2	2,0%	46
L2E8	99	3	VH3-13-19	0	0,0%	46
L2D10	101	3	VH3-13-10	1	1,0%	46
L2E7	98	3	VH3-13-10	1	1,0%	46
L3A10	100	3	VH3-13-24	0	0,0%	46
L2E5	97	3	VH3-13-2	1	1,0%	46
BUR	119	3	VH3-13-7	21	21,4%	67
s4D5	107	3	VH3-11-3	1	1,0%	46
19	116	3	VH3-13-16	4	4,1%	118
s5D4	99	3	VH3-13-1	0	0,0%	46
s6A8	100	3	VH3-13-1	0	0.0%	46
HIV-loop13	123	3	VH3-13-12	17	17,3%	12
TR1.32	112	3	VH3-11-8	18	18,6%	88
L2B10	97	3	VH3-11-3	1	1,0%	46
TR1.5	114	3	VH3-11-8	21	21,6%	88
s6H9	101	3	VH3-13-25	0	0,0%	46
8	112	3	VH3-13-1	6	6,1%	118
23	115	3	VH3-13-1	6	6,1%	118
7	115	3	VH3-13-1	4	4,1%	118 8

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
TRI.3	120	3	VH3-11-8	20	20.6%	88
18/2	125	3	VH3-13-10	0	0.0%	32
18/9	125	3	VH3-13-10	0	0,0%	31
30P1	119	3	VH3-13-10	0	0,0%	106
HF2-1/17	125	3	VH3-13-10	0	0,0%	8
A77	109	3	VH3-13-10	0	0,0%	44
B19.7	108	3	VH3-13-10	0	0,0%	44
M43	119	3	VH3-13-10	0	0,0%	103
1/17	125	3	VH3-13-10	0	0,0%	31
18/17	125	3	VH3-13-10	0	0,0%	31
E54 3.4	109	3	VH3-13-10	0	0,0%	26
LAMBDA-VH26	98	3	VH3-13-10	1	1,0%	95
E54 3.8	111	3	VH3-13-10	1	1,0%	26
GL16	106	3	VH3-13-10	1	1,0%	44
4G12	125	3	VH3-13-10	1	1,0%	56
A73	106	3	VH3-13-10	2	2,0%	44
AL1.3	111	3	VH3-13-10	3	3,1%	117
3.A290	118	3	VH3-13-10	2	2,0%	108
Ab18	127	3	VH3-13-8	2	2,0%	100
E54 3.3	105	3	VH3-13-10	3	3,1%	26
35G6	121	3	VH3-13-10	3	3,1%	57
A95	107	3	VH3-13-10	5	5,1%	44
Ab25	128	3	VH3-13-10	5	5,1%	100
N87	126	3	VH3-13-10	4	4,1%	77
ED8.4	99	3	VH3-13-10	6	6,1%	2
RF-KL1	122	3	VH3-13-10	6	6,1%	82
AL1.1	112	3	VH3-13-10	2	2,0%	117
AL3.11	102	3	VH3-13-10	1	1,0%	117
32.B9	127	3	VH3-13-8	6	6,1%	129
TK1	109	3	VH3-13-10	2	2,0%	117
POP	123	3	VH3-13-10	8	8,2%	115
9F2H	127	3	VH3-13-10	9	9,2%	127
VD	115	3	VH3-13-10	9	9,2%	10
Vh38C1.10	121	3	VH3-13-10	8	8,2%	74
Vh38Cl.9	121	3	VH3-13-10	8	8,2%	74
Vh38C1.8	121	3	VH3-13-10	8	8,2%	74
63P1	120	3	VH3-11-8	0	0,0%	104
60P2	117	3	VH3-11-8	0	0,0%	104
AL3.5	90	3	VH3-13-10	2	2,0%	117
GF4/1.1	123	3	VH3-13-10	10	10,2%	39
Ab21	126	3	VH3-13-10	12	12,2%	100
TD d Vp	118	3	VH3-13-17	2	2,0%	16
Vh38C1.4	119	3	VH3-13-10	8	8,2%	74
Vh38C1.5	119	3	VH3-13-10	8	8,2%	74
AL3.4	104	3	VH3-13-10	1	1,0%	117
FOG1-A3	115	3	VH3-13-19	2	2.0%	42
HA3DI	117	3	VH3-13-21	1	1,0%	81
E54 3.2	112	3	VH3-13-24	0	0,0%	26
mAb52	128	3	VH3-13-12	2	2,0%	51

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
mAb53	128	3	VH3-13-12	2	2,0%	51
mAb56	128	3	VH3-13-12	2	2,0%	51
mAb57	128	3	VH3-13-12	2	2,0%	51
mAb58	128	3	VH3-13-12	2	2,0%	51
mAb59	128	3	VH3-13-12	2	2,0%	51
mAb105	128	3	VH3-13-12	2	2,0%	51
mAb107	128	3	VH3-13-12	2	2,0%	51
E55 3.14	110	3	VH3-13-19	0	0,0%	26
F13-28	106	3	VH3-13-19	1	1,0%	94
mAb55	127	3	VH3-13-18	4	4,1%	51
YSE	117	3	VH3-13-24	6	6,1%	72
E553.23	106	3	VH3-13-19	2	2,0%	26
RF-TS5	101	3	VH3-13-1	3	3,1%	85
N42P5	124	3	VH3-13-2	7	7,1%	77
FOG1-H6	110	3	VH3-13-16	7	7,1%	42
O-81	115	3	VH3-13-19	11	11,2%	47
HIV-s8	122	3	VH3-13-12	11	11,2%	12
mAb114	125	3	VH3-13-19	12	12,2%	71
33.F12	116	3	VH3-13-2	4	4,1%	129
4B4	119	3	VH3-1X-3	0	0,0%	101
M26	123	3	VN3-1X-3	0	0,0%	103
VHGL 3.1	100	3	VH3-1X-3	0	0,0%	26
E553.13	113	3	VH3-1X-3	1	1,0%	26
SB5/D6	101	3	VH3-1X-6	3	3,0%	2
RAY4	101	3	VH3-1X-6	3	3,0%	2
82-D V-D	106	3	VH3-1X-3	5	5,0%	112
MAL	129	3	VH3-1X-3	5	5,0%	72
LOC	123	3	VH3-1X-6	5	5,0%	72
LSF2	101	3	VH3-IX-6	11	11,0%	2
HIB RC3	100	3	VH3-1X-6	11	11,0%	1
56P 1	119	3	VH3-13-7	0	0,0%	104
M72	122	3	VH3-13-7	0	0,0%	103
M74	121	3	VH3-13-7	0	0,0%	103
E54 3.5	105	3	VH3-13-7	0	0,0%	26
2E7	123	3	VH3-13-7	0	0,0%	63
2P1	117	3	VH3-13-7	0	0,0%	104
RF-SJ2	127	3	VH3-13-7	1	1,0%	83
PR-TS1	114	3	VH3-13-7	1	1,0%	85
KIM46H	127	3	VH3-13-13	0	0,0%	18
E55 3.6	108	3	VH3-13-7	2	2,0%	26
E55 3.10	107	3	VH3-13-13	1	1,0%	26
3.B6	114	3	VH3-13-13	1	1,0%	108
E54 3.6	110	3	VH3-13-13	1	1.0%	26
FL2-2	114	3	VH3-13-13	1	1.0%	80
RF-SJ3	112	3	VH3-13-7	2	2,0%	85
E55 3.5	105	3	VH3-13-14	1	1,0%	26
BSA3	121	3	VH3-13-13	1	1,0%	73
HMST-1	119	3	VH3-13-7	3	3,1%	130
RF-TS2	126	3	VH3-13-13	4	4,1%	82

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
E55 3.12	109	3	VH3-13-15	0	0,0%	26
19.E7	126	3	VH3-13-14	3	3,1%	129
11-50	119	3	VH3-13-13	6	6.1%	130
E29.1	120	3	VH3-13-15	2	2.0%	25
E55 3.16	108	3	VH3-13-7	6	6,1%	26
TNF-E1	117	3	VH3-13-7	7	7,1%	42
RF-SJ1	127	3	VH3-13-13	6	6,1%	83
FOG1-A4	116	3	VH3-13-7	8	8,2%	42
TNF-A1	117	3	VH3-13-15	4	4,1%	42
PR-SJ2	107	3	VH3-13-14	8	8,2%	85
HN.14	124	3	VH3-13-13	10	10,2%	33
CAM'	121	3	VH3-13-7	12	12,2%	65
HIV-B8	125	3	VH3-13-7	9	9,2%	12
HIV-b27	125	3	VH3-13-7	9	9,2%	12
HIV-b8	125	3	VH3-13-7	9	9,2%	12
HIV-s4	125	3	VH3-13-7	9	9,2%	12
HIV-B26	125	3	VH3-13-7	9	9,2%	12
HIV-B35	125	3	VH3-13-7	10	10,2%	12
HIV-bl8	125	3	VH3-13-7	10	10,2%	12
HIV-b22	125	3	VH3-13-7	11	11,2%	12
HIV-b13	125	3	VH3-13-7	12	12,2%	12
333	117	3	VH3-14-4	24	24,0%	24
1H1	120	3	VH3-14-4	24	24,0%	24
1B11	120	3	VH3-14-4	23	23,0%	24
CLL30 2-3	86	3	VH3-13-19	1	1,0%	29
GA	110	3	VH3-13-7	19	19,4%	36
JeB	99	3	VH3-13-14	3	3,1%	7
GAL	110	3	VH3-13-19	10	10,2%	126
K6H6	119	3	VH3-1X-6	18	18,0%	60
K4B8	119	3	VH3-1X-6	18	18,0%	60
K5B8	119	3	VH3-1X-6	18	18,0%	60
K5C7	119	3	VH3-1X-6	19	19,0%	60
K5G5	119	3	VH3-1X-6	19	19,0%	60
K6F5	119	3	VH3-1X-6	19	19,0%	60
AL3.16	98	3	VH3-13-10	1	1,0%	117
N86P2	98	3	VH3-13-10	3	3,1%	77
N54P6	95	3	VH3-13-16	7	7,1%	77
LAMBDA HT112-1	126	4	VH4-11-2	0	0,0%	3
HY18	121	4	VH4-11-2	0	0,0%	43
mAb63	126	4	VH4-11-2	0	0,0%	45
FS-3	105	4	VH4-11-2	0	0,0%	86
FS-5	111	4	VH4-11-2	0	0,0%	86
FS-7	107	4	VH4-11-2	0	0,0%	86
FS-8	110	4	VH4-11-2	0	0,0%	86
PR-TS2	105	4	VH4-11-2	0	0,0%	85
RF-TMC	102	4	VH4-11-2	0	0,0%	85
mAb216	122	4	VH4-11-2	1	1,0%	15
mAb410.7.F91	122	4	VH4-11-2	1	1,0%	52
mAbA6H4C5	124	4	VH4-11-2	1	1,0%	15

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
Ab44	127	4	VH4-11-2	2	2,1%	100
6H-3C4	124	4	VH4-11-2	3	3,1%	59
FS-6	108	4	VH4-11-2	6	6,2%	86
FS-2	114	4	VH4-11-2	6	6,2%	84
HIG1	126	4	VH4-11-2	7	7,2%	62
FS-4	105	4	VH4-11-2	8	8,2%	86
SA-4A	123	4	VH4-11-2	9	9,3%	125
LES-C	119	4	VH4-11-2	10	10,3%	99
DI	78	4	VH4-11-9	16	16,5%	58
Ab26	126	4	VH4-31-4	8	8,1%	100
TS2	124	4	VH4-31-12	15	15,2%	110
265-695	115	4	VH4-11-7	16	16,5%	5
WAH	129	4	VH4-31-13	19	19,2%	93
268-D	122	4	VH4-11-8	22	22,7%	6
58P2	118	4	VH4-11-8	0	0,0%	104
mAb67	128	4	VH4-21-4	1	1,0%	45
4.L39	115	4	VH4-11-8	2	2,1%	108
mF7	111	4	VH4-31-13	3	3,0%	75
33.C9	122	4	VH4-21-5	7	7,1%	129
Pag-1	124	4	VH4-11-16	5	5,2%	50
B3	123	4	VH4-21-3	8	8,2%	53
IC4	120	4	VH4-11-8	6	6,2%	70
C6B2	127	4	VH4-31-12	4	4,0%	48
N78	118	4	VH4-11-9	11	11,3%	77
B2	109	4	VH4-11-8	12	12,4%	53
WRD2	123	4	VH4-11-12	6	6,2%	90
mAb426.4.2F20	126	4	VH4-11-8	2	2,1%	52
E54 4.58	115	4	VH4-11-8	1	1,0%	26
WRD6	123	4	VH4-11-12	10	10,3%	90
mAb426.12.3F1.4	122	4	VH4-11-9	4	4,1%	52
E54 4.2	108	4	VH4-21-6	2	2,0%	26
W1L	127	4	VH4-31-13	0	0,0%	90
COF	126	4	VH4-31-13	0	0,0%	90
LAR	122	4	VH4-31-13	2	2,0%	90
WAT	125	4	VH4-31-13	4	4,0%	90
mAb61	123	4	VH4-31-13	5	5,1%	45
WAG	127	4	VH4-31-4	0	0,0%	90
RF-SJ4	108	4	VH4-31-12	2	2,0%	85
E54 4.4	110	4	VH4-11-7	0	0,0%	26
E55 4.A1	108	4	VH4-11-7	0	0,0%	26
PR-SJ1	103	4	VH4-11-7	1	1,0%	85
E54 4.23	111	4	VH4-11-7	1	1,0%	26
CLL77-2	97	4	VH4-11-12	0	0,0%	29
37P1	95	4	VH4-11-12	0	0,0%	104
ALL52 30-2	91	4	VH4-31-12	4	4,0%	29
EBV-21	98	5	VHS-12-1	0	0,0%	13
CB-4	98	5	VH5-12-1	0	0,0%	13
CLL-12	98	5	VHS-12-1	0	0,0%	13
L3-4	98	5	VH5-12-1	0	0,0%	13

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
CLL11	98	5	VHS-12-1	0	0,0%	17
CORD3	98	5	VHS-12-1	0	0.0%	17
CORD4	98	5	VHS-12-1	0	0,0%	17
CORD8	98	5	VHS-12-1	0	0,0%	17
CORD9	98	5	VH5-12-1	0	0,0%	17
CD+1	98	5	VHS-12-1	0	0,0%	17
CD+3	98	5	VHS-12-1	0	0,0%	17
CD+4	98	5	VH5-12-1	0	0,0%	17
CD-1	98	5	VH5-12-1	0	0,0%	17
CD-5	98	5	VH5-12-1	0	0,0%	17
VERG14	98	5	VH5-12-1	0	0,0%	17
PBL1	98	5	VH5-12-1	0	0,0%	17
PBL10	98	5	VH5-12-1	0	0,0%	17
STRAb SA-1A	127	5	VH5-12-1	0	0,0%	125
DOB'	122	5	VH5-12-1	0	0,0%	97
VERG5	98	5	VH5-12-1	0	0,0%	17
PBL2	98	5	VH5-12-1	1	1.0%	17
Tu16	119	5	VH5-12-1	1	1,0%	49
PBL12	98	5	VH5-12-1	1	1,0%	17
CD+2	98	5	VH5-12-1	1	1,0%	17
CORD10	98	5	VH5-12-1	1	1,0%	17
PBL9	98	5	VH5-12-1	1	1,0%	17
CORD2	98	5	VH5-12-1	2	2,0%	17
PBL6	98	5	VH5-12-1	2	2,0%	17
CORD5	98	5	VH5-12-1	2	2,0%	17
CD-2	98	5	VH5-12-1	2	2,0%	17
CORD1	98	5	VH5-12-1	2	2,0%	17
CD-3	98	5	VH5-12-1	3	3,1%	17
VERG4	98	5	VH5-12-1	3	3,1%	17
PBL13	98	5	VH5-12-1	3	3,1%	17
PBL7	98	5	VH5-12-1	3	3,1%	17
HAN	119	5	VH5-12-1	3	3,1%	97
VERG3	98	5	VH5-12-1	3	3,1%	17
PBL3	98	5	VH5-12-1	3	3,1%	17
VERG7	98	5	VH5-12-1	3	3,1%	17
PBL5	94	5	VH5-12-1	0	0,0%	17
CD-4	98	5	VH5-12-1	4	4,1%	17
CLL10	98	5	VH5-12-1	4	4,1%	17
PBL11	98	5	VH5-12-1	4	4,1%	17
CORD6	98	5	VHS-12-1	4	4,1%	17
VERG2	98	5	VHS-12-1	5	5,1%	17
83P2	119	5	VH5-12-1	0	0,0%	103
VERG9	98	5	VH5-12-1	6	6,1%	17
CLL6	98	5	VHS-12-1	6	6,1%	17
PBL8	98	5	VH5-12-1	7	7,1%	17
Ab2022	120	5	VHS-12-1	3	3.1%	100
CAV	127	5	VHS-12-4	0	0,0%	97
HOW'	120	5	VH5-12-4	0	0,0%	97
PET	127	5	VHS-12-4	0	0.0%	97

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
ANG	121	5	VH5-12-4	0	0,0%	97
KER	121	5	VH5-12-4	0	0.0%	97
5.M13	118	5	VH5-12-4	0	0,0%	107
Au2.1	118	5	VH5-12-4	1	1,0%	49
WS1	126	5	VH5-12-1	9	9.2%	110
TD Vn	98	5	VHS-12-4	1	1.0%	16
TEL13	116	5	VH5-12-1	9	9.2%	73
E55 5.237	112	5	VH5-12-4	2	20%	26
VERG1	98	5	VH5-12-1	10	10,2%	17
CD4-74	117	5	VH5-12-1	10	10,2%	42
257-D	125	5	VHS-12-1	11	11,2%	6
CLL4	98	5	VH5-12-1	11	11,2%	17
CLL8	98	5	VH5-12-1	11	11,2%	17
Ab2	124	5	VHS-12-1	12	12,2%	120
Vh383ex	98	5	VHS-12-1	12	12,2%	120
CLL3	98	5	VHS-12-2	11	11,2%	17
Au59.1	122	5	VH5-12-1	12	12,2%	49
TEL16	117	5	VH5-12-1	12	12,2%	73
M61	104	5	VH5-12-1	0	0,0%	103
Tu0	99	5	VH5-12-1	5	5,1%	49
P2-51	122	5	VHS-12-1	13	13,3%	121
P2-54	122	5	VHS-12-1	11	11,2%	121
P1-56	119	5	VH5-12-1	9	9,2%	121
P2-53	122	5	VH5-12-1	10	10,2%	121
P1-51	123	5	VH5-12-1	19	19,4%	121
P1-54	123	5	VH5-12-1	3	3,1%	121
P3-69	127	5	VH5-12-1	4	4,1%	121
P3-9	119	5	VH5-12-1	4	4,1%	121
1-185-37	125	5	VH5-12-4	0	0,0%	124
1-187-29	125	5	VH5-12-4	0	0,0%	124
P1-58	128	5	VHS-12-4	10	10,2%	121
P2-57	118	5	VHS-12-4	3	3,1%	121
P2-55	123	5	VHS-12-1	5	5,1%	121
P2-56	123	5	VH5-12-1	20	20,4%	121
P2-52	122	5	VHS-12-1	11	11,2%	121
P3-60	122	5	VHS-12-1	8	8,2%	121
P1-57	123	5	VH5-12-1	4	4,1%	121
P1-55	122	5	VH5-12-1	14	14,3%	121
MD3-4	128	5	VHS-12-4	12	12,2%	5
P1-52	121	5	VHS-12-1	11	11,2%	121
CLL5	98	5	VH5-12-1	13	13,3%	17
CLL7	98	5	VH5-12-1	14	14.3%	17
L2F10	100	5	VH5-12-1	1	1,0%	46
L3B6	98	5	VH5-12-1	1	1,0%	46
VH6.A12	119	6	VH6-35-1	13	12,9%	122
s5A9	102	6	VH6-35-1	1	1,0%	46
s6G4	99	6	VH6-35-1	1	1,0%	46
ss3	99	6	VH6-35-1	1	1,0%	46
6-IG1	101	6	VH6-35-1	0	0,0%	14

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
F19L16	107	6	VH6-35-1	0	0,0%	68
L16	120	6	VH6-35-1	0	0,0%	69
M71	121	6	VH6-35-1	0	0,0%	103
ML1	120	6	VH6-35-1	0	0,0%	69
F19ML1	107	6	VH6-35-1	0	0,0%	68
15P1	127	6	VH6-35-1	0	0,0%	104
VH6.N1	121	6	VH6-35-1	0	0,0%	122
VH6.N11	123	6	VH6-35-1	0	0,0%	122
VH6.N12	123	6	VH6-35-1	0	0,0%	122
VH6.N2	125	6	VH6-35-1	0	0,0%	122
VH6.N5	125	6	VH6-35-1	0	0,0%	122
VH6.N6	127	6	VH6-35-1	0	0,0%	122
VH6.N7	126	6	VH6-35-1	0	0,0%	122
VH6.N8	123	6	VH6-35-1	0	0,0%	122
VH6.N9	123	6	VH6-35-1	0	0,0%	122
VH6.N10	123	6	VH6-35-1	0	0,0%	122
VH6.A3	123	6	VH6-35-1	0	0,0%	122
VH6.A 1	124	6	VH6-35-1	0	0,0%	122
VH6.A4	120	6	VH6-35-1	0	0,0%	122
E55 6.16	116	6	VH6-35-1	0	0,0%	26
E55 6.17	120	6	VH6-35-1	0	0,0%	26
E556.6	120	6	VH6-35-1	0	0,0%	26
VHGL 6.3	102	6	VH6-35-1	0	0,0%	26
CB-201	118	6	VH6-35-1	0	0,0%	109
VH6.N4	122	6	VH6-35-1	0	0,0%	122
E54 6.4	109	6	VH6-35-1	1	1,0%	26
VH6.A6	126	6	VH6-35-1	1	1,0%	122
E556.14	120	6	VH6-35-1	1	1,0%	26
E546.6	107	6	VH6-35-1	1	1,0%	26
E556.10	112	6	VH6-35-1	1	1,0%	26
E546.1	107	6	VH6-35-1	2	2,0%	26
E55 6.13	120	6	VH6-35-1	2	2,0%	26
E55 6.3	120	6	VH6-35-1	2	2,0%	26
E55 6.7	116	6	VH6-35-1	2	2,0%	26
E55 6.2	120	6	VH6-35-1	2	2,0%	26
E55 6.X	111	6	VH6-35-1	2	2,0%	26
E55 6.11	111	6	VH6-35-1	3	3,0%	26
VH6.A11	118	6	VH6-35-1	3	3,0%	122
A10	107	6	VH6-35-1	3	3,0%	68
E55 6.1	120	6	VH6-35-1	4	4,0%	26
FK-001	124	6	VH6-35-1	4	4,0%	65
VH6.A5	121	6	VH6-35-1	4	4.0%	122
VH6.A7	123	6	VH6-35-1	4	4,0%	122
HBp2	119	6	VH6-35-1	4	4,0%	4
Au46.2	123	6	VH6-35-1	5	5,0%	49
A431	106	6	VH6-35-1	5	5,0%	68
VH6.A2	120	6	VH6-35-1	5	5.0%	122
VH6.A9	125	6	VH6-35-1	8	7,9%	122
VH6.A8	118	6	VH6-35-1	10	9,9%	122

(continued)

Name[1]	aa[2]	Computed family[3]	Germline gene[4]	Diff. to germline[5]	% diff. to germline[6]	Reference[7]
VH6-FF3	118	6	VH6-35-1	2	2,0%	123
VH6.A10	126	6	VH6-35-1	12	11,9%	122
VH6-EB10	117	6	VH6-35-1	3	3,0%	123
VH6-E6	119	6	VH6-35-1	6	5,9%	123
VH6-FE2	121	6	VH6-35-1	6	5,9%	123
VH6-EE6	116	6	VH6-35-1	6	5,9%	123
VH6-FD10	118	6	VH6-35-1	6	5,9%	123
VH6-EX8	113	6	VH6-35-1	6	5,9%	123
VH6-FG9	121	6	VH6-35-1	8	7.9%	123
VH6-E5	116	6	VH6-35-1	9	8,9%	123
VEI6-EC8	122	6	VH6-35-1	9	8.9%	123
VH6-E10	120	6	VH6-35-1	10	9,9%	123
VHG-FF11	122	6	VH6-35-1	11	10,9%	123
VH6-FD2	115	6	VH6-35-1	11	10.9%	123
CLL10 17-2	88	6	VH6-35-1	4	4,0%	29
VHG-BB11	94	6	VH6-35-1	4	4,0%	123
VH6-B41	93	6	VH6-35-1	7	6,9%	123
JU17	102	6	VHG-35-1	3	3.0%	114
VH6-BD9	96	6	VH6-35-1	11	10,9%	123
VH6-BB9	94	6	VH6-35-1	12	11,9%	123

Table 3A: assignment of rearranged V kappa sequences to their germline counterparts

Family[1]	Name	Rearranged[2]	Sum
1	Vk1-1	28	
1	Vk1-2	0	
1	Vk1-3	1	
1	Vk1-4	0	
1	Vk1-5	7	
1	Vk1-6	0	
1	Vk1-7	0	
1	Vk1-8	2	
1	Vk1-9	9	
1	Vk1-10	0	
1	Vk1-11	1	
1	Vk1-12	7	
1	Vk1-13	1	
1	Vk1-14	7	
1	Vk1-15	2	
1	Vk1-16	2	
1	Vk1-17	16	
1	Vk1-18	1	
1	Vk1-19	33	
1	Vk1-20	1	
1	Vk1-21	1	
1	Vk1-22	0	
1	Vk1-23	0	***119 entries***
2	Vk2-1	0	

(continued)

Family[1]	Name	Rearranged[2]	Sum
2	Vk2-2	1	
2	Vk2-3	0	
2	Vk2-4	0	
2	Vk2-5	0	
2	Vk2-6	16	
2	Vk2-7	0	
2	Vk2-8	0	
2	Vk2-9	1	
2	Vk2-10	0	
2	Vk2-11	7	
2	Vk2-12	0	***25 entries***
3	Vk3-1	1	
3	Vk3-2	0	
3	Vk3-3	35	
3	Vk3-4	115	
3	Vk3-5	0	
3	Vk3-6	0	
3	Vk3-7	1	
3	Vk3-8	40	***192 entries***
4	Vk4-1	33	***33 entries***
5	Vk5-1	1	***1 entry***
6	Vk6-1	0	
6	Vk6-2	0	***0 entries***
7	Vk7-1	0	***0 entries***

Table 3B: assignment of rearranged V lambda sequences to their germline counterparts

Family[1]	Name	Rearranged[2]	Sum
1	DPL1	1	
1	DPL2	14	
1	DPL3	6	
1	DPL4	1	
1	HUMLV117	4	
1	DPL5	13	
1	DPL6	0	
1	DPL7	0	
1	DPL8	3	
1	DPL9	0	***42 entries***
2	DPL10	5	
2	VLAMBDA 2.1	0	
2	DPL1	23	
2	DPL12	15	
2	DPL 13	0	
2	DPL14	0	***43 entries***
3	DPL16	10	
3	DPL23	19	
3	Humlv318	9	***38 entries***

(continued)

Family[1]	Name	Rearranged[2]	Sum
7	DPL18	1	
7	DPL19	0	***1 entries***
8	DPL21	2	
8	HUMLV801	6	***8 entries***
9	DPL22	0	***0 entries***
unassigned	DPL24	0	***0 entries***
10	gVLX-4.4	0	***0 entries***

Table 3C: assignment of rearranged V heavy chain sequences to their germline counterparts

Family[1]	Name	Rearranged[2]	Sum
1	VH1-12-1	38	
1	VH1-12-8	2	
1	VH1-12-2	2	
1	VH1-12-9	2	
1	VH1-12-3	0	
1	VH1-12-4	0	
1	VH1-12-5	3	
1	VH1-12-6	0	
1	VH1-12-7	23	
1	VH1-13-1	1	
1	VH1-13-2	1	
1	VH1-13-3	0	
1	VH1-13-4	0	
1	VH1-13-5	0	
1	VH1-13-6	17	
1	VH1-13-7	0	
1	VH1-13-8	3	
1	VH1-13-9	0	
1	VH1-13-10	0	
1	VH1-13-11	0	
1	VH1-13-12	10	
1	VH1-13-13	0	
1	VH1-13-14	0	
1	VH1-13-15	4	
1	VH1-13-16	2	
1	VH1-13-17	0	
1	VH1-13-18	1	
1	VH1-13-19	0	
1	VH1-1X-1	1	***110 entries***
2	VH2-21-1	0	
2	VH2-31-1	0	
2	VH2-31-2	1	
2	VH2-31-3	1	
2	VH2-31-4	0	
2	VH2-31-5	2	
2	VH2-31-6	0	

(continued)

Family[1]	Name	Rearranged[2]	Sum
2	VH2-31-7	0	
2	VH2-31-14	1	
2	VH2-31-8	0	
2	VH2-31-9	0	
2	VH2-31-10	0	
2	VH2-31-11	1	
2	VH2-31-12	0	
2	VH2-31-13	1	***7 entries***
3	VH3-11-1	0	
3	VH3-11-2	0	
3	VH3-11-3	5	
3	VH3-11-4	0	
3	VH3-11-5	1	
3	VH3-11-6	1	
3	VH3-11-7	0	
3	VH3-11-8	5	
3	VH3-13-1	9	
3	VH3-13-2	3	
3	VH3-13-3	0	
3	VH3-13-4	0	
3	VH3-13-5	0	
3	VH3-13-6	0	
3	VH3-13-7	32	
3	VH3-13-8	4	
3	VH3-13-9	0	
3	VH3-13-10	46	
3	VH3-13-11	0	
3	VH3-13-12	11	
3	VH3-13-13	17	
3	VH3-13-14	8	
3	VH3-13-15	4	
3	VH3-13-16	3	
3	VH3-13-17	2	
3	VH3-13-18	1	
3	VH3-13-19	13	
3	VH3-13-20	1	
3	VH3-13-21	1	
3	VH3-13-22	0	
3	VH3-13-23	0	
3	VH3-13-24	4	
3	VH3-13-25	1	
3	VH3-13-26	6	
3	VH3-14-1	1	
3	VH3-14-4	15	
3	VH3-14-2	0	
3	VH3-14-3	0	
3	VH3-1X-1	0	
3	VH3-1X-2	0	
3	VH3-1X-3	6	

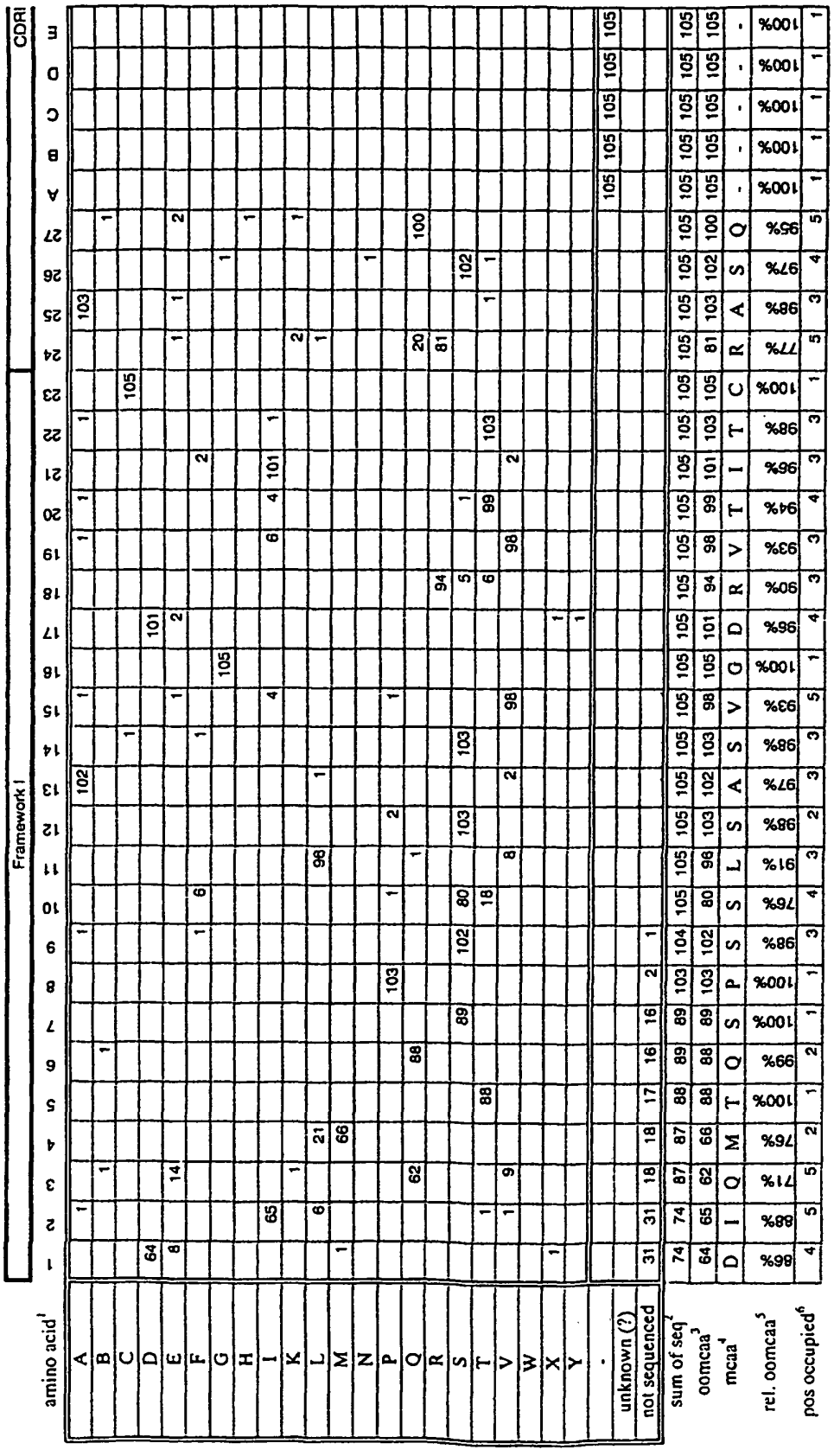
(continued)

Family[1]	Name	Rearranged[2]	Sum
3	VH3-1X-4	0	
3	VH3-1X-5	0	
3	VH3-1X-6	11	
3	VH3-1X-7	0	
3	VH3-1X-8	1	
3	VH3-1X-9	0	***212 entries***
4	VH4-11-1	0	
4	VH4-11-2	20	
4	VH4-11-3	0	
4	VH4-11-4	0	
4	VH4-11-5	0	
4	VH4-11-6	0	
4	VH4-11-7	5	
4	VH4-11-8	7	
4	VH4-11-9	3	
4	VH4-11-10	0	
4	VH4-11-11	0	
4	VH4-11-12	4	
4	VH4-11-13	0	
4	VH4-11-14	0	
4	VH4-11-15	0	
4	VH4-11-16	1	
4	VH4-21-1	0	
4	VH4-21-2	0	
4	VH4-21-3	1	
4	VH4-21-4	1	
4	VH4-21-5	1	
4	VH4-21-6	1	
4	VH4-21-7	0	
4	VH4-21-8	0	
4	VH4-21-9	0	
4	VH4-31-1	0	
4	VH4-31-2	0	
4	VH4-31-3	0	
4	VH4-31-4	2	
4	VH4-31-5	0	
4	VH4-31-6	0	
4	VH4-31-7	0	
4	VH4-31-8	0	
4	VH4-31-9	0	
4	VH4-31-10	0	
4	VH4-31-11	0	
4	VH4-31-12	4	
4	VH4-31-13	7	
4	VH4-31-14	0	
4	VH4-31-15	0	
4	VH4-31-16	0	
4	VH4-31-17	0	
4	VH4-31-18	0	

(continued)

Family[1]	Name	Rearranged[2]	Sum
4	VH4-31-19	0	
4	VH4-31-20	0	***57 entries***
5	VHS-12-1	82	
5	VHS-12-2	1	
5	VHS-12-3	0	
5	VHS-12-4	14	***97 entries***
6	VH6-35-1	74	***74 entries***

Table 4A: Analysis of V kappa subgroup 1



amino acid[1]	Framework I																								CDRI							
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	A	B	C	D	E
A		1	1						1				102		1					1	1				103							
B			1			1																					1					
C														1									105								
D	64																101														
E	8		14												1		2							1	1		2				
F						1			1	6				1								2									
G																	105						1								
H																											1				
I		65													4				6	4	101	1									
K			1																					2							
L		6		21							96		1											1							
M	1			66																											
N																										1					
P								103		1		2			1																
Q			62				88				1													20			100				
R												103							94	81											
S								89	102	80				103					5			102									
T		1			88					18									6	99		1	103		1	1					
V		1	9								8		2		98				98		2										
W																															
X	1																														
Y																			1												
-																															
unknown (?)																												105	105	105	105	105
not sequenced	31	31	18	18	17	16	16	16	7	2	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	A	B	C	D	E
sum of seq[2]	74	74	87	87	88	89	89	89	104	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105	105
oomcaa[3]	64	65	62	66	88	88	89	89	102	80	96	103	102	103	98	105	80	98	94	99	101	103	105	81	103	102	100	105	105	105	105	105
mcaa[4]	D	I	Q	M	T	Q	S	P	S	S	L	S	A	S	V	G	D	R	V	T	I	T	C	R	A	S	Q	-	-	-	-	-
rel. oomcaa[5]	98%		88%		76%		100%		100%	76%	91%	98%	97%	98%	93%	100%	96%	90%	93%	94%	96%	98%	100%	77%	98%	97%	95%	100%	100%	100%	100%	100%
pos occupied[6]	4	5	5	2	1	2	1	1	3	4	3	2	3	3	5	1	4	3	3	4	3	3	1	5	3	4	5	1	1	1	1	1

Table 4A: Analysis of V kappa subgroup 1

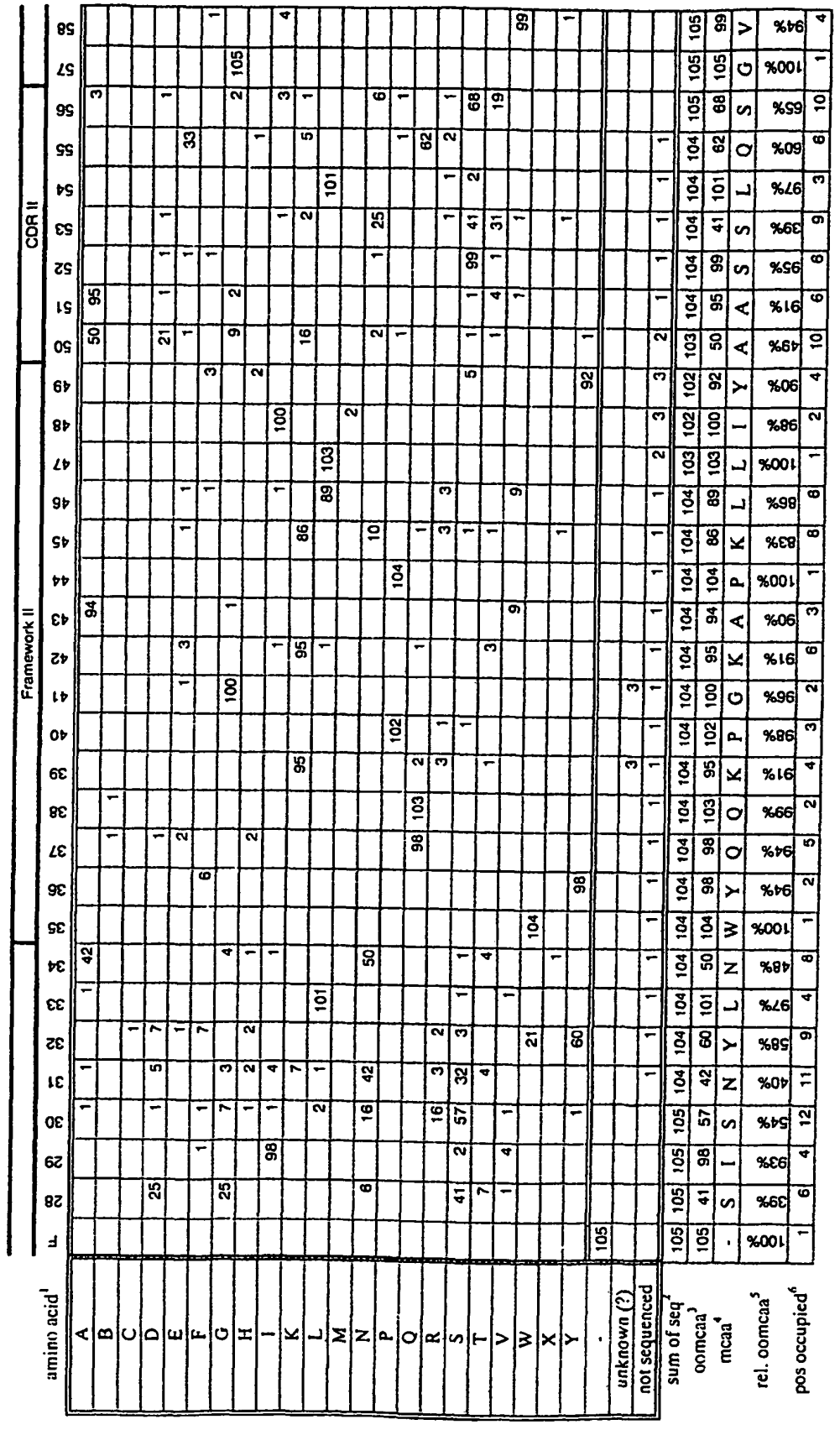


Table 4A: Analysis of V kappa subgroup 1

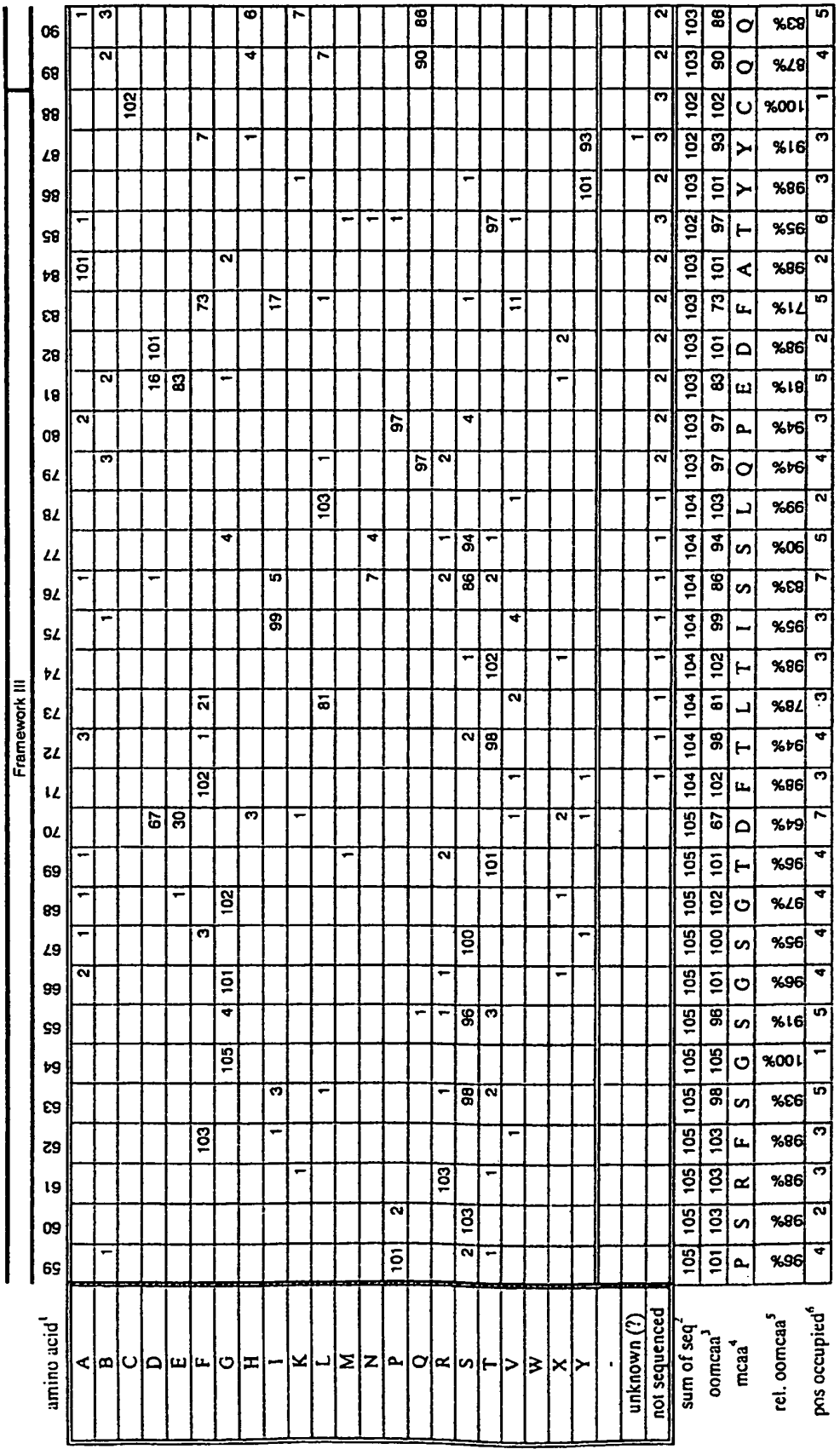


Table 4A: Analysis of V kappa subgroup 1

	CDR III													Framework IV												
amino acid[1]	91	92	93	94	95	A	B	C	D	E	F	96	97	98	99	100	101	102	103	104	105	106	A	107	108	sum
A	7	1		5	1							1														627
B																1					1					19
C																										209
D		23	5	1								1									15					459
E		1	1		1	1										2					65					258
F	3			13								6		86								2				451
G	1		1	2	1		1								87	29	87								2	894
H	7	3	1									2	1													40
I		4	1	2	1							5								1		72				606
K		1										1	1						77					79		480
L	6	2		18	2							18	1	1						22	4	2				793
M													1									5				77
N	6	31	19	1								1										1		2		232
P				1	82	6						6				7									1	620
Q	1	2										1				48					1					865
R	1		2	2								6							6					2	70	413
S	27	3	58	5	10							2	2													1636
T	3	1	15	25								2	82					87	3					2		1021
V				5								2							1	63		3				440
W				1								15														141
X																										14
Y	42	32	1	23								16														564
-					3	82	88	89	89	89	89	4	1										85		1	1250
unknown (?)																										7
not sequenced	1	1	1	1	4	16	16	16	16	16	16	16	16	18	18	18	18	18	18	19	19	20	20	20	31	589
sum of seq[2]	104	104	104	104	101	89	89	89	89	89	89	89	89	87	87	87	87	87	87	86	86	85	85	85	74	
oomcaa[3]	42	32	58	25	82	82	88	89	89	89	89	18	82	86	87	48	87	87	77	63	65	72	85	79	70	
mcaa[4]	Y	Y	S	T	P	-	-	-	-	-	-	L	T	F	G	G	G	T	K	V	E	I	-	K	R	
rel. oomcaa[5]	40%	31%	56%	24%	81%	92%	99%	100%	100%	100%	100%	20%	92%	99%	100%	55%	100%	100%	89%	73%	76%	85%	100%	93%	95%	
pos occupied[6]	11	12	10	14	8	3	2	1	1	1	1	17	7	2	1	5	1	1	4	3	5	6	1	4	4	

Table 4B: Analysis of V kappa subgroup 2

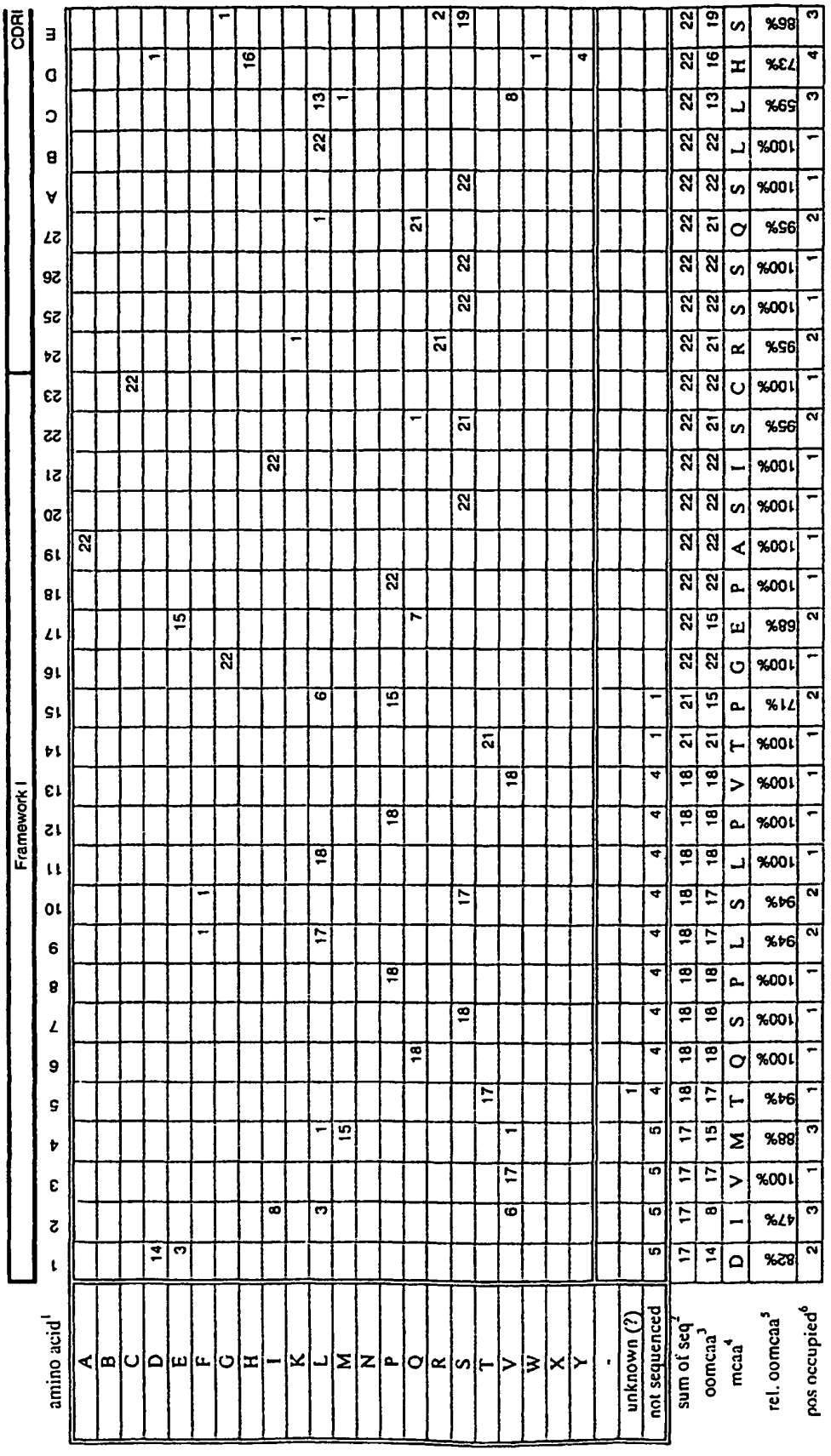


Table 4B: Analysis of V kappa subgroup 2

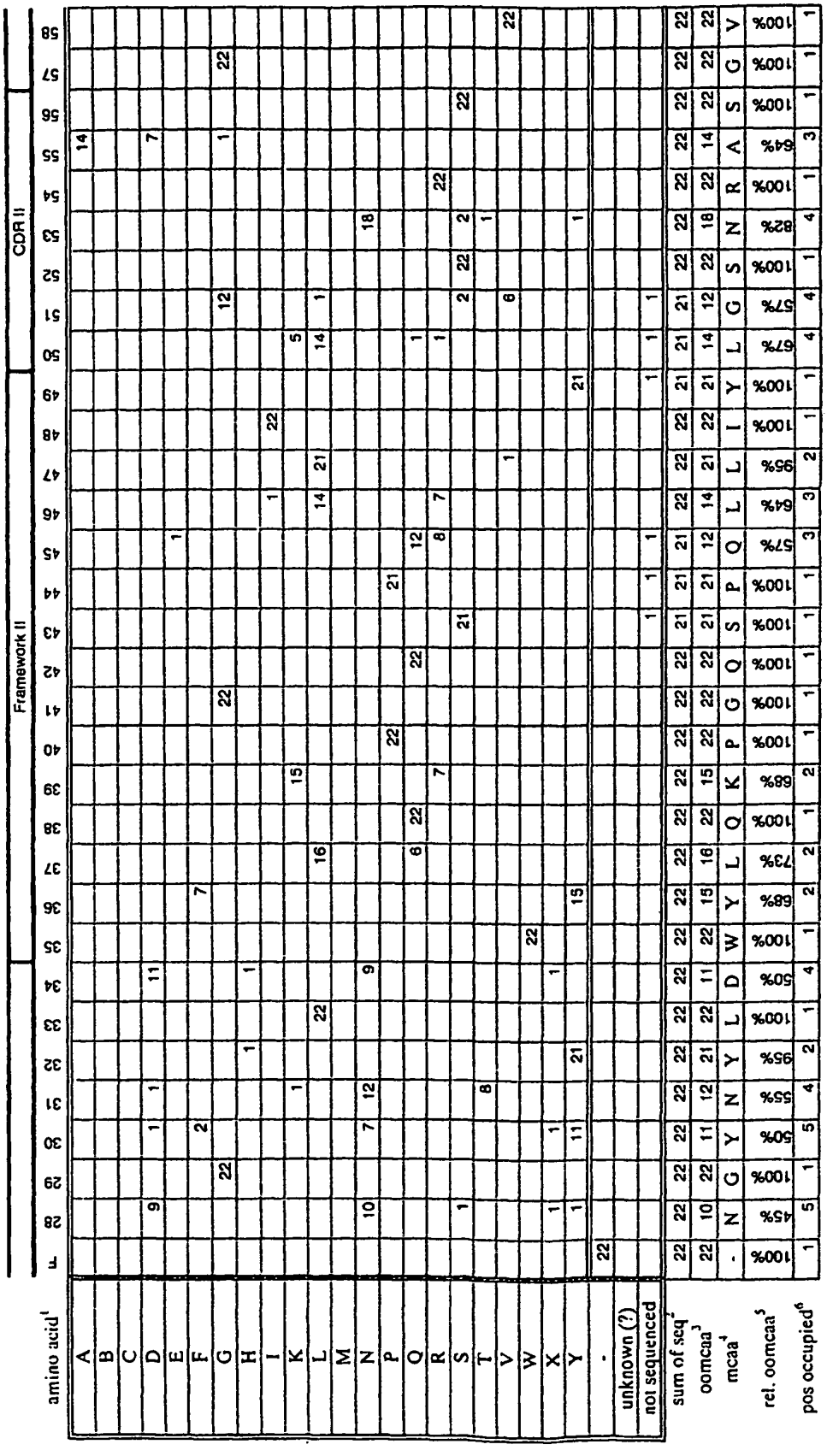


Table 4B: Analysis of V kappa subgroup 2

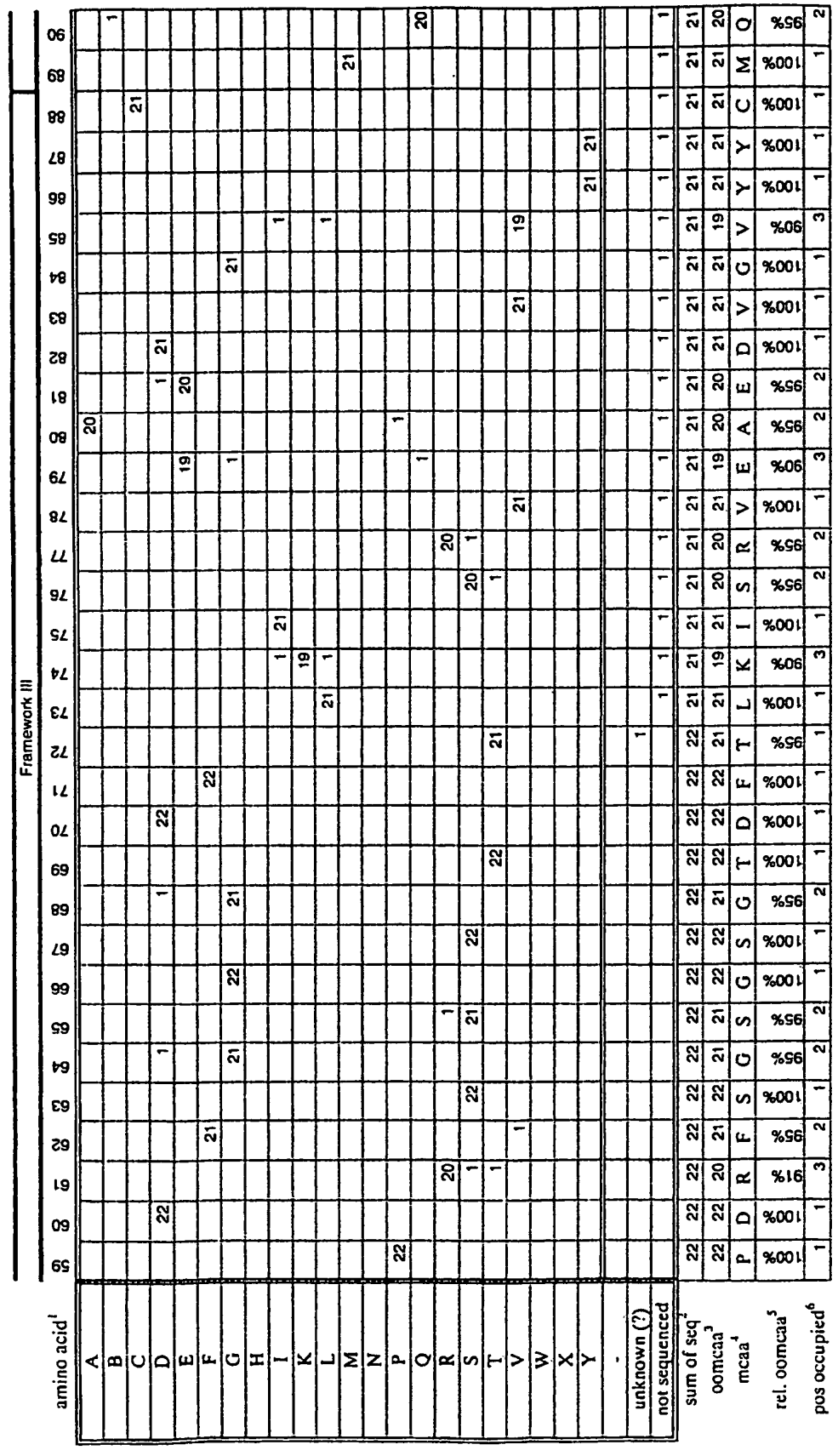


Table 4B: Analysis of V kappa subgroup 2

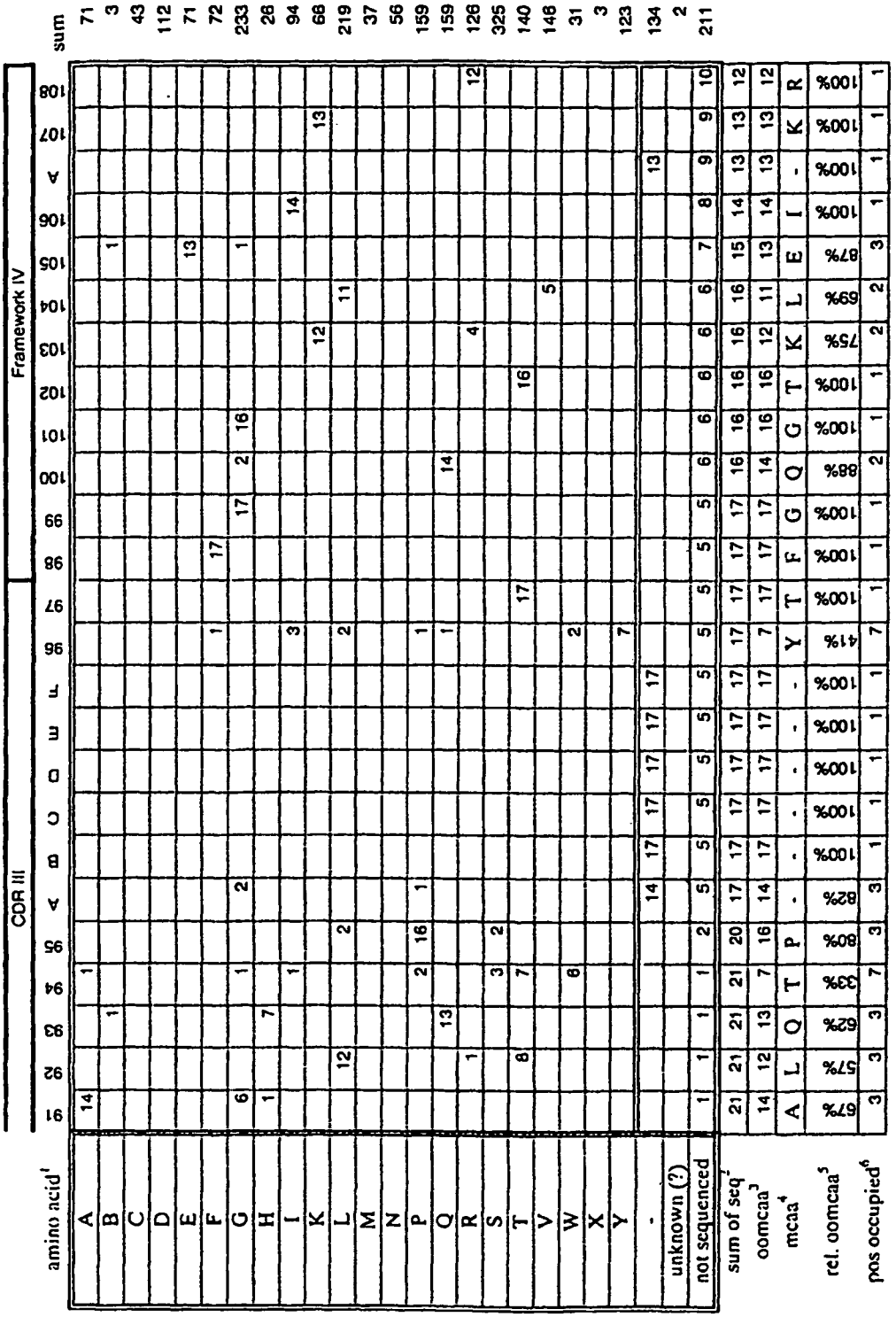


Table 4C: Analysis of V kappa subgroup 3

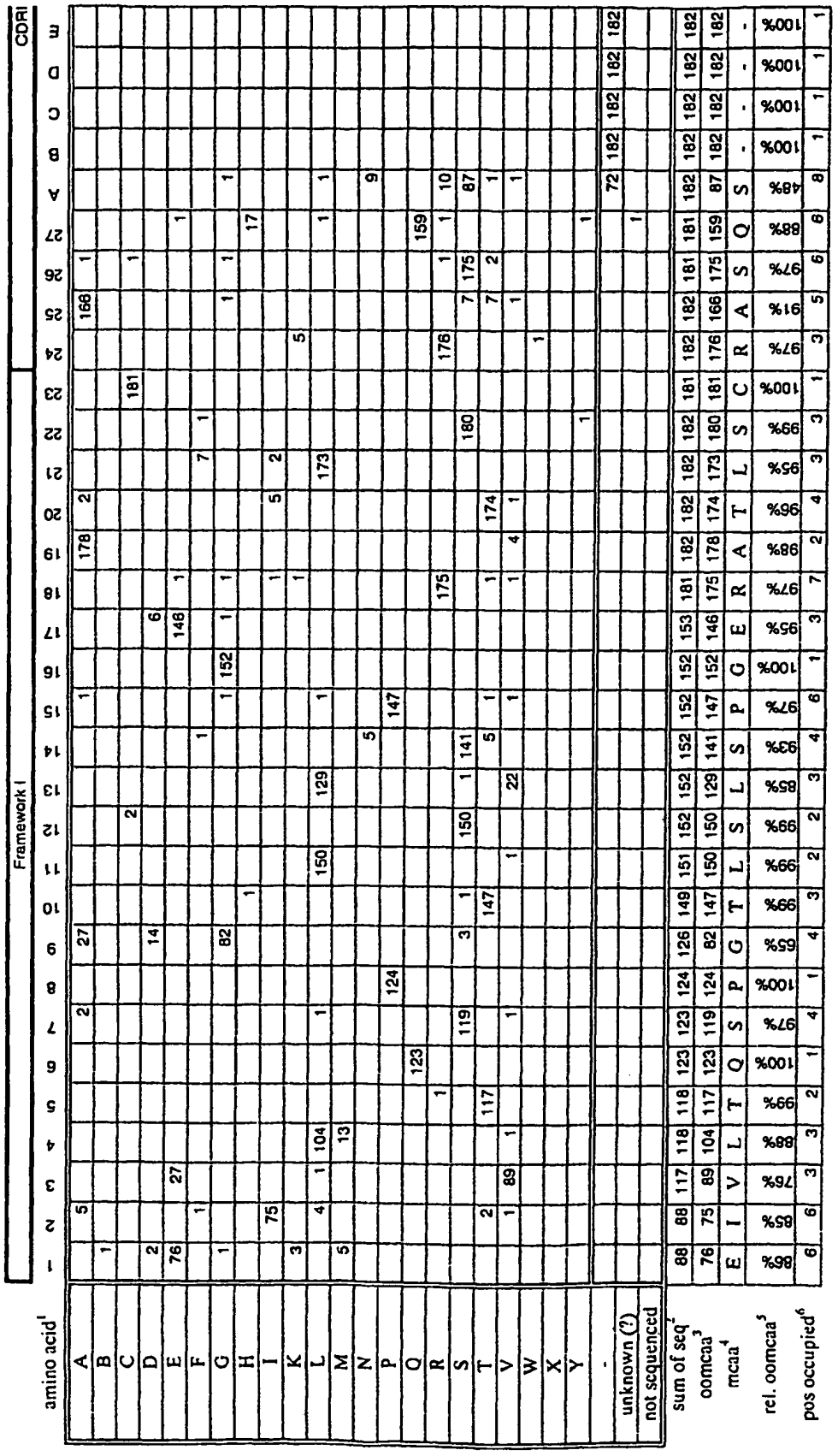


Table 4C: Analysis of V kappa subgroup 3

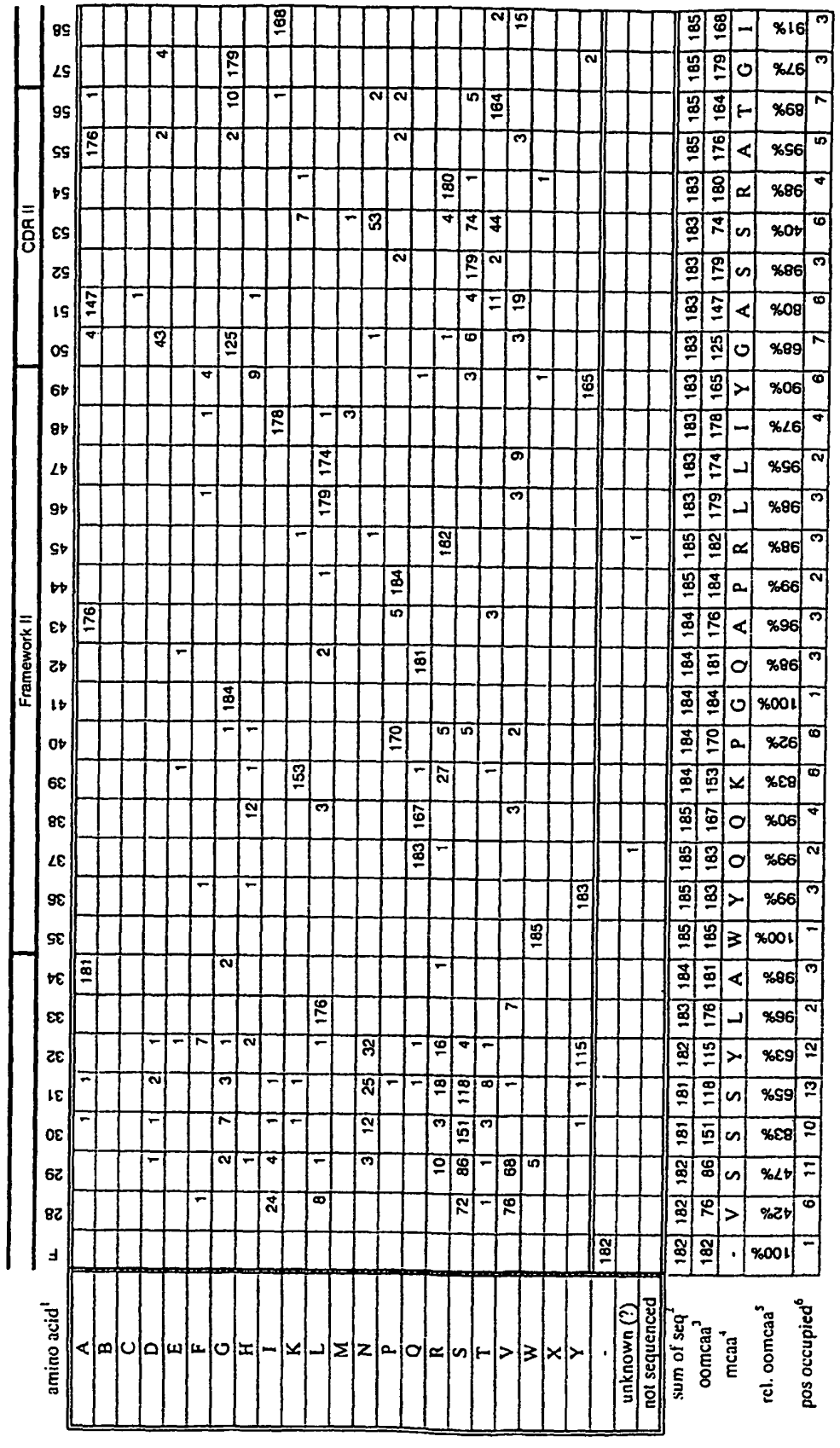


Table 4C: Analysis of V kappa subgroup 3

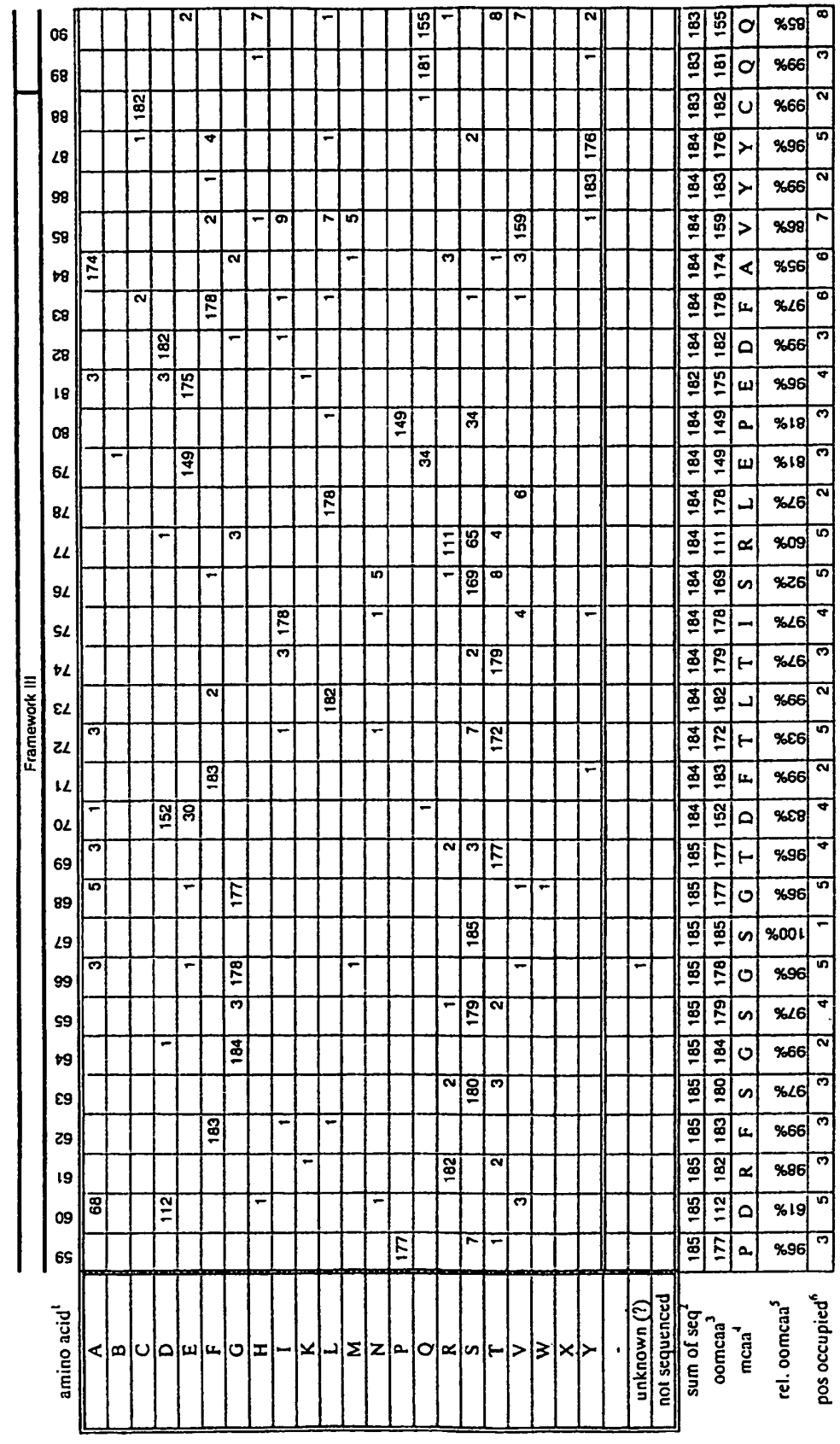


Table 4C: Analysis of V kappa subgroup 3

amino acid[1]	CDR III													Framework IV												sum
	91	92	93	94	95	A	B	C	D	E	F	96	97	98	99	100	101	102	103	104	105	106	A	107	108	
A		1	8	3	3											1										1345
B																										2
C	2			1								2														375
D		8	5										1								23					564
E		2										1							3		141					759
F	5		2									7		166								6				765
G	1	104	15		1	1	2					1			166	41	166								1	1804
H	4	1										2									1					64
I			1			1						4										143				803
K			2			1						1				1			152					157		489
L				2	7	5						42								54		1			2	1596
M		1			1	2																3				36
N		28	71									1						1						3		255
P				1	139	24						7	2			9		1		1						1147
Q	1		1		3	1						3				114			1		1					1314
R	34	2	3		2	2						19							9			2		4	134	1326
S	2	33	58	102	15	2						1	8					2								2629
T		2	13	1	1	2						1	154					162	1					1		1593
V					3	1						2								111		11				646
W				69								24														287
X																										
Y	134	1	1									43							1							1014
-			3	3	7	127	167	169	169	169	169	8	1	1	1	1	1	1	1	1	1	1	166	1	1	2151
unknown (?)																										4
not sequenced						14	14	14	14	14	14	14	17	16	16	16	16	16	15	16	16	16	17	17	45	337
sum of seq[2]	183	183	183	182	182	169	169	169	169	169	169	169	166	167	167	167	167	167	168	167	167	167	166	166	138	
oomcaa[3]	134	104	71	102	139	127	167	169	169	169	169	43	154	166	166	114	166	162	152	111	141	143	166	157	134	
mcaa[4]	Y	G	N	S	P	-	-	-	-	-	-	Y	T	F	G	Q	G	T	K	V	E	I	-	K	R	
rel. oomcaa[5]	73%	57%	39%	56%	76%	75%	99%	100%	100%	100%	100%	25%	93%	99%	99%	68%	99%	97%	90%	66%	84%	86%	100%	95%	97%	
pos occupied[6]	8	11	13	8	11	12	2	1	1	1	1	18	5	2	2	6	2	5	7	4	5	7	1	5	4	

Table 4D: Analysis of V kappa subgroup 4

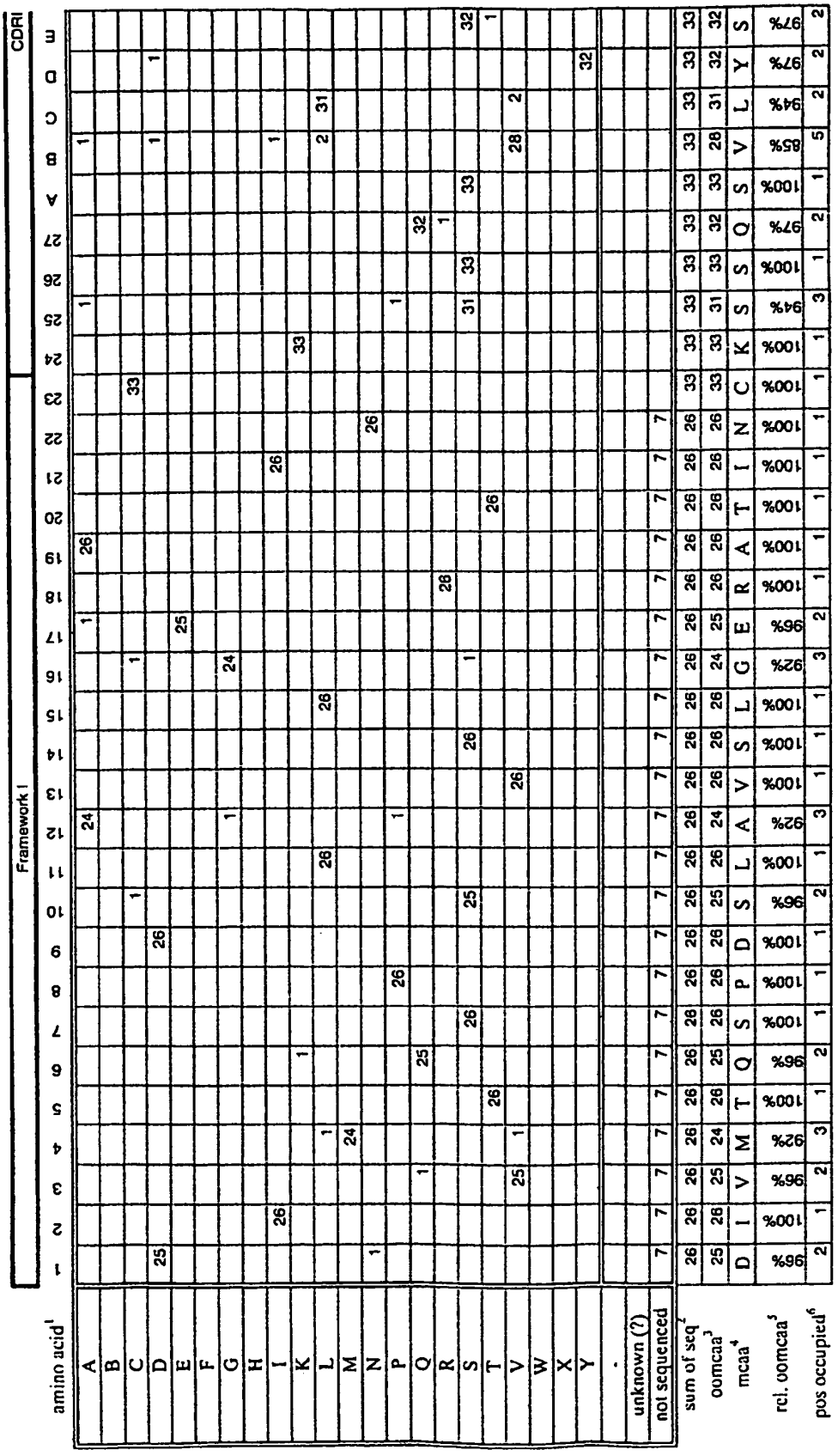


Table 4D: Analysis of V kappa subgroup 4

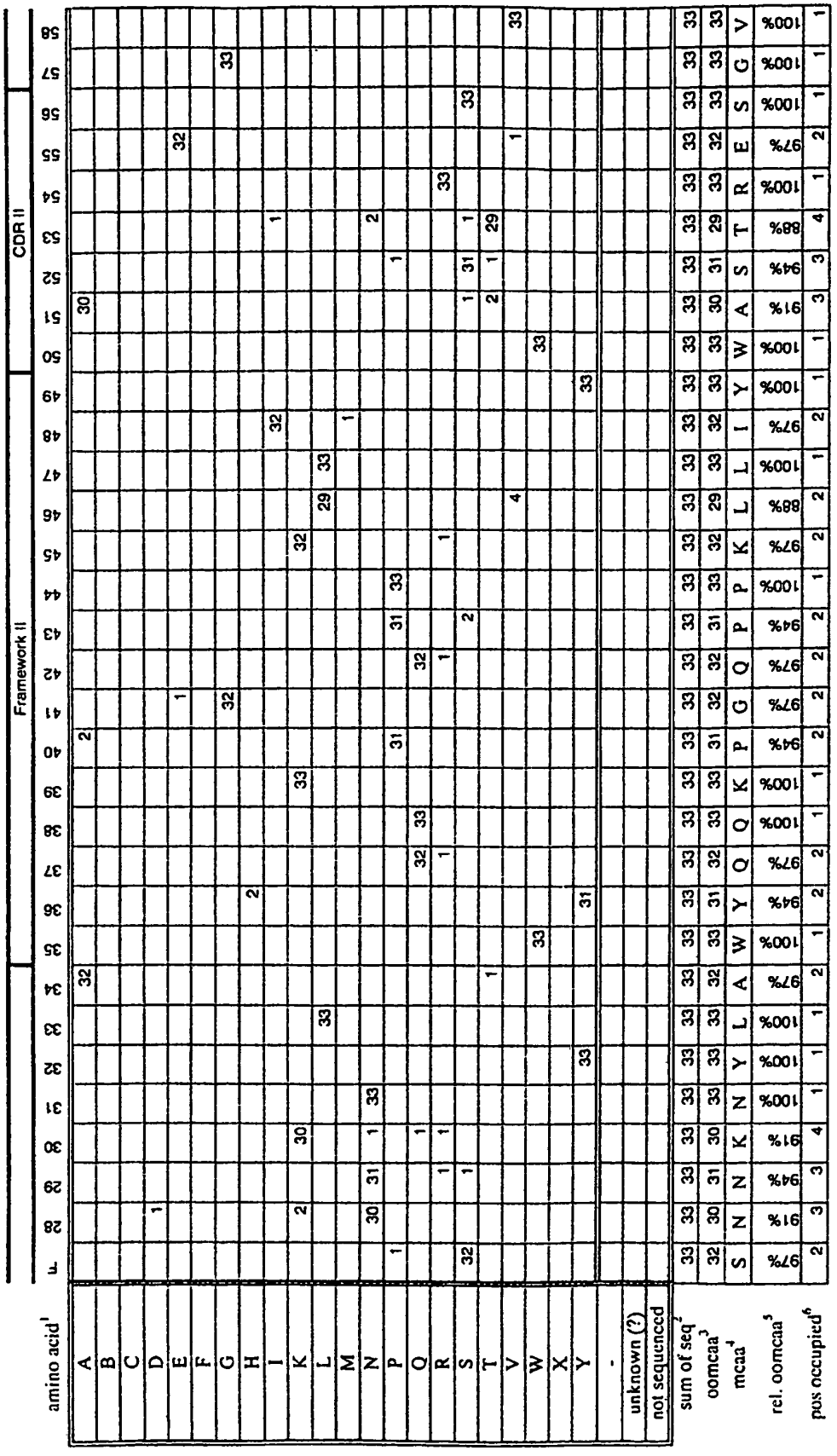


Table 4D: Analysis of V kappa subgroup 4

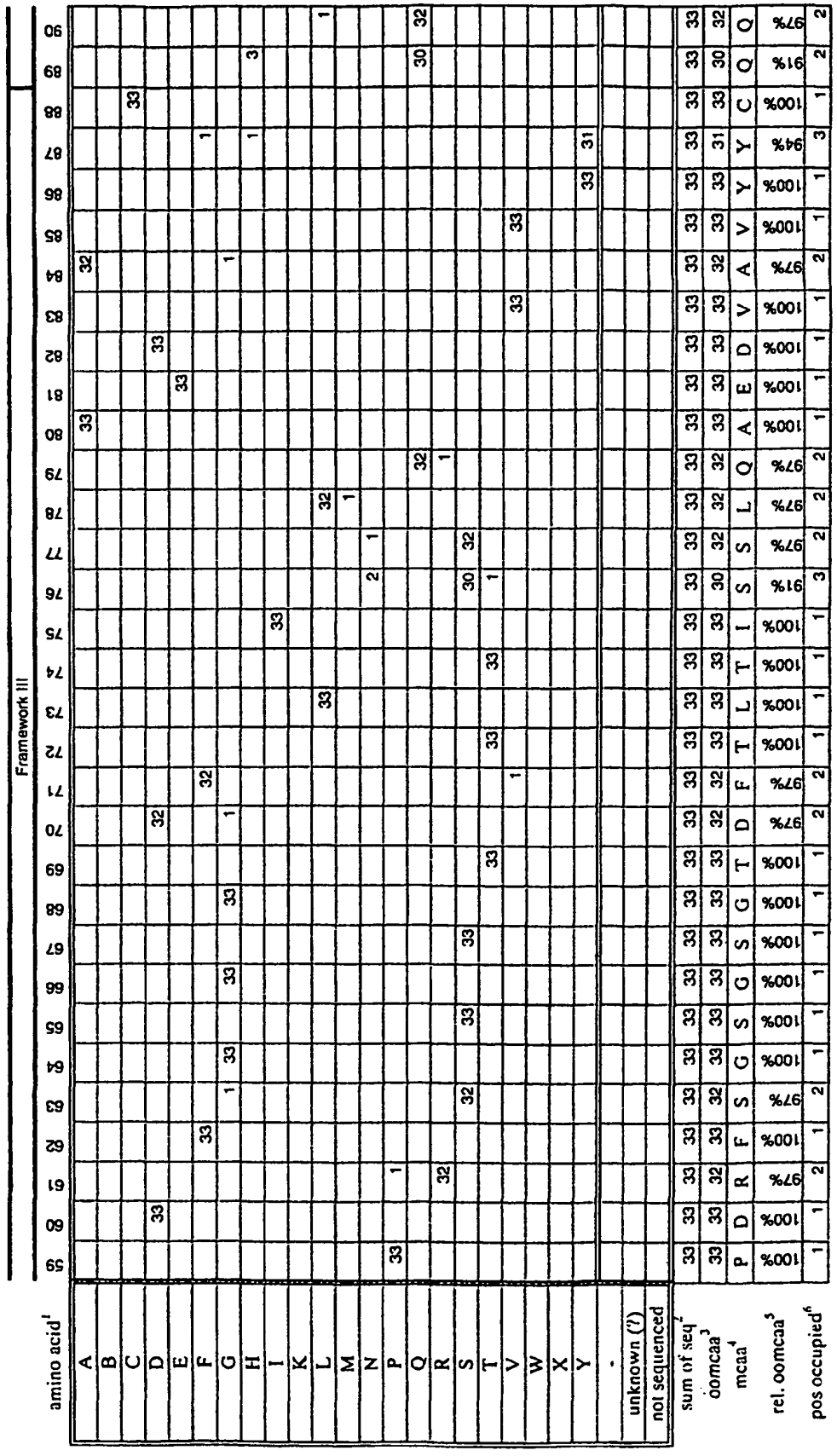


Table 4D: Analysis of V kappa subgroup 4

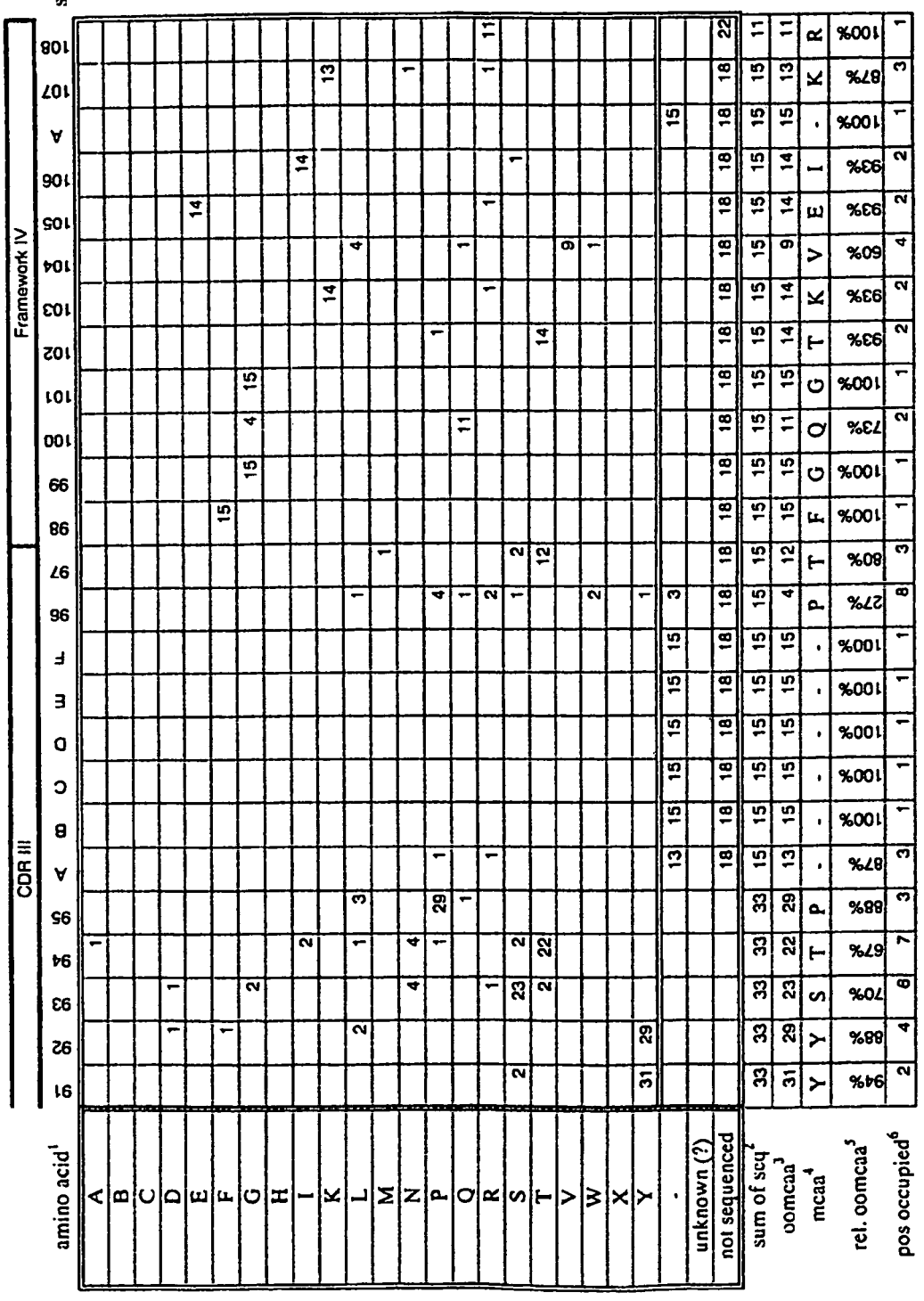


Table 5A: Analysis of V lambda subgroup 1

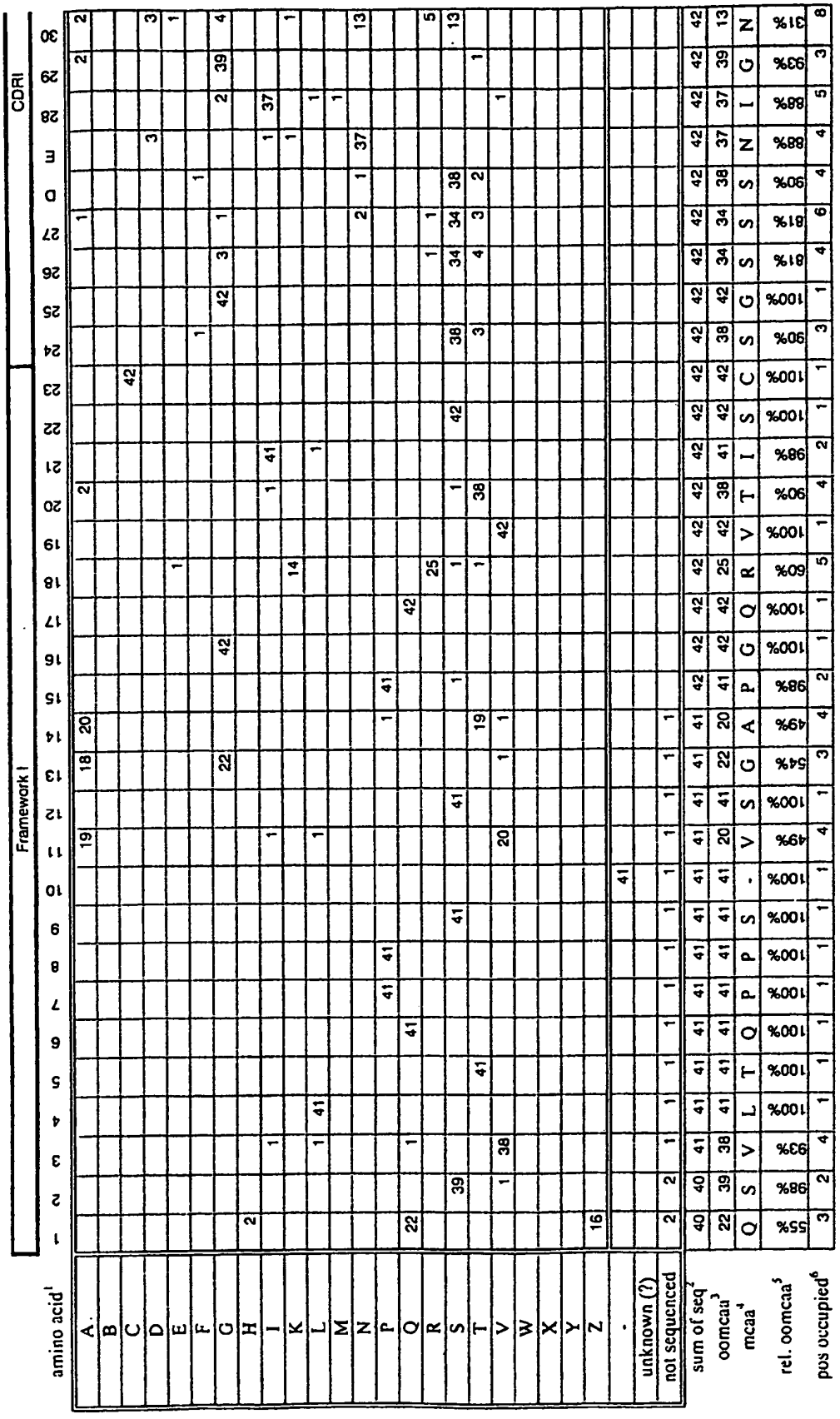


Table 5A: Analysis of V lambda subgroup 1

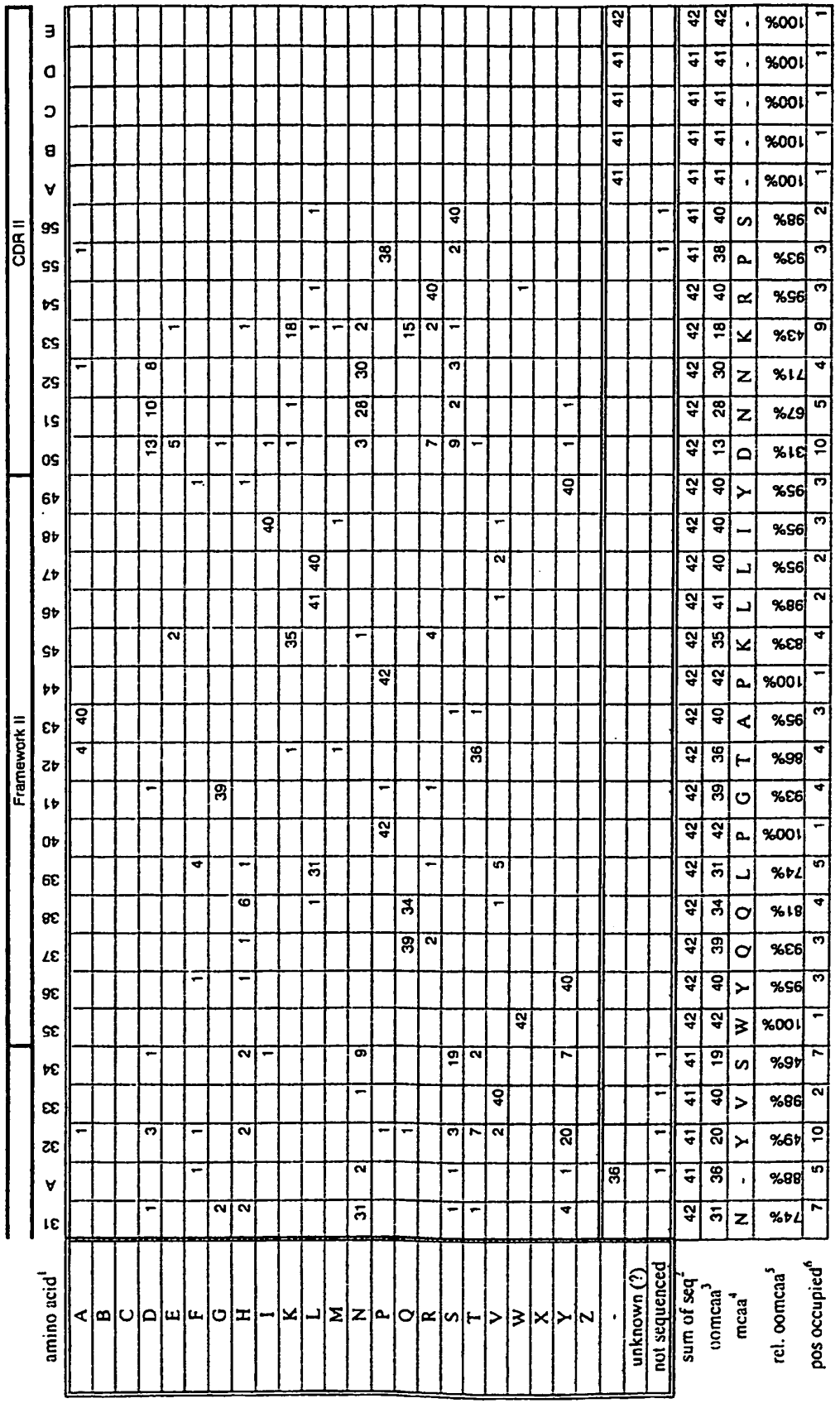


Table 5A: Analysis of V lambda subgroup 1

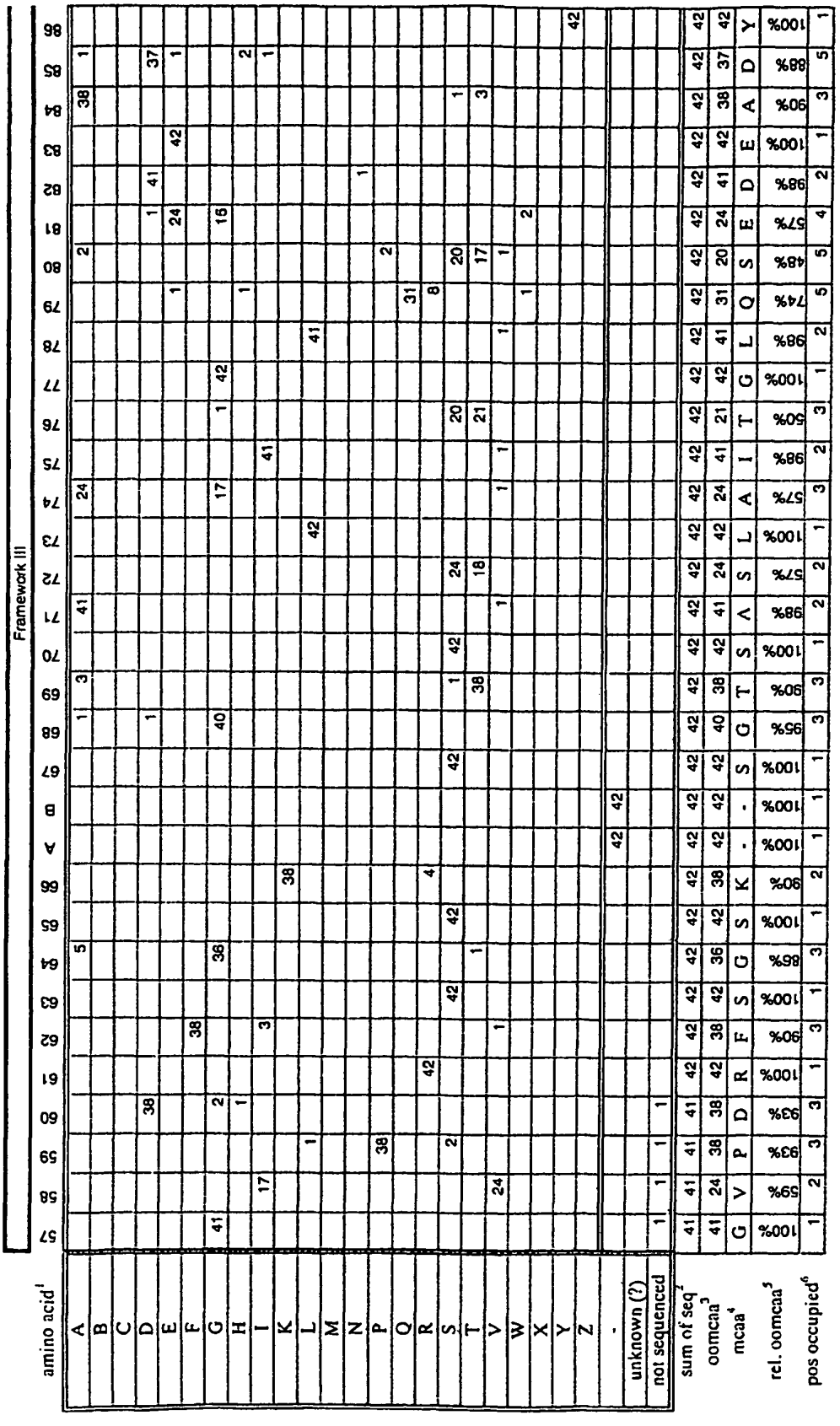


Table 5A: Analysis of V lambda subgroup 1

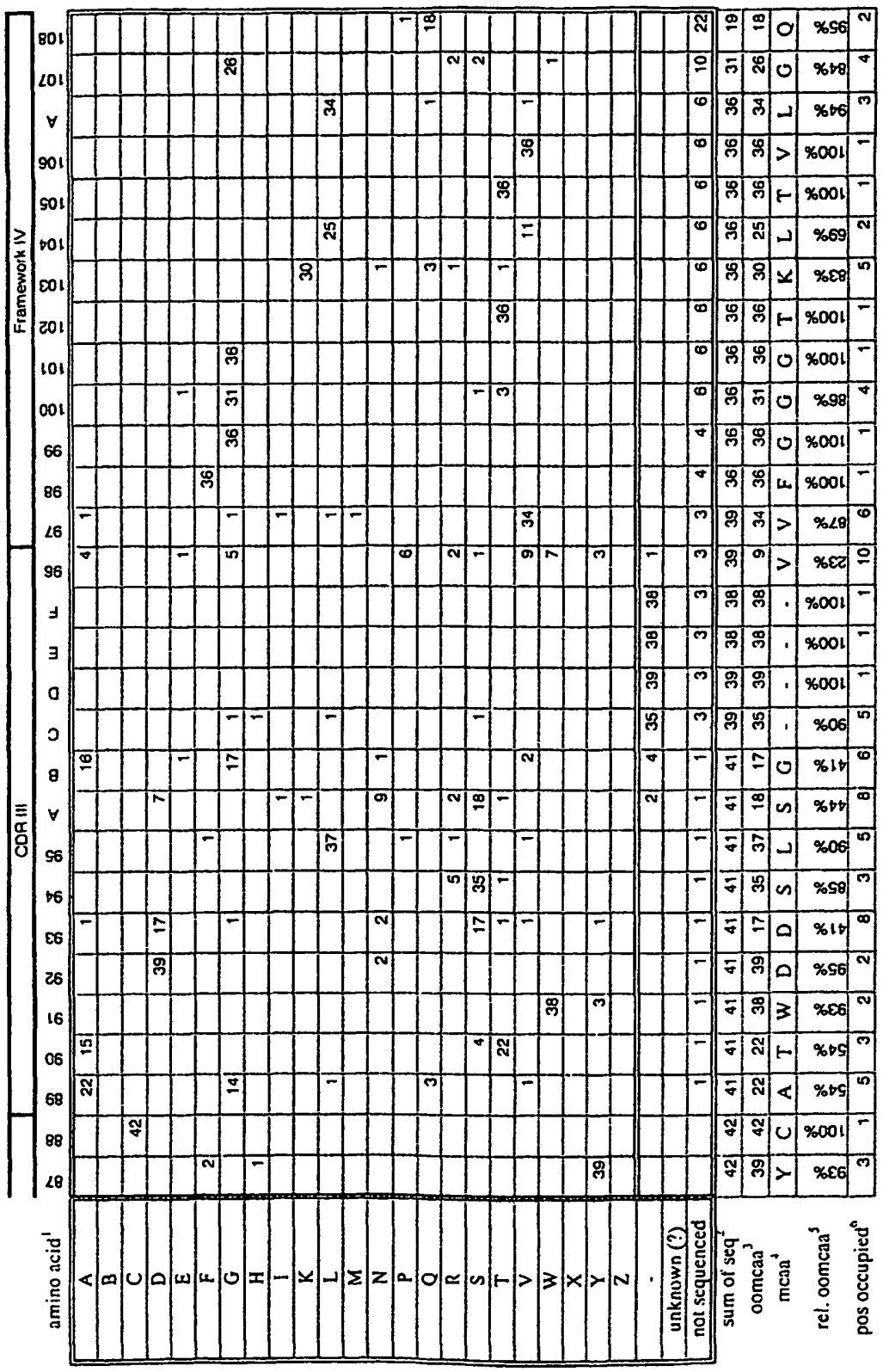


Table 5B: Analysis of V lambda subgroup 2

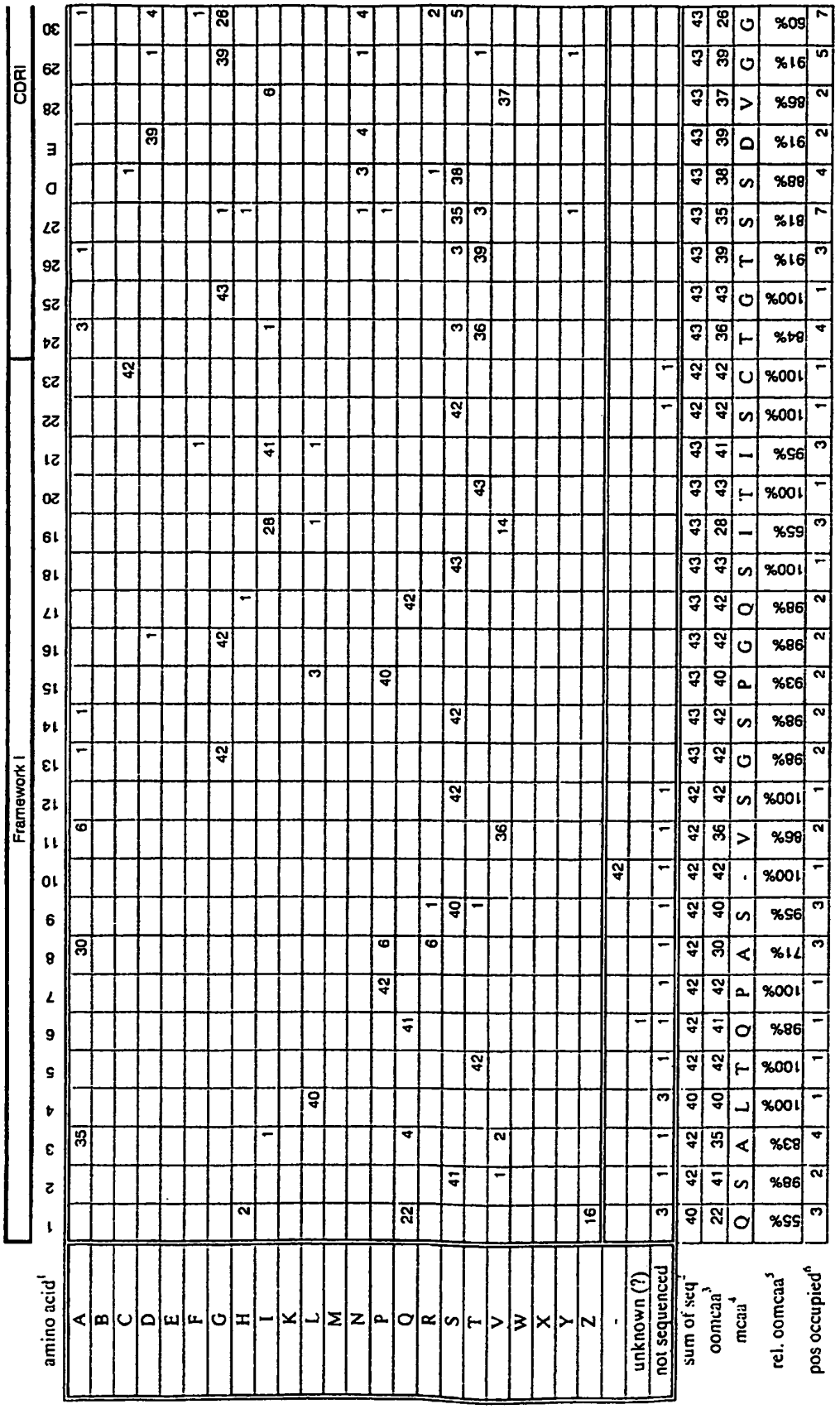


Table 5B: Analysis of V lambda subgroup 2

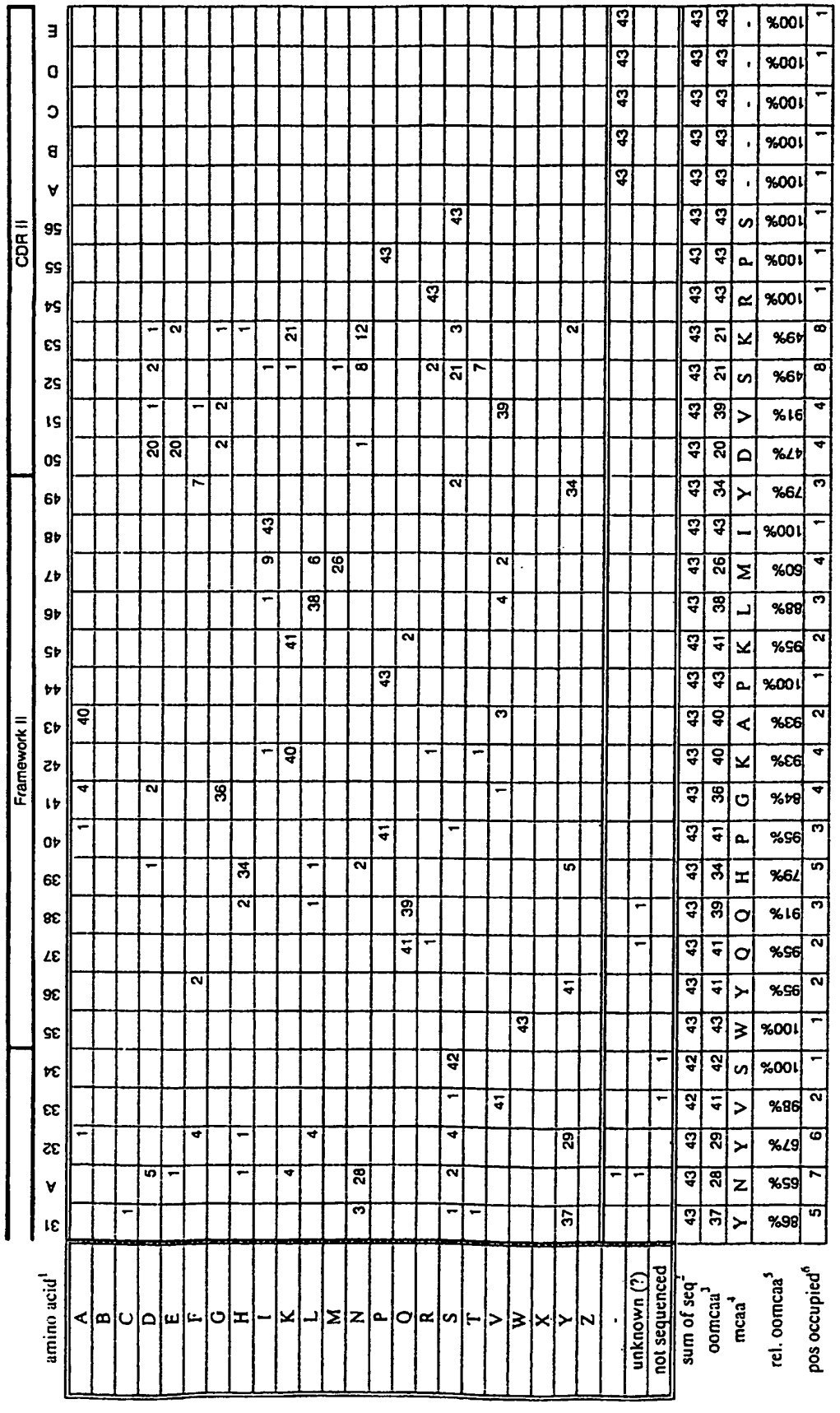


Table 5B: Analysis of V lambda subgroup 2

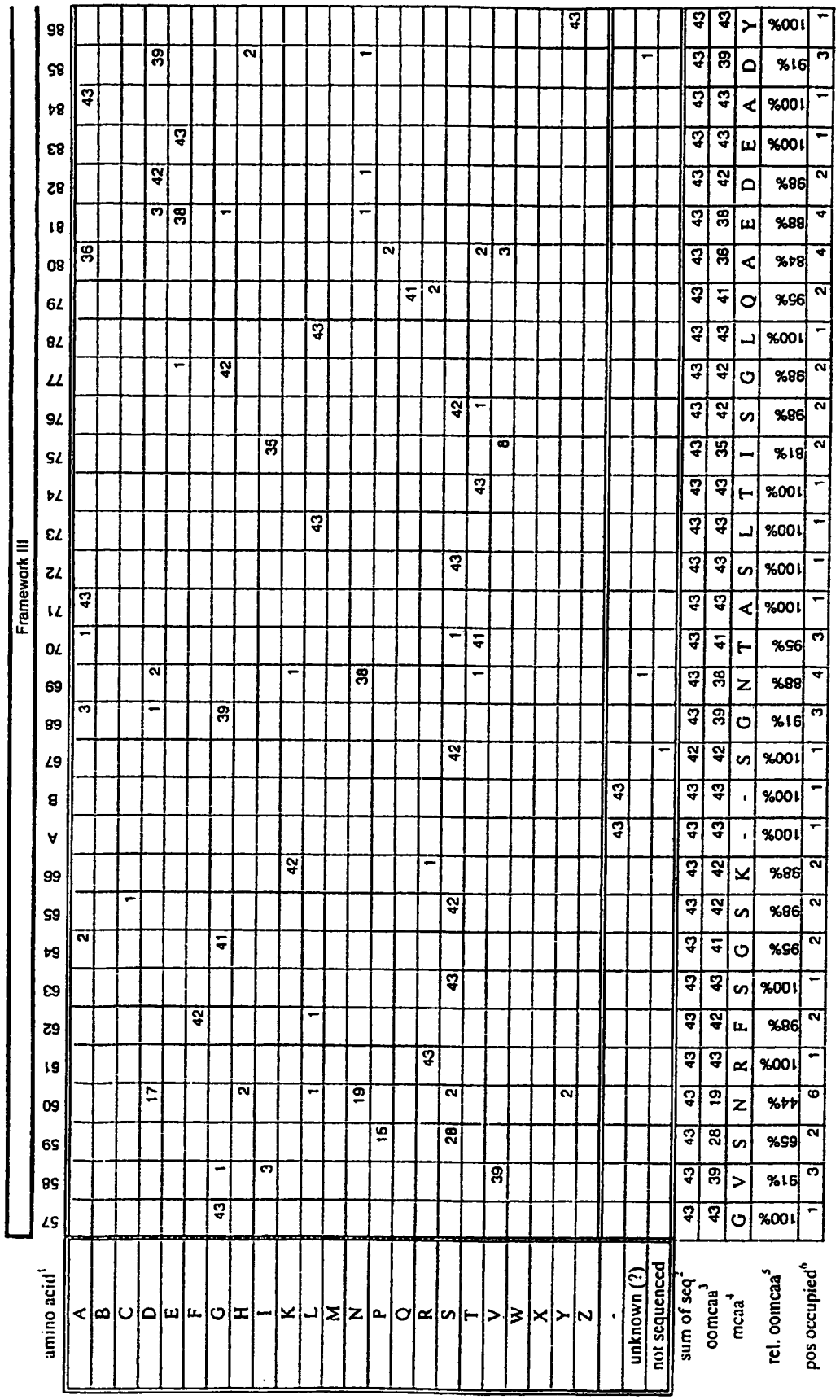


Table 5B: Analysis of V lambda subgroup 2

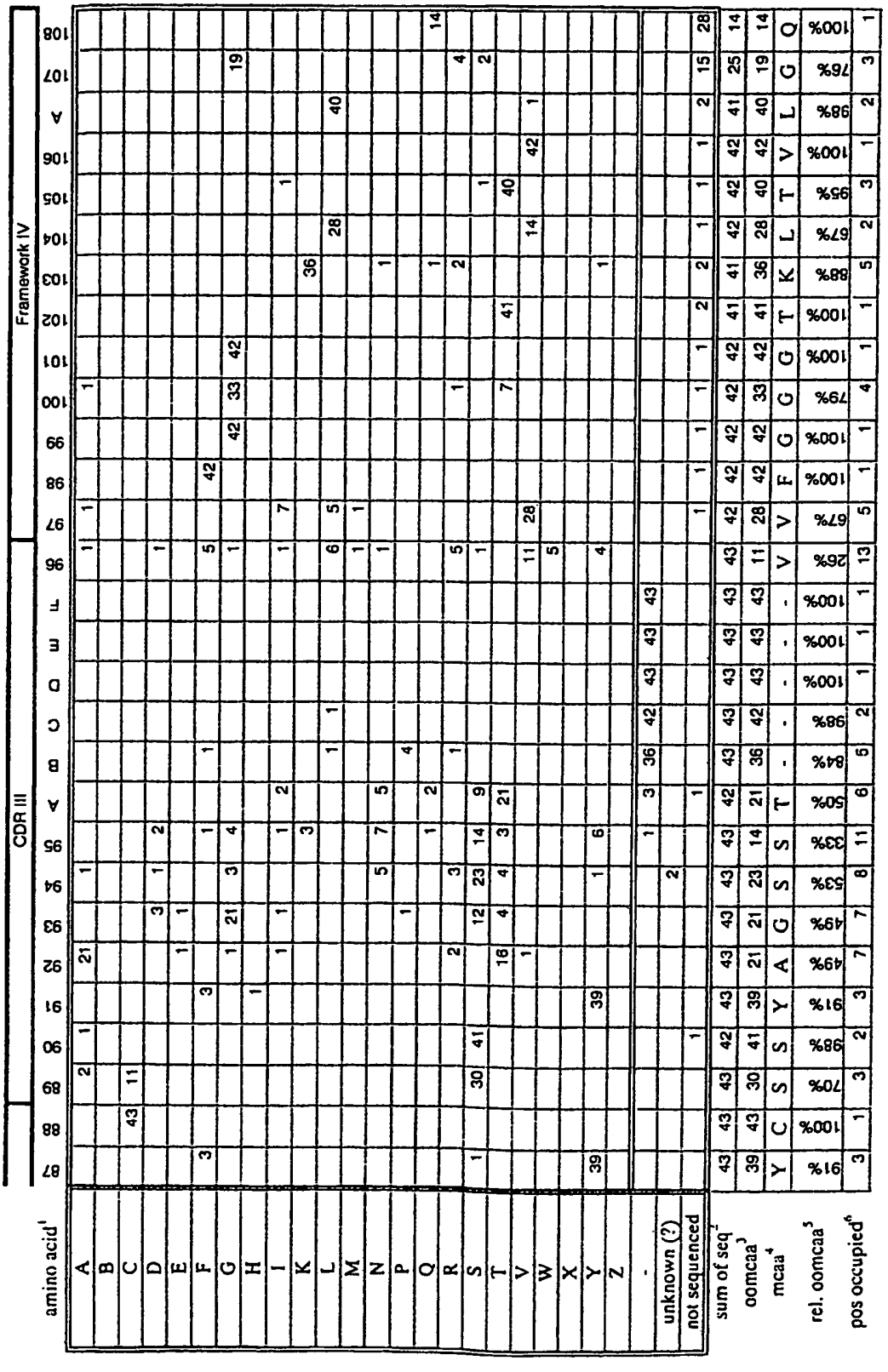


Table 5C: Analysis of V lambda subgroup 3

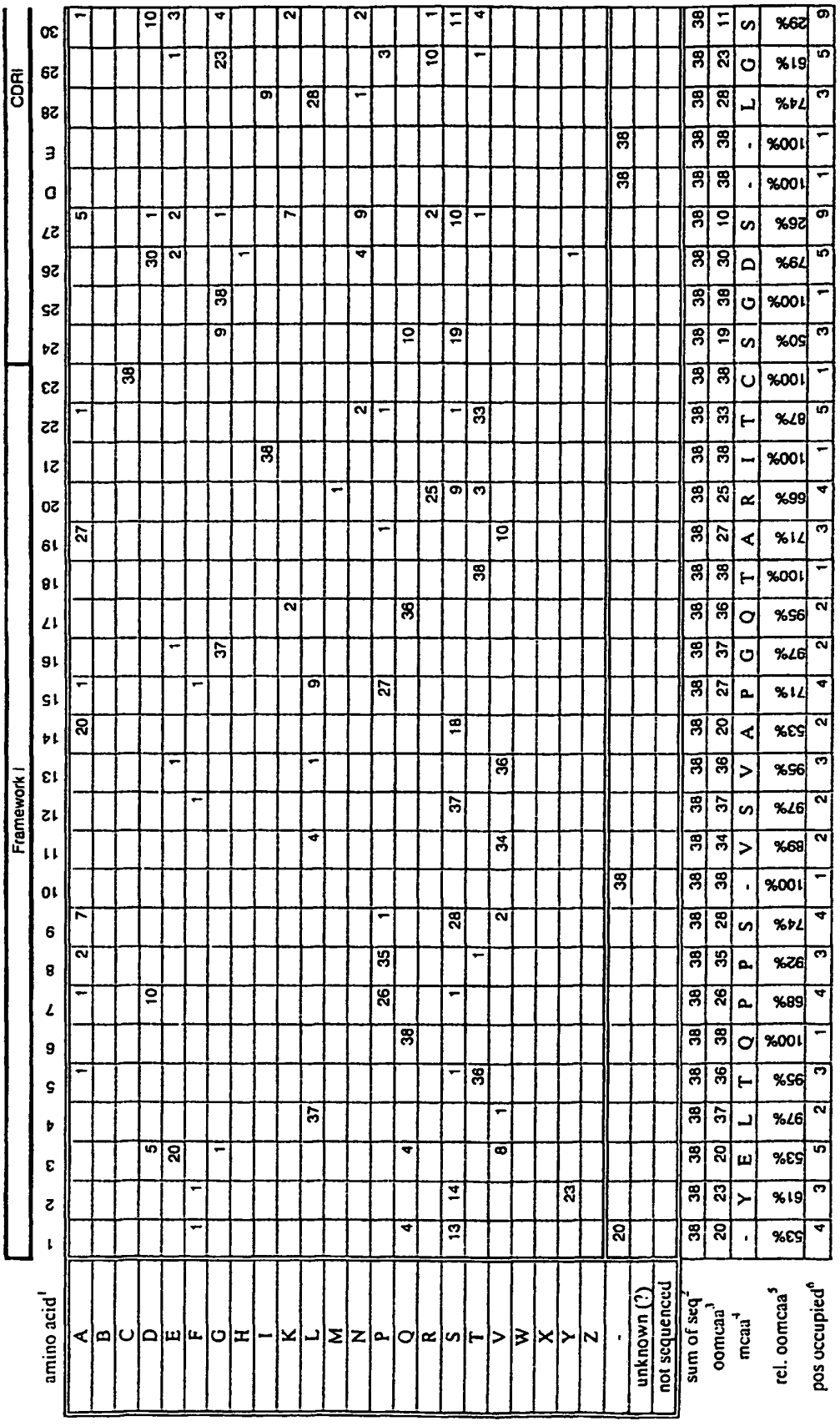


Table 5C: Analysis of V lambda subgroup 3

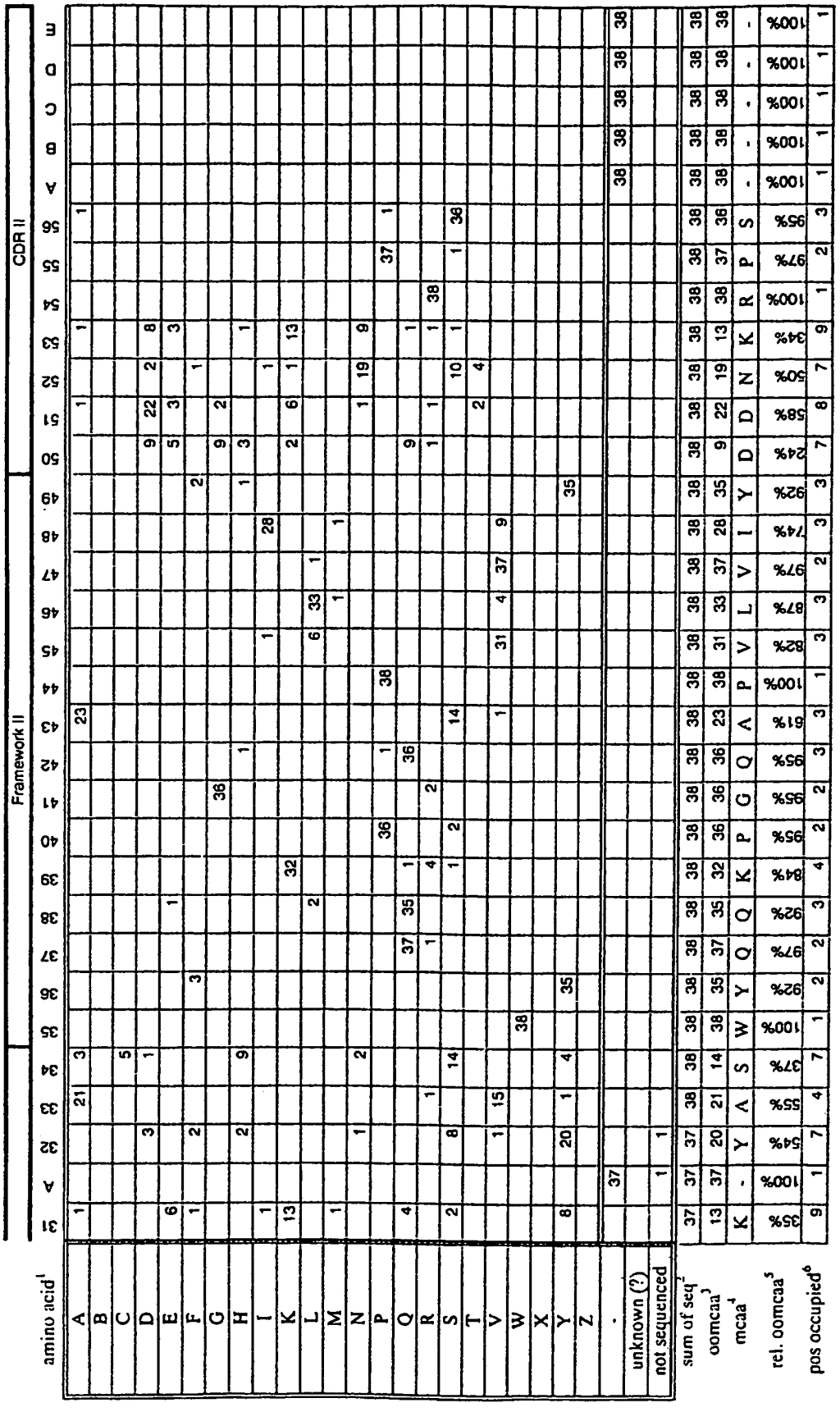


Table 5C: Analysis of V lambda subgroup 3

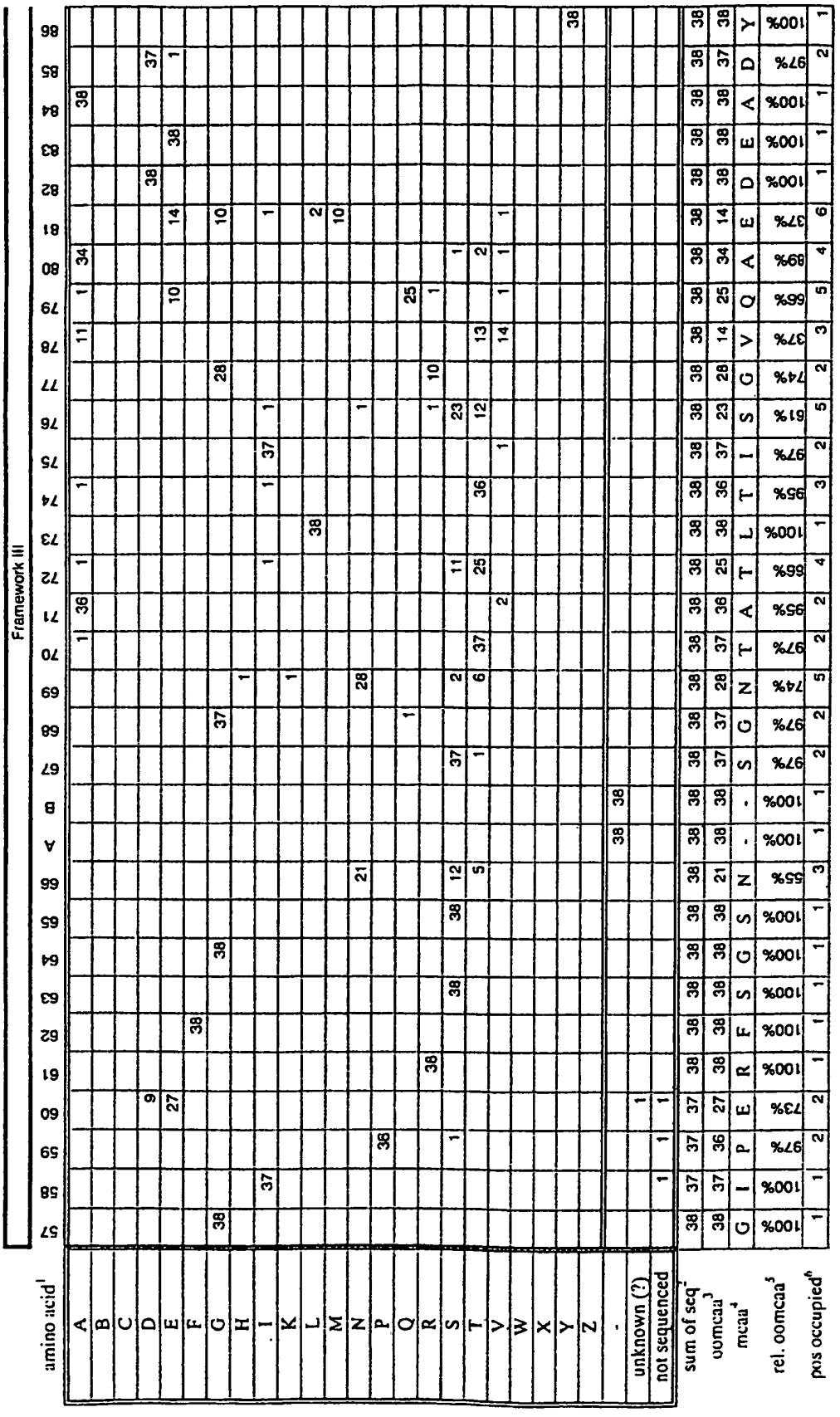


Table 5C: Analysis of V lambda subgroup 3

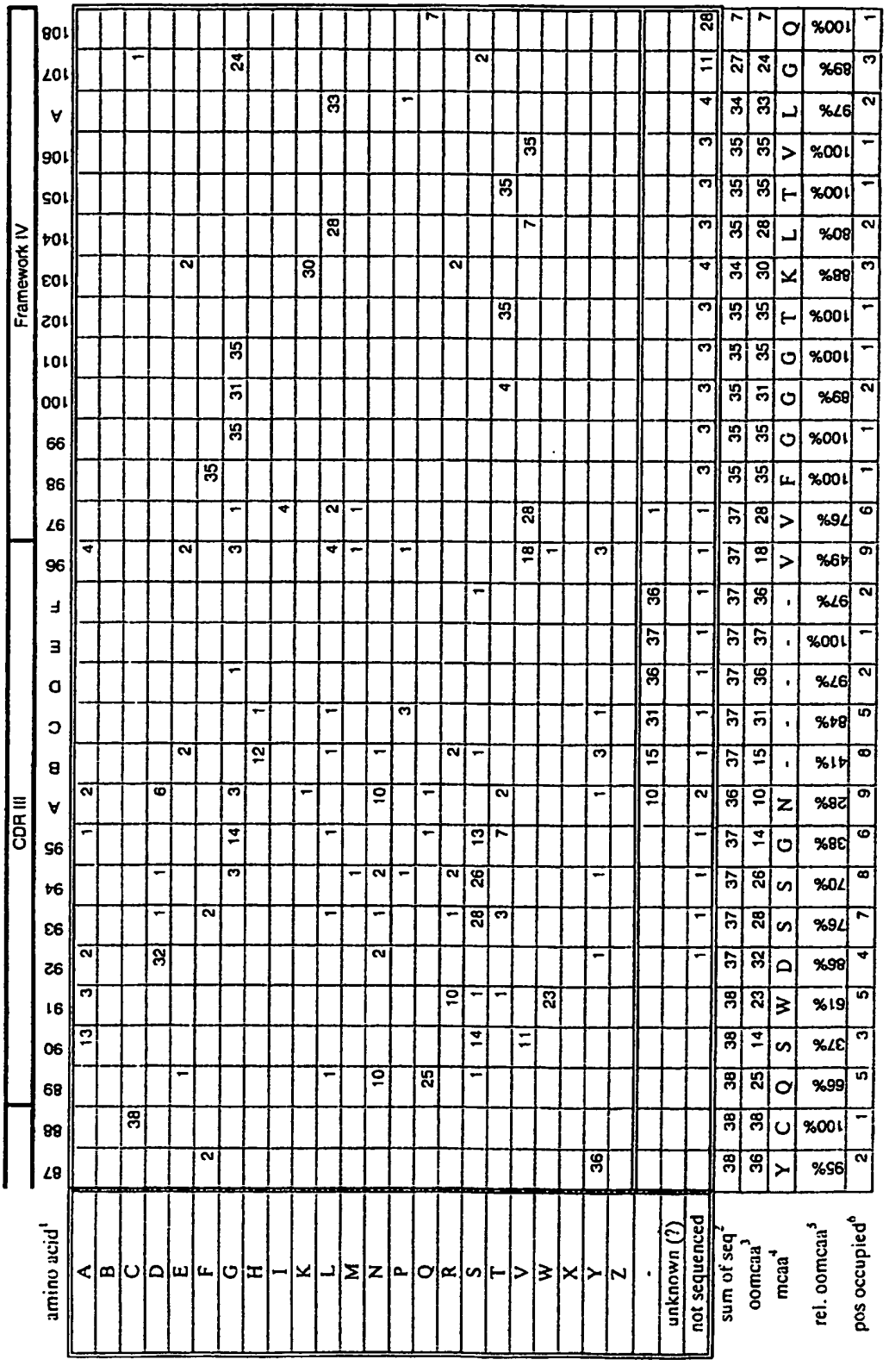


Table 6A: Analysis of V heavy chain subgroup 1A

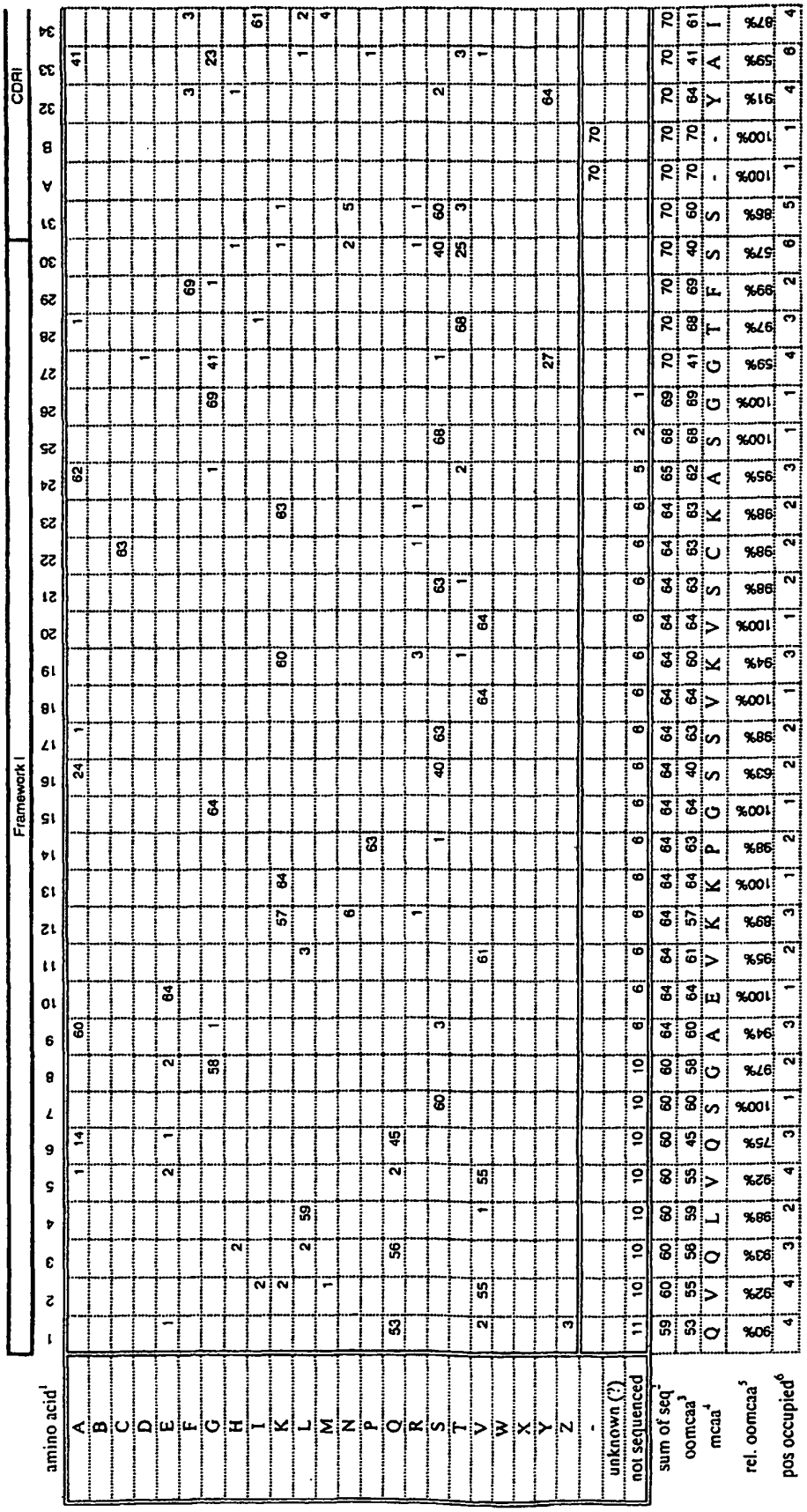


Table 6A: Analysis of V heavy chain subgroup 1A

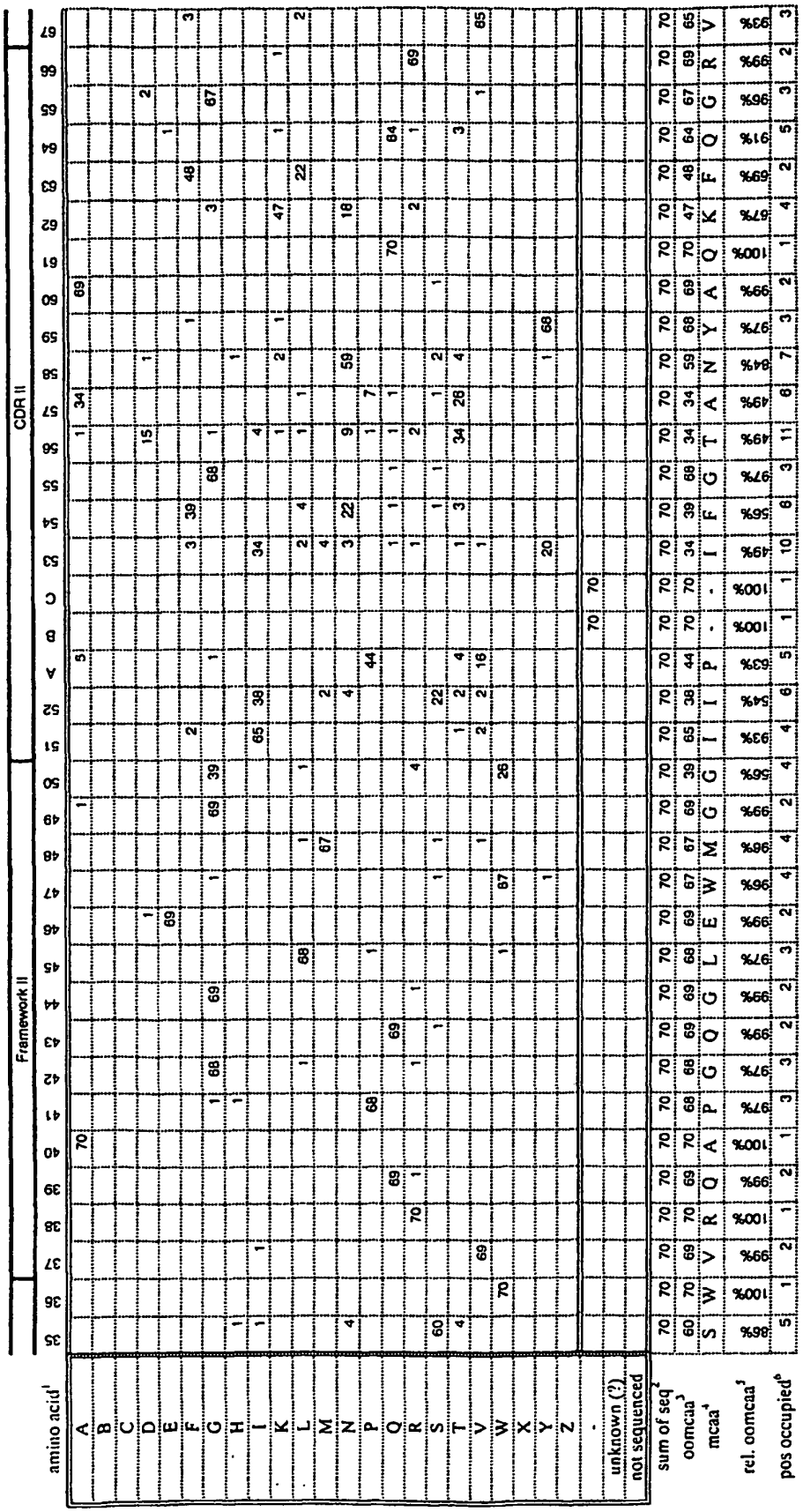


Table 6A: Analysis of V heavy chain subgroup 1A

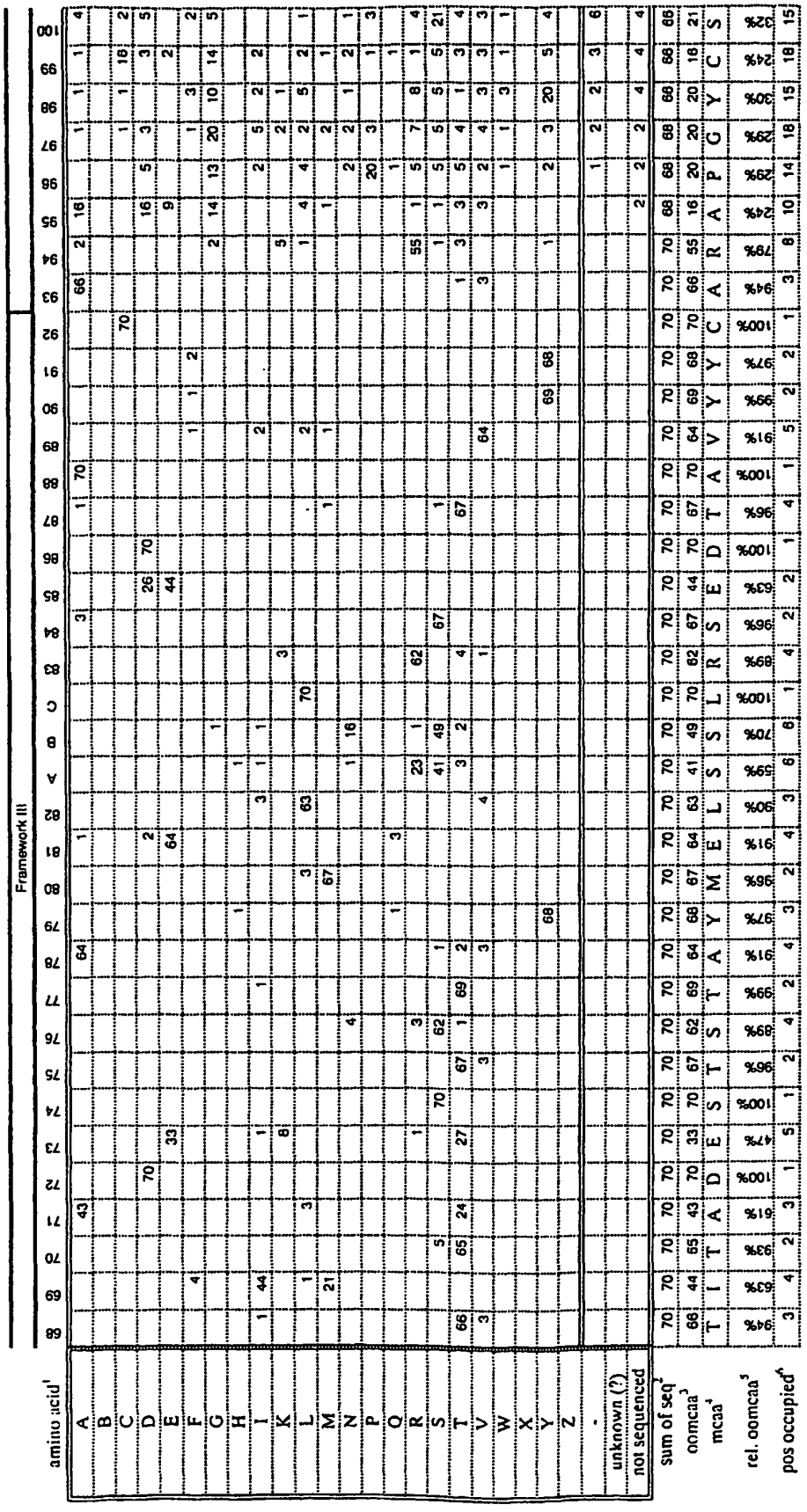


Table 6A: Analysis of V heavy chain subgroup 1A

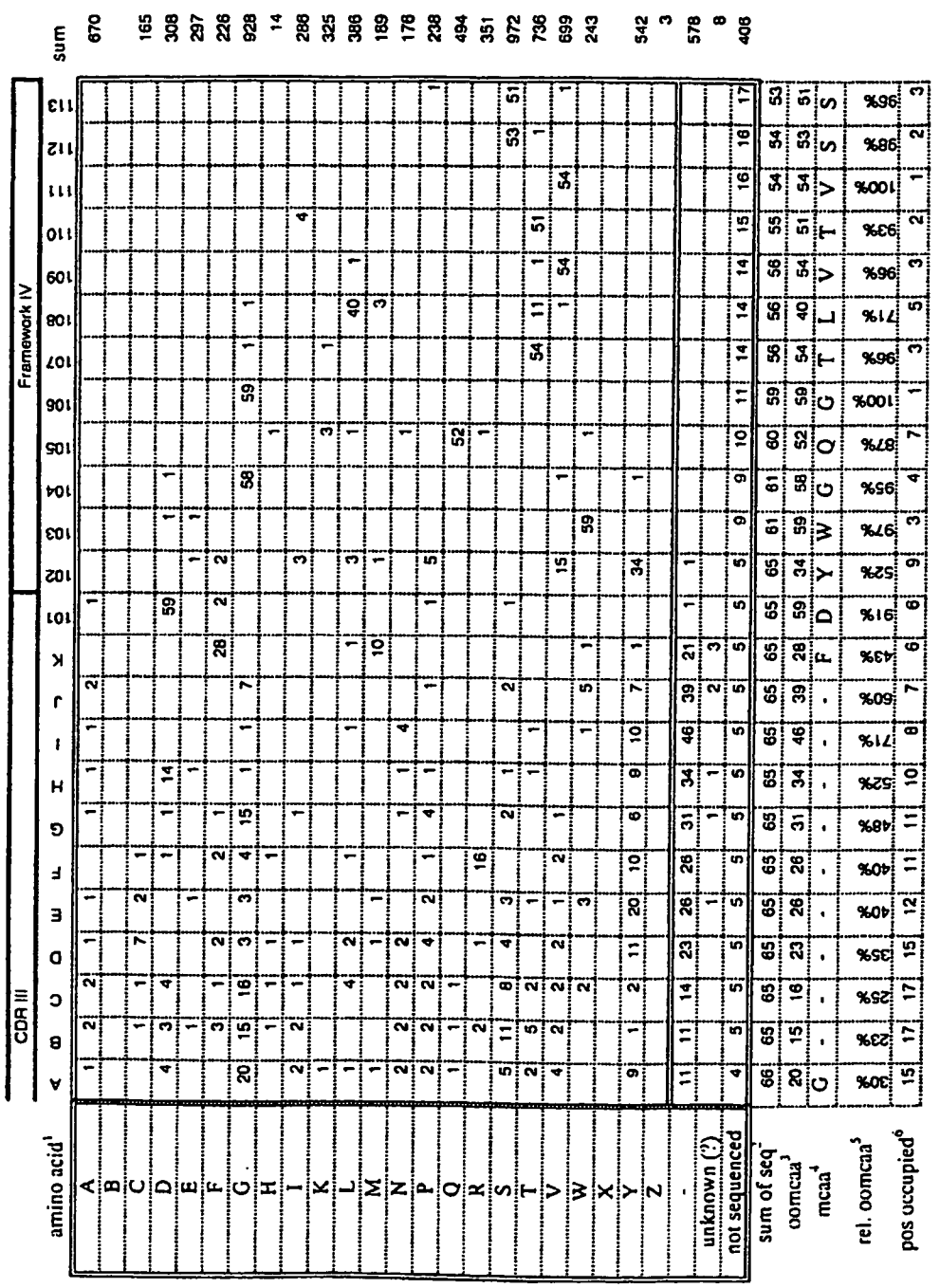


Table 6B: Analysis of V heavy chain subgroup 1B

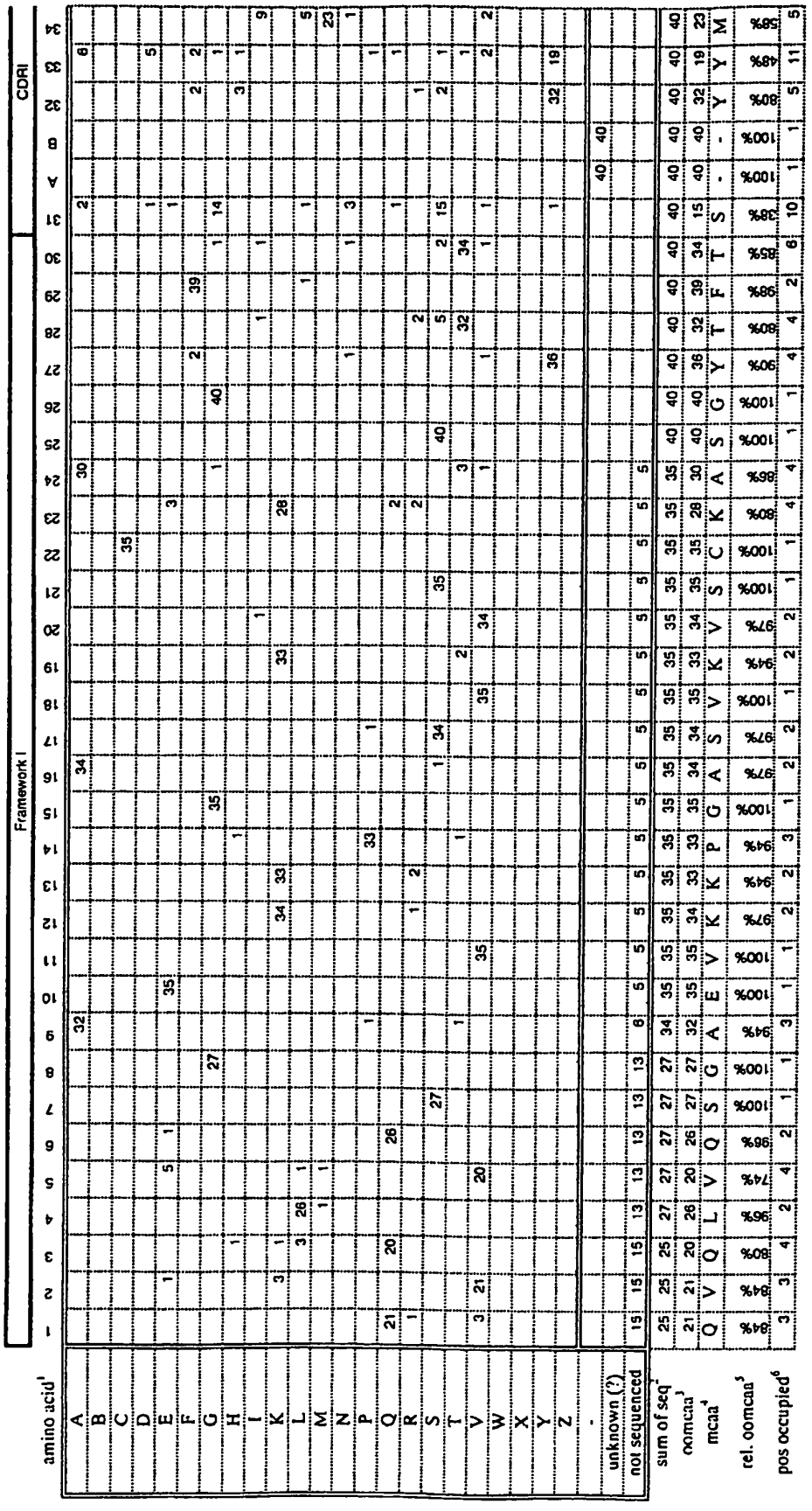


Table 6B: Analysis of V heavy chain subgroup 1B

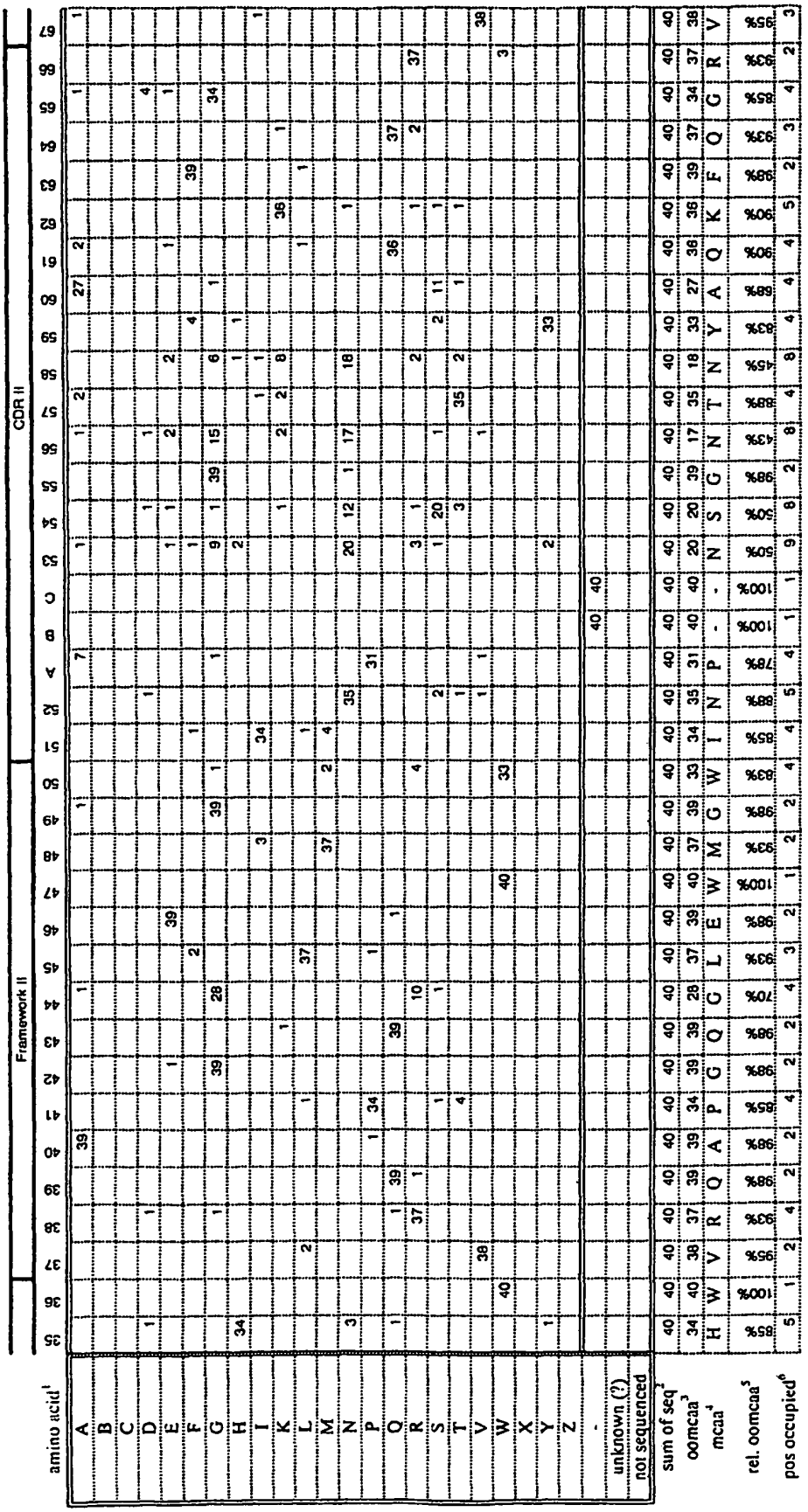


Table 6B: Analysis of V heavy chain subgroup 1B

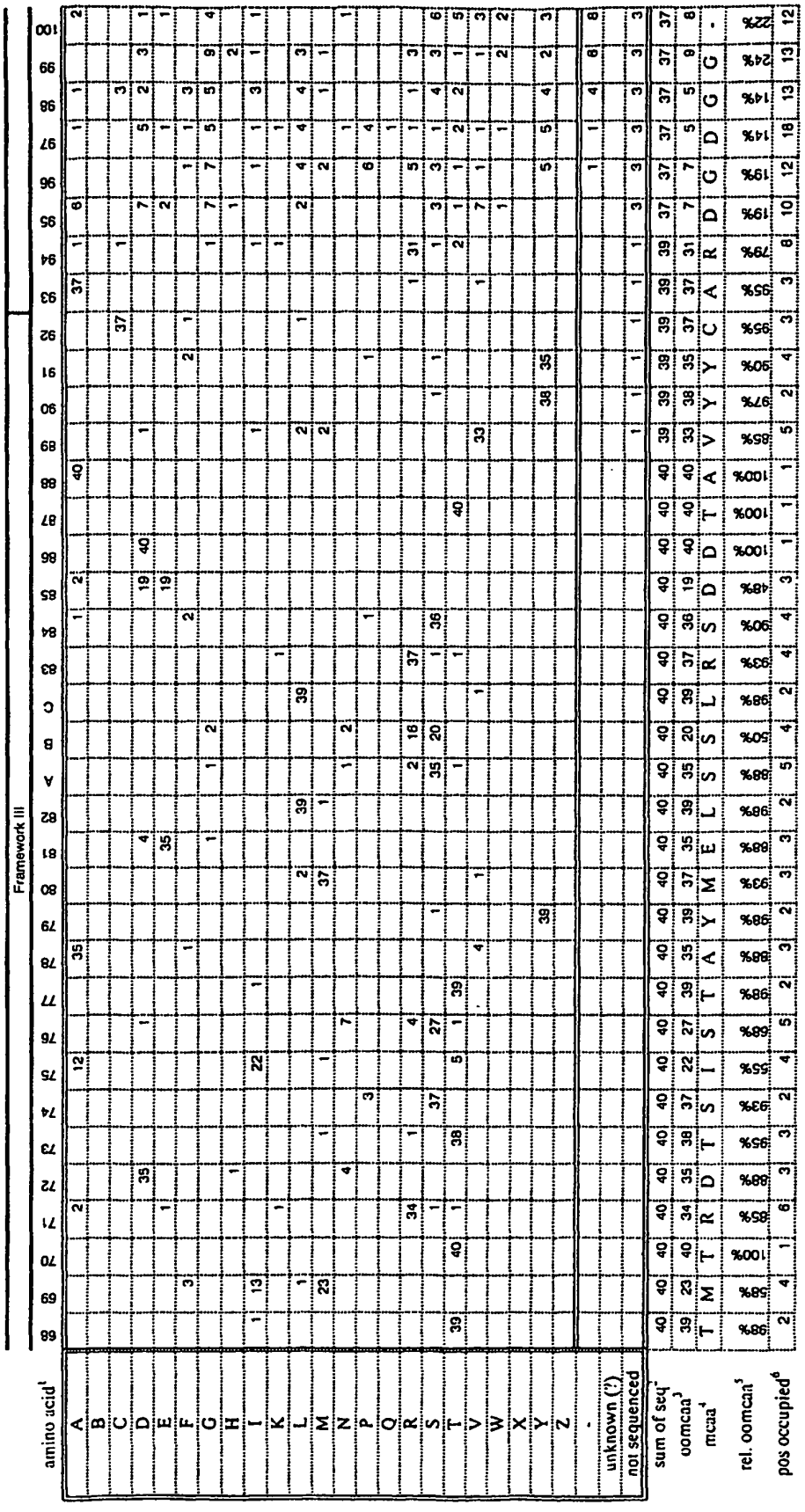


Table 6B: Analysis of V heavy chain subgroup 1B

amino acid[1]	A	B	C	D	E	F	G	H	I	J	K	101	102	103	104	105	106	107	108	109	110	111	112	113	sum
	CDR III												Framework IV												
A	3	1	3		1					5															340
B																									
C		2	1																						79
D	5	4		1		2	2	1	2			27	2												179
E	1		2		1		1									1									159
F	2	1	1	1	1					2	15		1												130
G	7	1	3		2	2	1		1	3		1			27		26					1			450
H		1	1										1												51
I	1	1	1	1							1		7								3				113
K	1	1		1		1			1							2									194
L		1	2	1	1	2		1			2								12			1			204
M							1				4								2						144
N		1	1	1			3		1			1	1												138
P	1	1		3	2				1				1			1									128
Q				1	2	1										23									253
R				1		1				1									1						247
S	3	2	2	1		1							3								1		18	18	432
T	1	1	1		1			1		1								21	6		16		1		390
V	1	2		1			1	2	1			1	6							21		18			342
W		1	1					1		4				29											158
X																									
Y		4	3	3	2	1	2	5	6	2			11												294
Z																									
-	10	11	14	20	23	25	25	25	23	18	11	6	3												394
unknown (?)											3														3
not sequenced	4	4	4	4	4	4	4	4	4	4	4	4	4	11	13	13	14	19	19	19	20	20	21	22	458
sum of seq[2]	36	36	36	36	36	36	36	36	36	36	36	36	36	29	27	27	26	21	21	21	20	20	19	18	
oomcaa[3]	10	11	14	20	23	25	25	25	23	18	15	27	11	29	27	23	26	21	12	21	16	18	18	18	
mcaa[4]	-	-	-	-	-	-	-	-	-	-	F	D	Y	W	G	Q	G	T	L	V	T	V	S	S	
rel. oomcaa[5]	28%	31%	39%	56%	64%	69%	69%	69%	64%	50%	42%	75%	31%	100%	100%	85%	100%	100%	57%	100%	80%	90%	95%	100%	
pos occupied[6]	12	17	14	13	10	9	8	7	8	8	5	5	10	1	1	4	1	1	4	1	3	3	2	1	

Table 6C: Analysis of V heavy chain subgroup 2

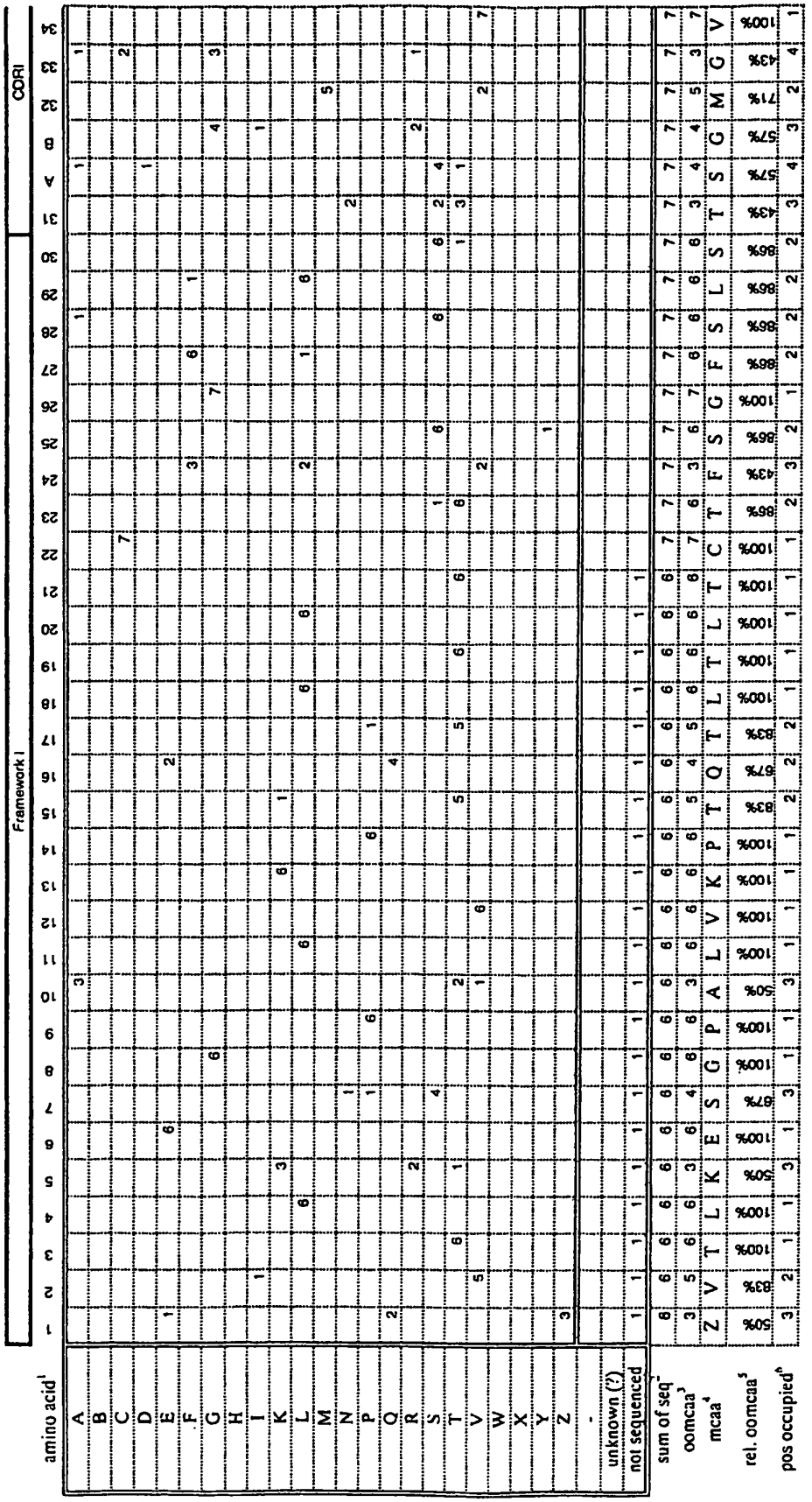


Table 6C: Analysis of V heavy chain subgroup 2

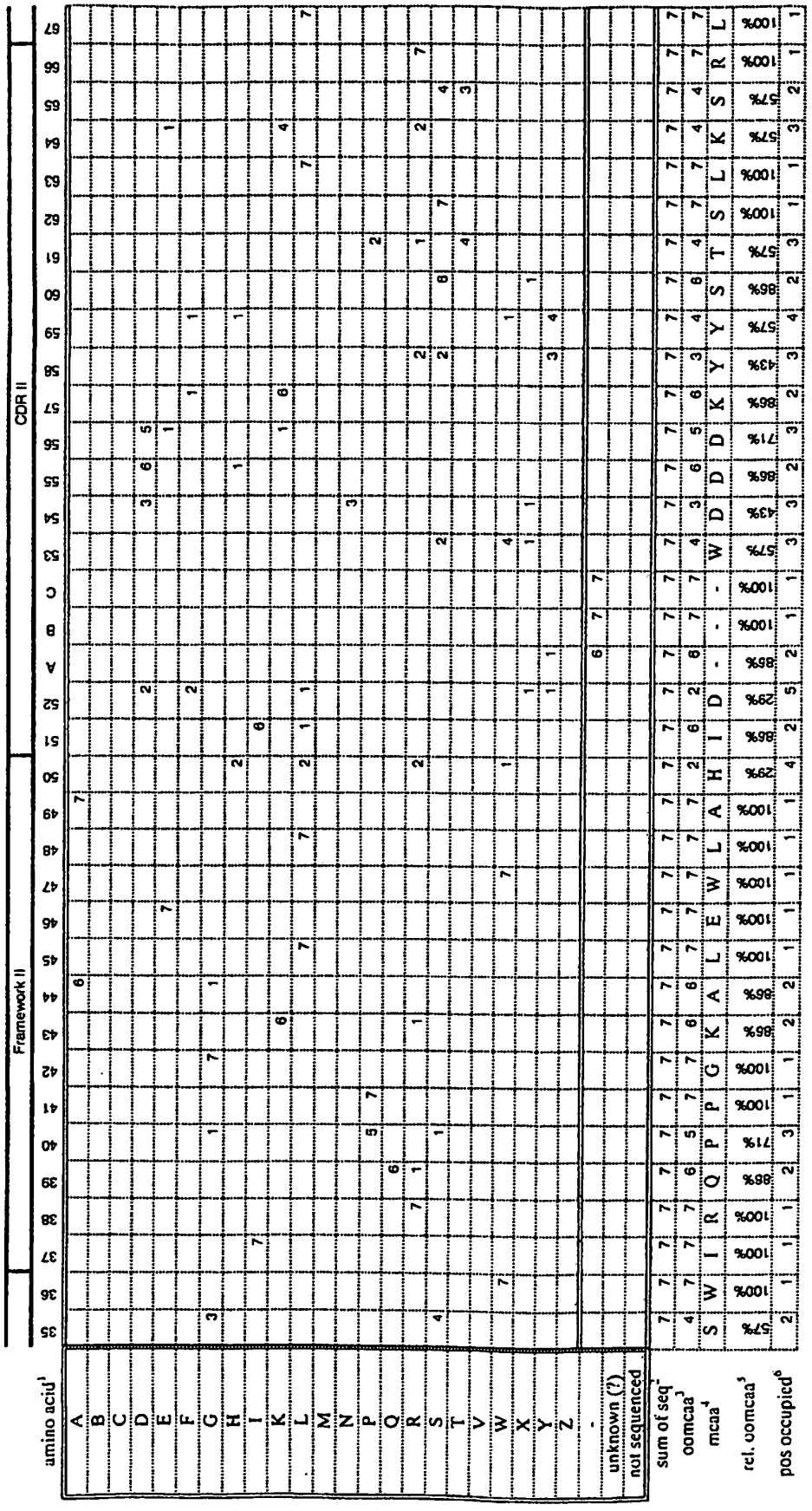


Table 6C: Analysis of V heavy chain subgroup 2

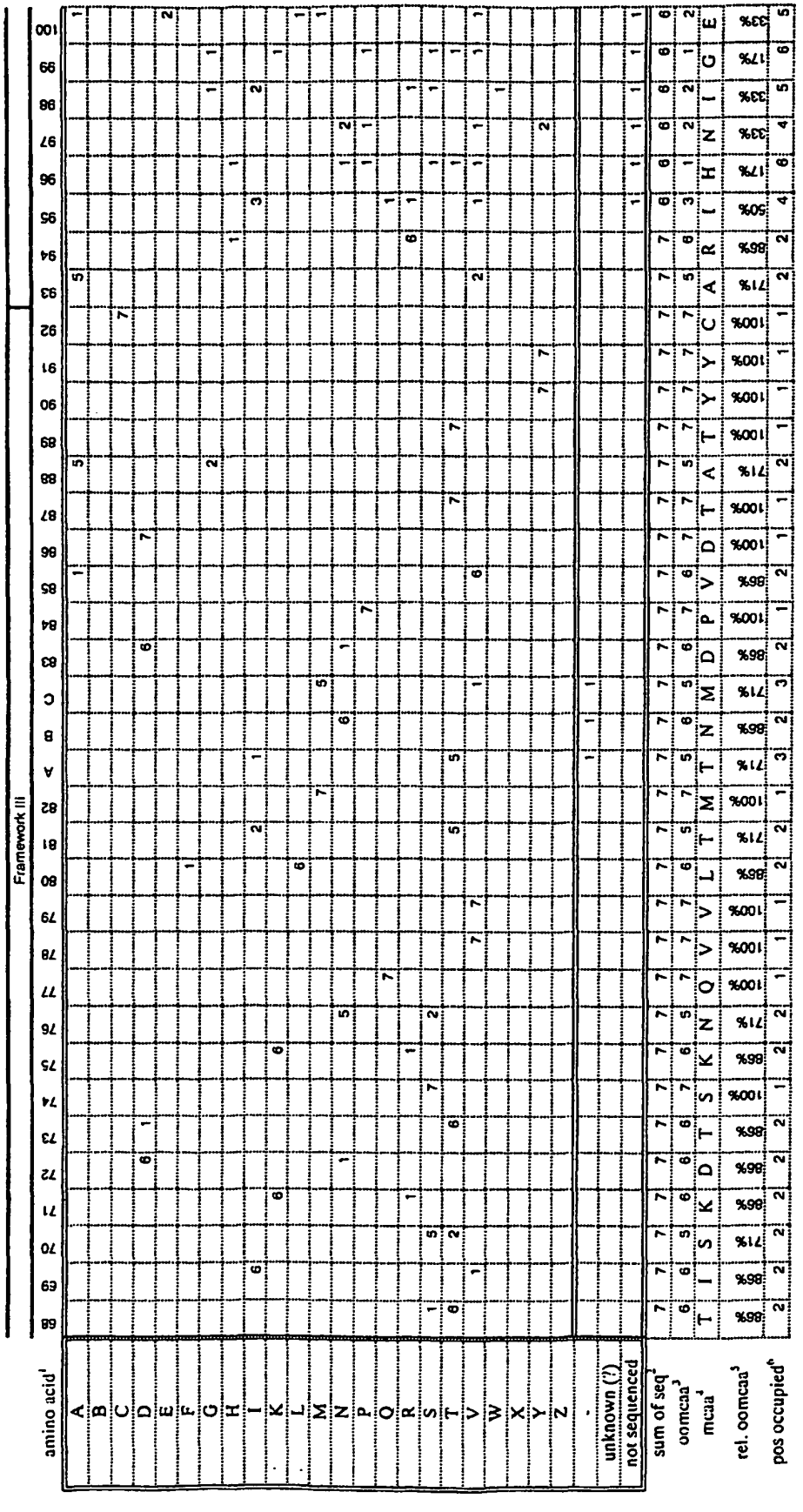


Table 6C: Analysis of V heavy chain subgroup 2

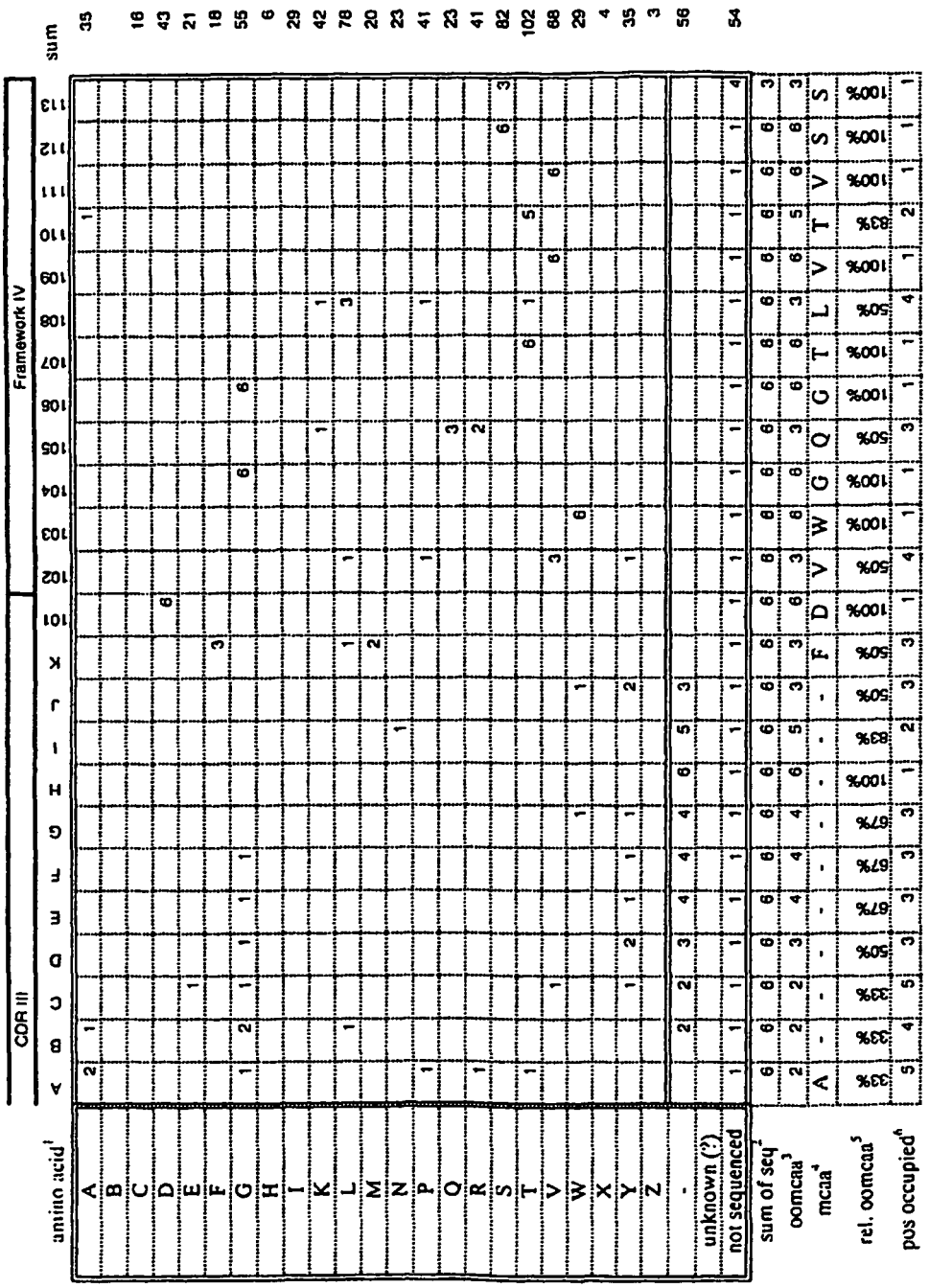


Table 6D: Analysis of V heavy chain subgroup 3

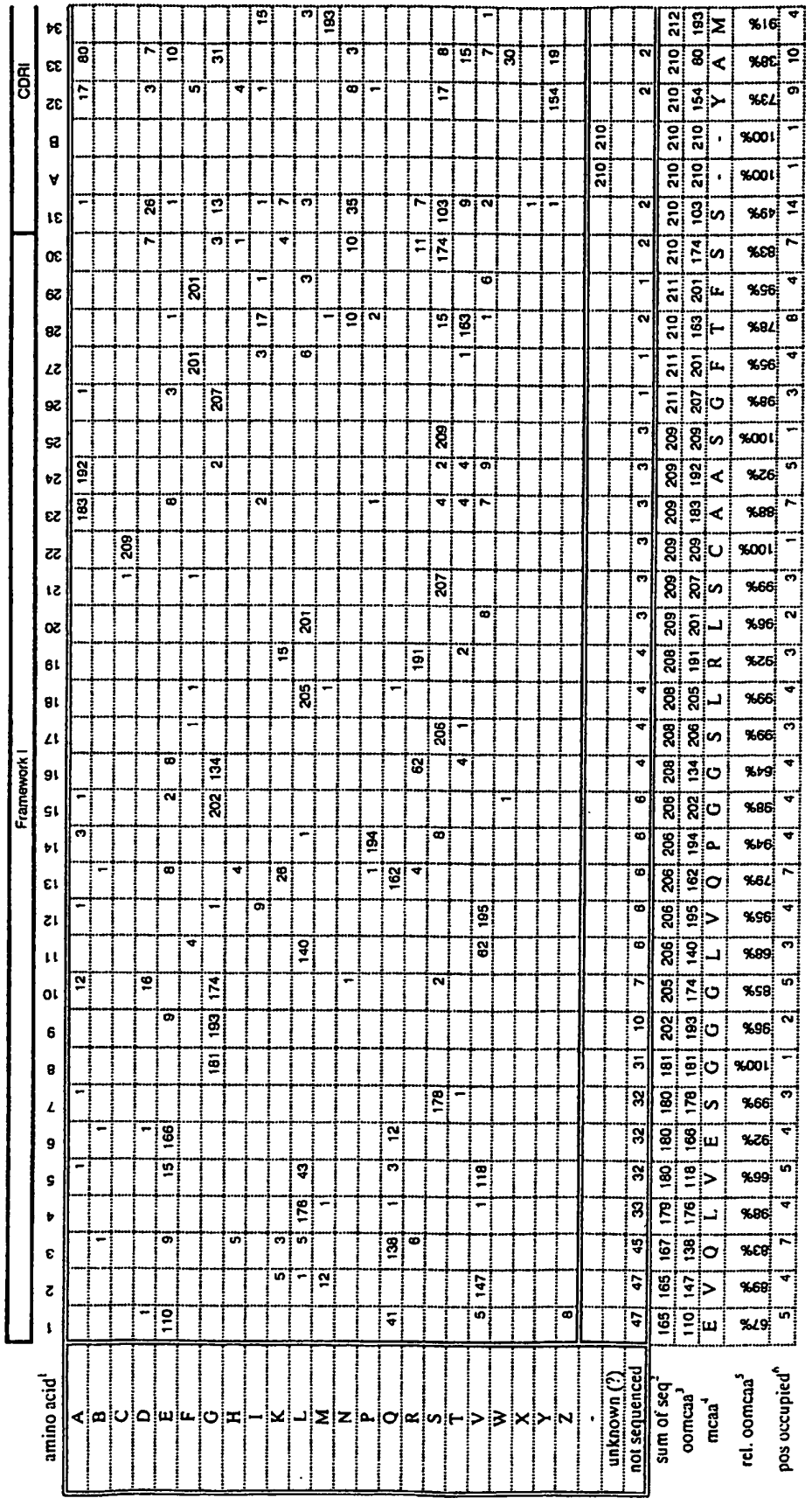


Table 6D: Analysis of V heavy chain subgroup 3

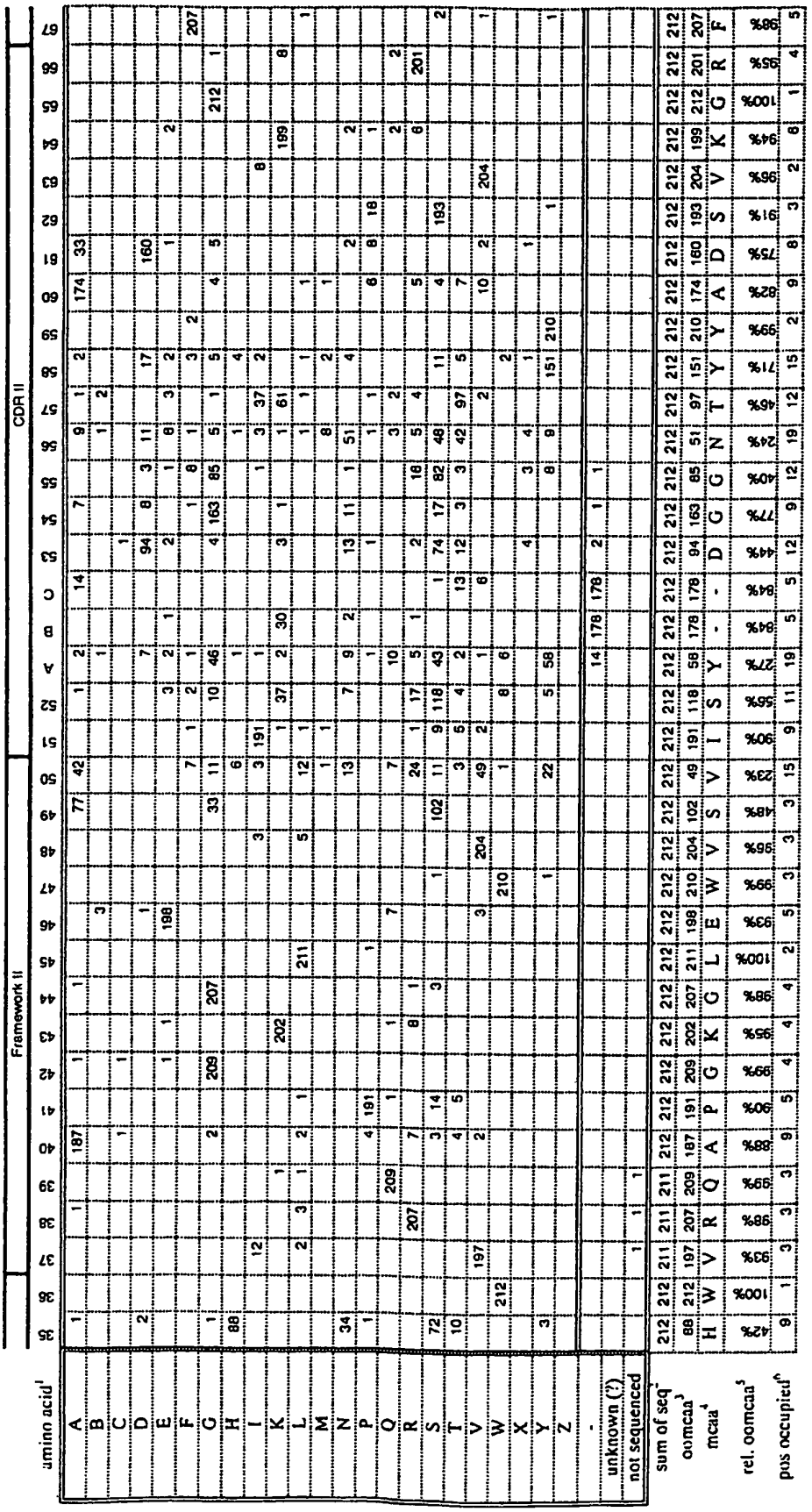


Table 6D: Analysis of V heavy chain subgroup 3

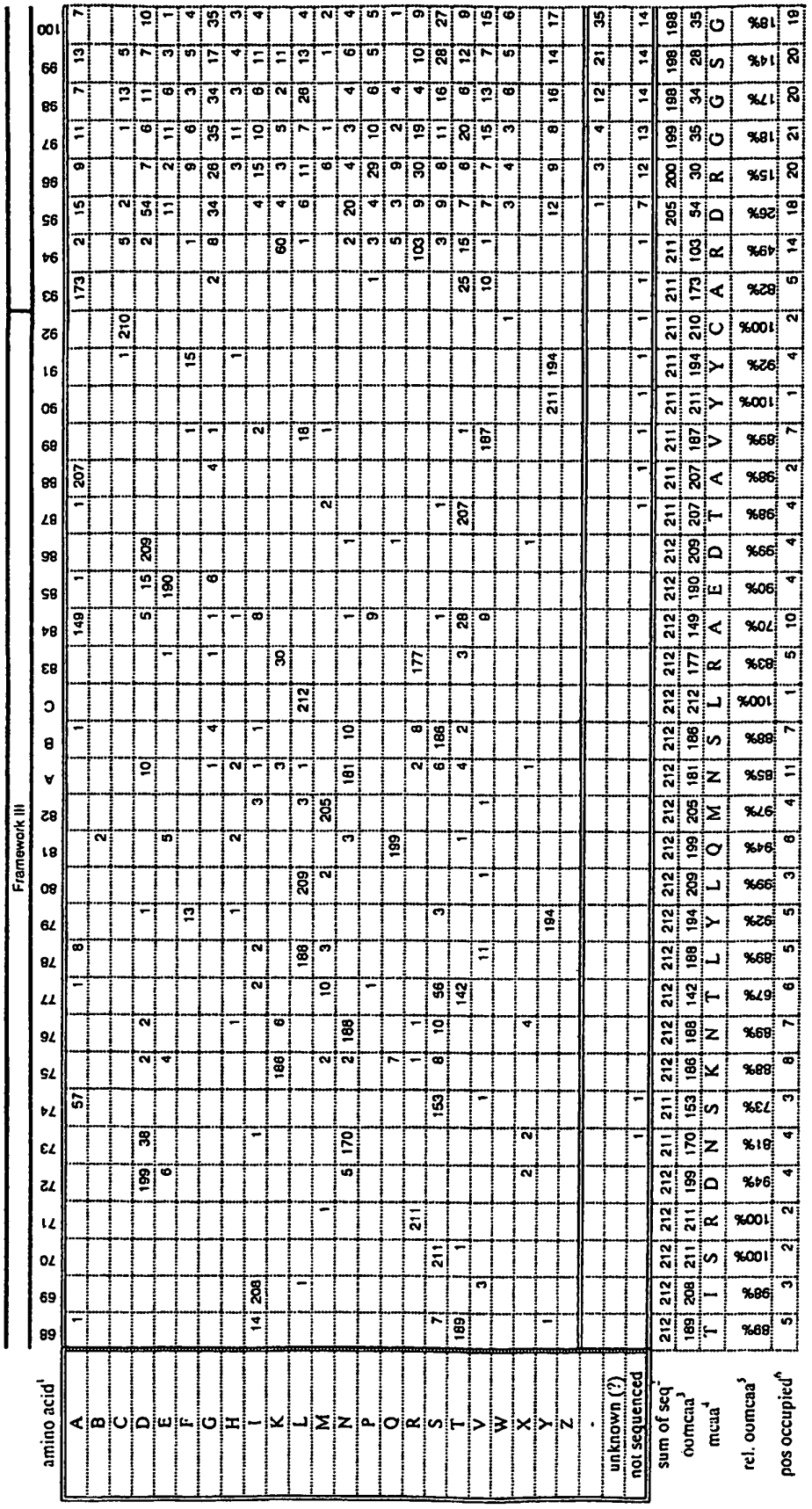
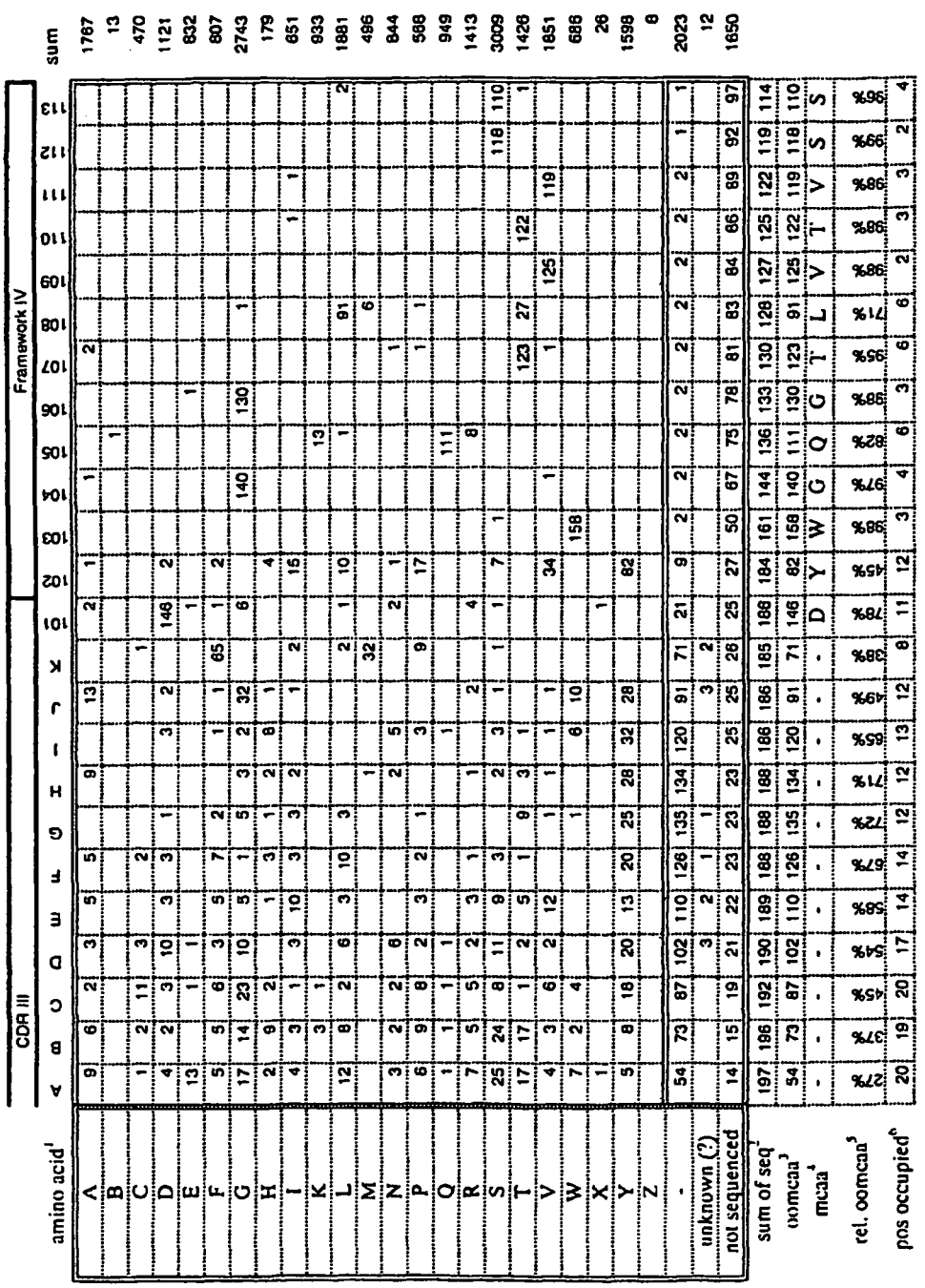


Table 6D: Analysis of V heavy chain subgroup 3



amino acid[1]	CDR III													Framework IV												sum
	A	B	C	D	E	F	G	H	I	K	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	-	
A	9	6	2	3	5	5		9		13				2	1					2		1				1767
B																										13
C	1	2	11	3		2					1															470
D	4	2	3	10	3	3	1							146	2											1121
E	13		1	1										1						1						832
F	5	5	6	3	5	7	2			1	65			1	2											807
G	17	14	23	10	5	1	5	3		2	32			6						1						2743
H	2	9	2		1	3	1	2		8	1				4											179
I	4	3	1	3	10	3	3	2			2				15											651
K		3	1																							933
L	12	8	2	6	3	10	3				2	32		2	1	10				91					2	1881
M								1												6						496
N	3	2	2	6				2		5				2	1					1						844
P	6	9	8	2	3	2	1			3				9		17				1						568
Q	1	1	1	1						1										1						949
R	7	5	5	2	3	1		1		2				4												1413
S	25	24	8	11	9	3		2		3	1			1	1	7	1									3009
T	17	17	1	2	5	1	9	3		1										123	27					118	110
V	4	3	6	2	12		1	1		1	1					34	1			1						1426
W	7	2	4				1			6	10						158									1851
X	1														1											686
Y	5	8	18	20	13	20	25	28	32	28						82										1598
Z																										8
-	54	73	87	102	110	126	135	134	120	91	71	21	9	2	2	2	2	2	2	2	2	2	2	2	1	2023
unknown (?)				3	2	1	1			3	2															12
not sequenced	14	15	19	21	22	23	23	23	25	25	26	25	27	25	27	50	67	75	78	81	83	84	86	89	92	97	1650
sum of seq[2]	197	196	192	190	189	188	188	188	186	186	185	188	184	161	144	136	133	130	128	127	125	122	119	114		
oomcan[3]	54	73	87	102	110	126	135	134	120	91	71	146	82	158	140	111	130	123	91	125	122	119	118	110		
mcaa[4]	-	-	-	-	-	-	-	-	-	-	2	-	Y	W	G	Q	G	T	L	V	T	V	S	S		
rel. oomcan[5]	27%	37%	45%	54%	58%	67%	72%	71%	65%	49%	38%	78%	45%	86%	97%	82%	98%	95%	91%	98%	98%	98%	99%	96%		
pos occupied[6]	20	19	20	17	14	14	12	12	13	12	8	11	12	3	4	6	3	6	6	2	3	3	2	4		

Table 6E: Analysis of V heavy chain subgroup 4

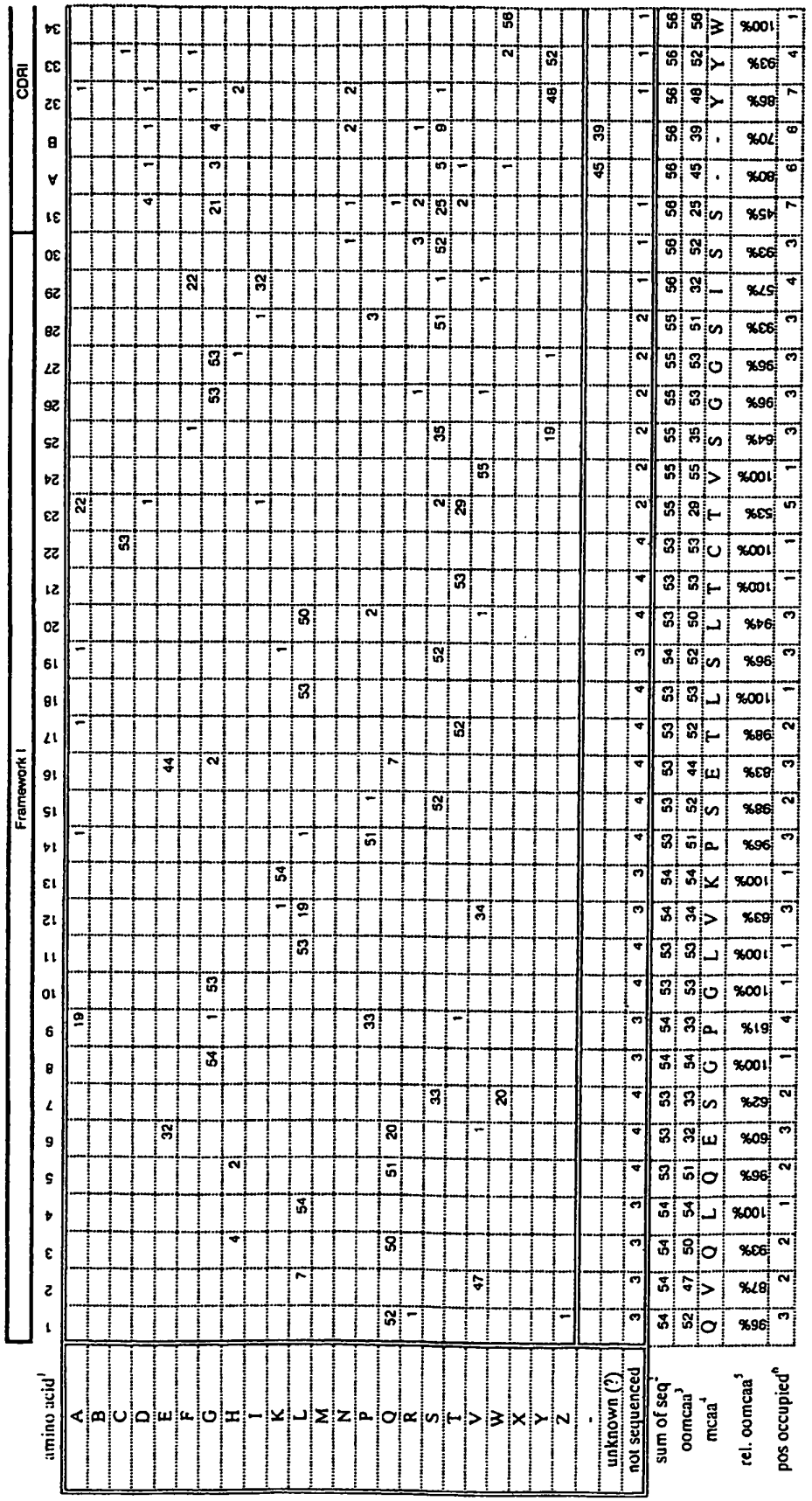


Table 6E: Analysis of V heavy chain subgroup 4

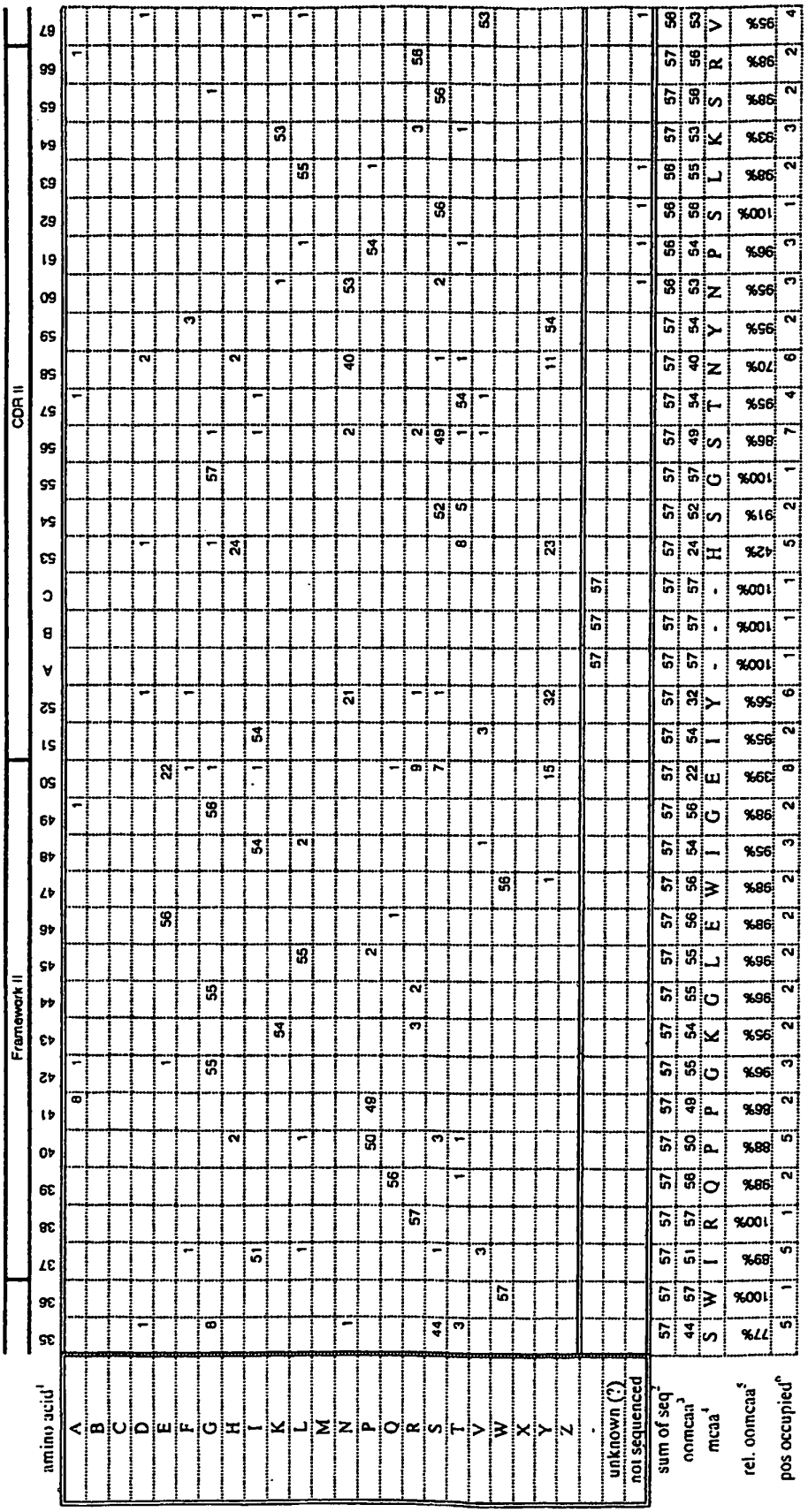


Table 6E: Analysis of V heavy chain subgroup 4

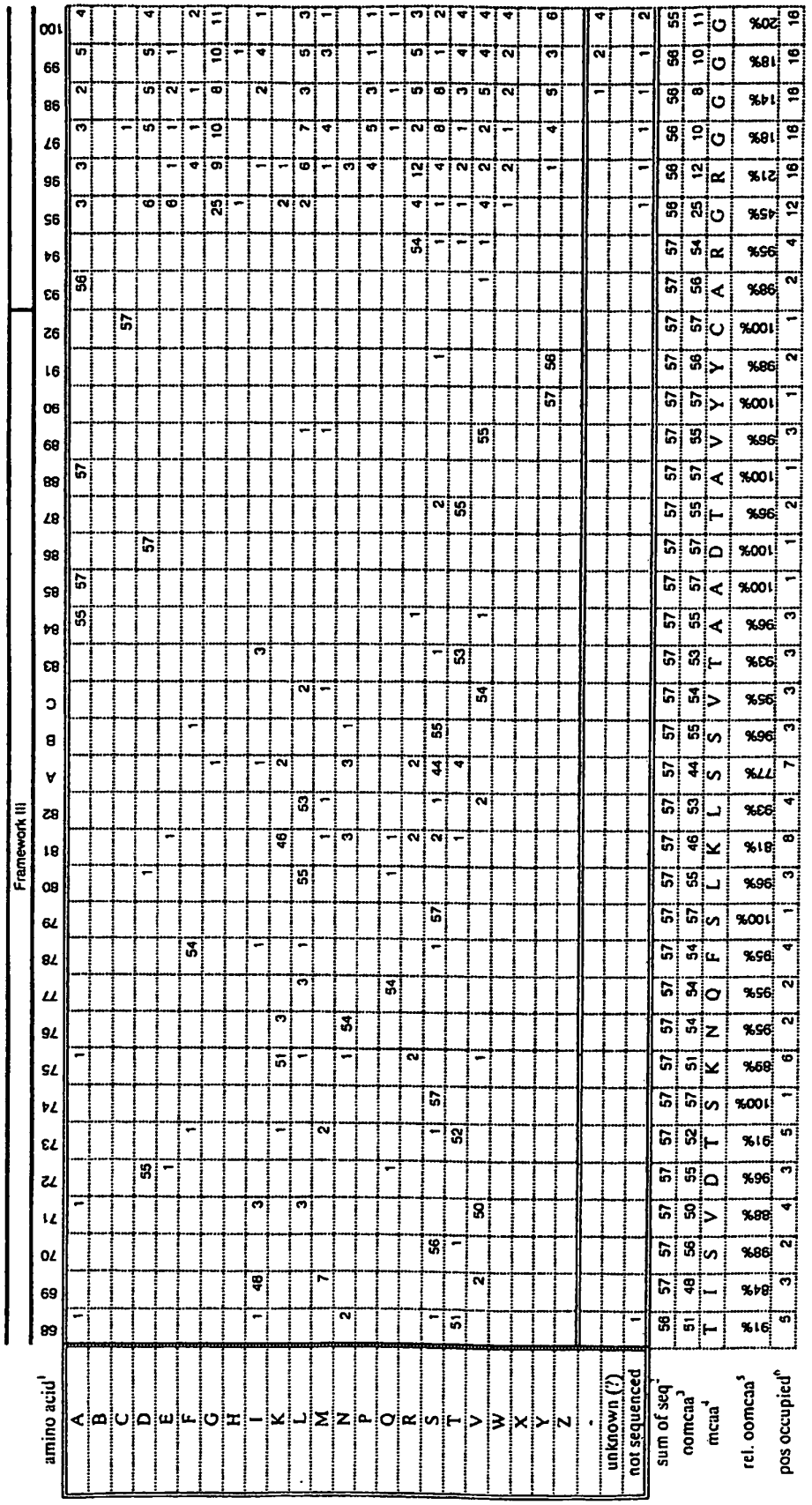


Table 6E: Analysis of V heavy chain subgroup 4

	CDR III											Framework IV													
amino acid[1]	A	B	C	D	E	F	G	H	I	J	K	101	102	103	104	105	106	107	108	109	110	111	112	113	sum
A	2	2	4		2	1		1	1	12								1			1				332
B																									
C	1																								113
D	3	2	4	3	1		1	2	1			41													210
E		1	3	1	2	1																			176
F	3	2	2		1	1					31														135
G	4	7	7	6	1	1	1	2	1	9					41		40	1							674
H					1			1				2	1								1				45
I	3	2	3		1						1		9					1							282
K		2	2			1										3									278
L	2	4	1	5	3	3		1					4						19						540
M		2	1								9								9						43
N	2	1	1	5	1	1			2									1							204
P	2	1	1	1	2	3	1	2	1				3			2								2	281
Q			1	1			3					1				29									334
R	2	3	1	2			2	1					1			4			1						250
S	5	7	4	2	1	1	1						1			1							36	33	986
T	3	3			1	1	1									1		33	8		34				532
V	7	3	1	2	1								12							36		36			488
W	5	1	1	2		2	1		3	2				46											267
X																									
Y	4	2	3	4	8	4	8	3	5	8		2	16												455
Z																									1
-	6	9	11	16	23	27	29	34	31	14	4														466
unknown (?)						1			1	1	1														4
not sequenced	3	3	6	7	8	9	9	10	11	11	11	11	10	11	16	17	17	20	20	21	21	21	21	22	426
sum of seq[2]	54	54	51	50	49	48	48	47	46	46	46	46	47	46	41	40	40	37	37	36	36	36	36	35	
oomcaa[3]	7	9	11	16	23	27	29	34	31	14	31	41	16	46	41	29	40	33	19	36	34	36	36	33	
mcaa[4]	V	-	-	-	-	-	-	-	-	-	F	D	Y	W	G	Q	G	T	L	V	T	V	S	S	
rel. oomcaa[5]	13%	17%	22%	32%	47%	56%	60%	72%	67%	30%	67%	89%	34%	100%	100%	73%	100%	89%	51%	100%	94%	100%	100%	94%	
pos occupied[6]	16	18	18	13	15	13	10	9	8	5	4	4	8	1	1	6	1	5	4	1	3	1	1	2	

Table 6F: Analysis of V heavy chain subgroup 5

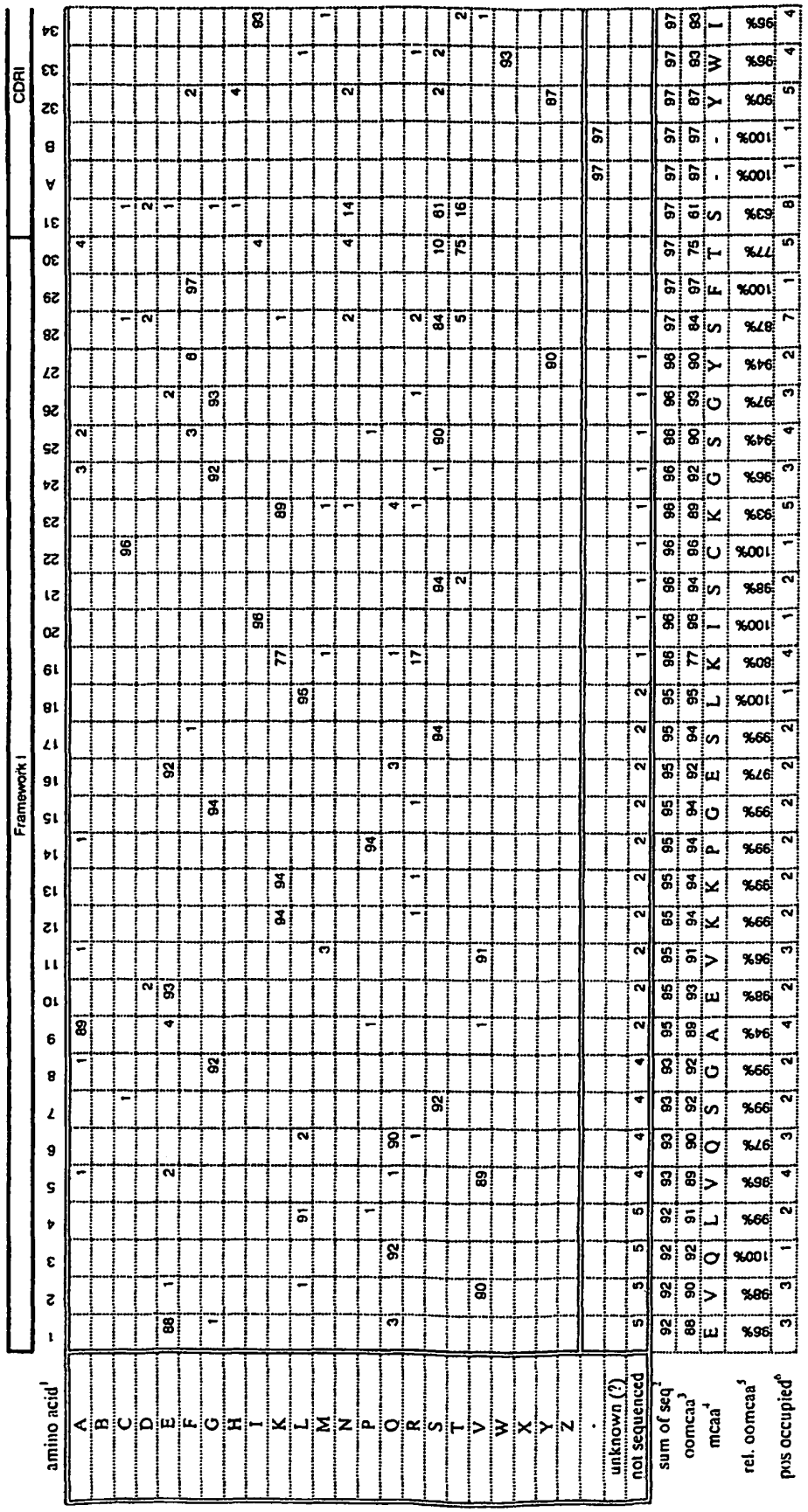


Table 6F: Analysis of V heavy chain subgroup 5

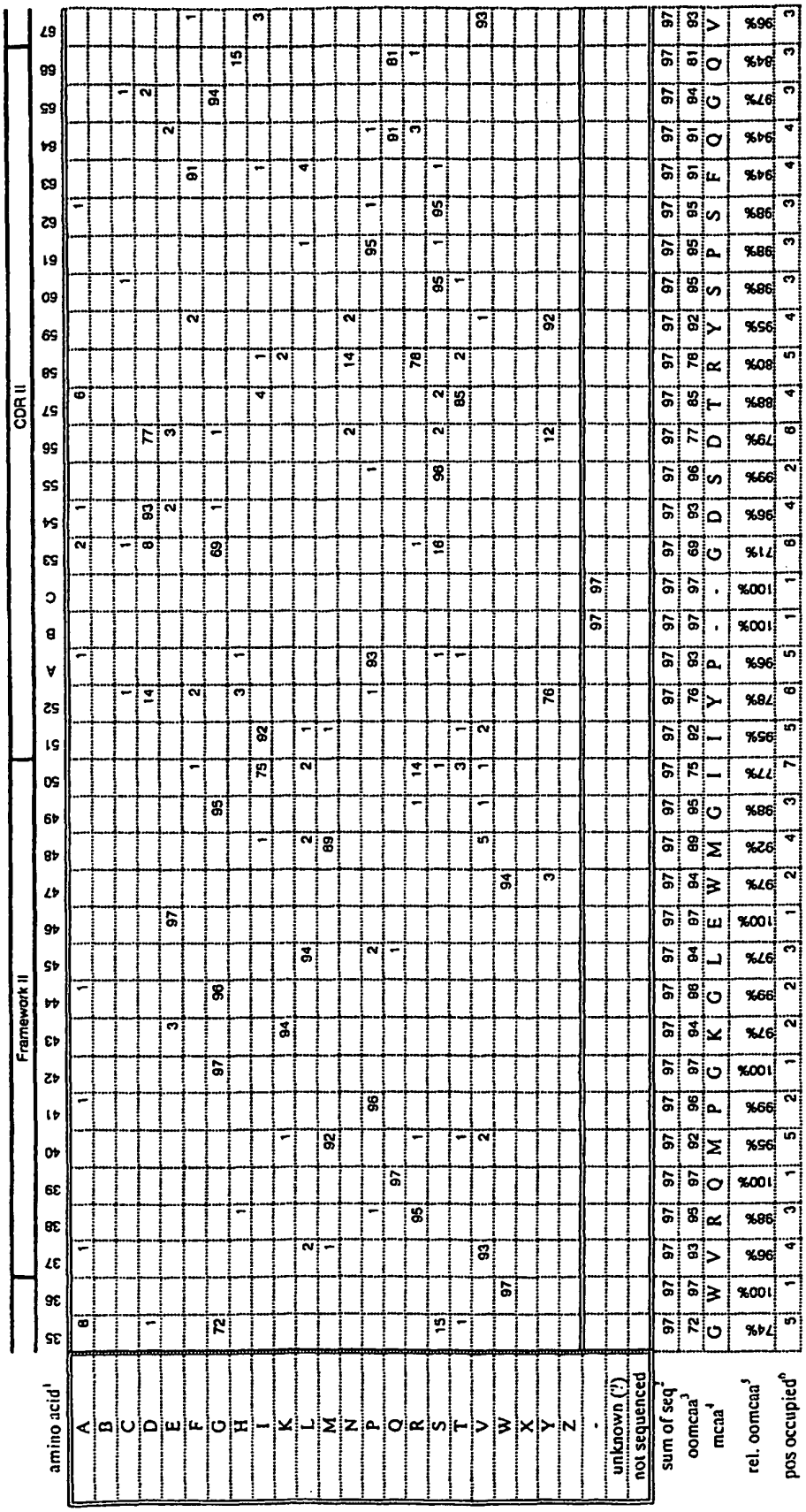


Table 6F: Analysis of V heavy chain subgroup 5

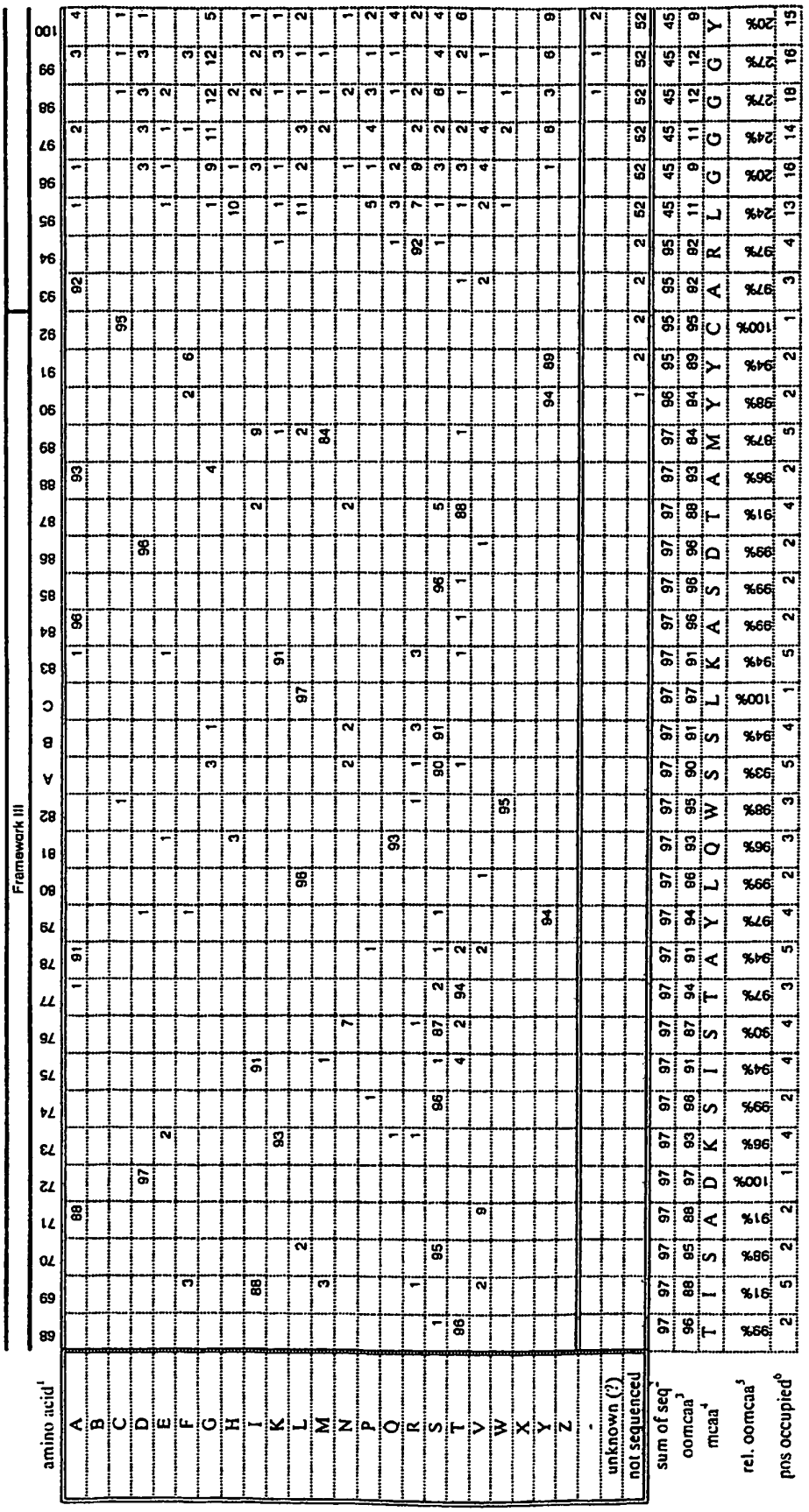


Table 6F: Analysis of V heavy chain subgroup 5

	CDR III											Framework IV													
amino acid[1]	A	B	C	D	E	F	G	H	I	J	K	101	102	103	104	105	106	107	108	109	110	111	112	113	sum
A	3	2		1			1			4		2												1	611
B																									
C			2		1																				205
D	2	1	1	2		2	1	1	2			37	1												458
E	1	1				1			1							1									404
F		3	2		1						26		2												256
G	2	4	3	10	2	1				5					41		41								1065
H	1	1		1																					44
I	1	4	1	1		1	1						9								2				588
K								2								3									650
L	5		1		1		1						2						25	1					549
M	1	1	1	1							10								8						303
N	1	2			1					2															64
P				1	1	1	1						2					1					1		414
Q	2	1	2									3				34									612
R	1		2													3									351
S	5	3	5	3	2	2			1		1		2										40	39	1545
T	3	3	6	1		1							1					40	8		39				604
V	1	2			1								11							40		41			594
W			1		2		1		1	1				43											432
X																									
Y	8	7	2	1	2	6	8	9	9	10		1	13												738
Z																									
-	8	10	16	23	30	30	31	32	30	22	7	2	2												635
unknown (?)					1			1	1	1															4
not sequenced	52	52	52	52	52	52	52	52	52	52	53	52	52	54	56	56	56	56	56	56	56	56	56	57	1678
sum of seq[2]	45	45	45	45	45	45	45	45	45	45	44	45	45	43	41	41	41	41	41	41	41	41	41	40	
oomcaa[3]	8	10	16	23	30	30	31	32	30	22	26	37	13	43	41	34	41	40	25	40	39	41	40	39	
mcaa[4]	Y	-	-	-	-	-	-	-	-	-	F	D	Y	W	G	Q	G	T	L	V	T	V	S	S	
rel. oomcaa[5]	18%	22%	36%	51%	67%	67%	69%	71%	67%	49%	59%	82%	29%	100%	100%	83%	100%	98%	61%	98%	95%	100%	98%	98%	
pos occupied[6]	16	15	14	11	11	9	8	4	6	6	4	5	10	1	1	4	1	2	3	2	2	1	2	2	

Table 6G: Analysis of V heavy chain subgroup 6

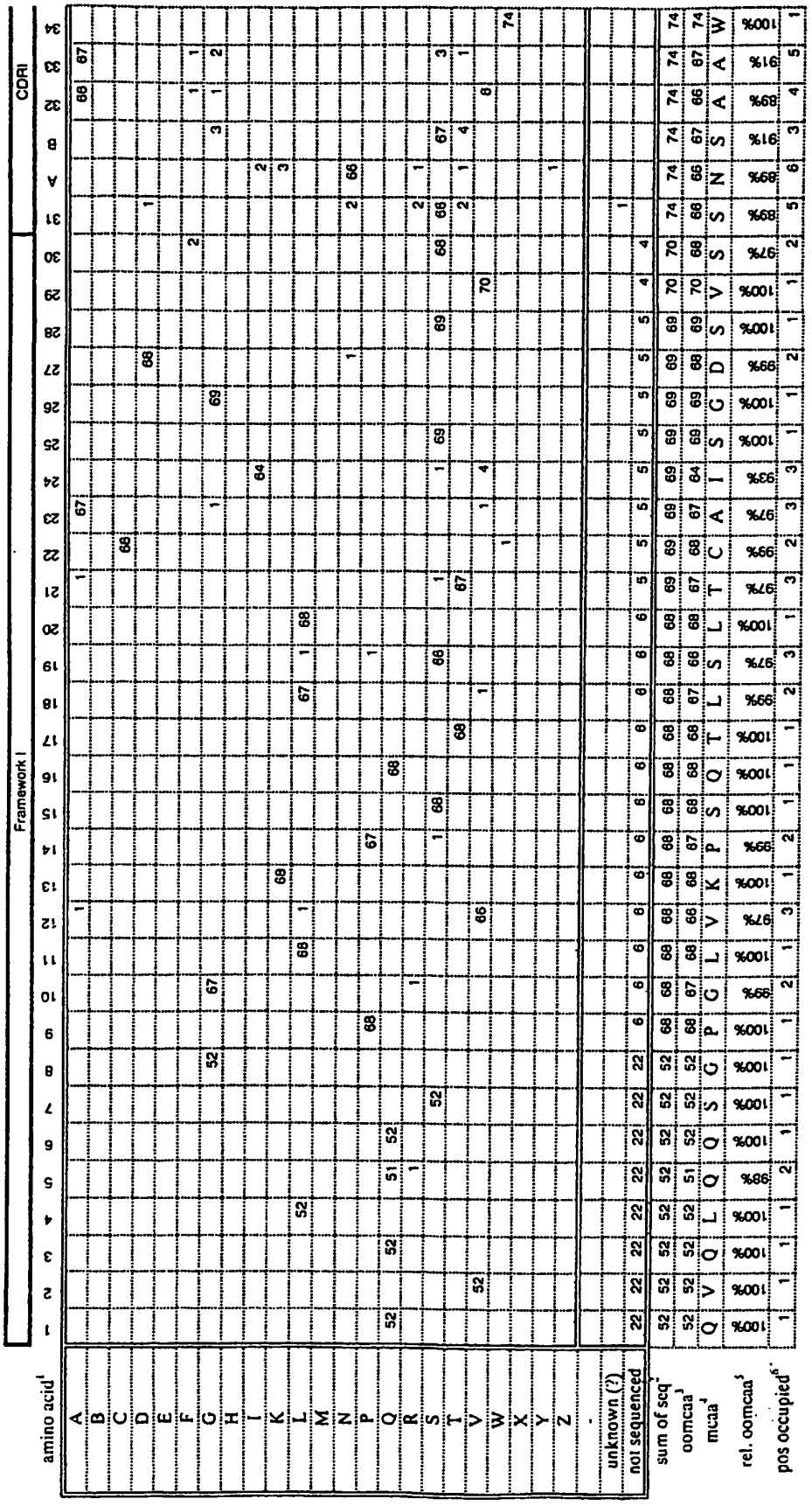


Table 6G: Analysis of V heavy chain subgroup 6

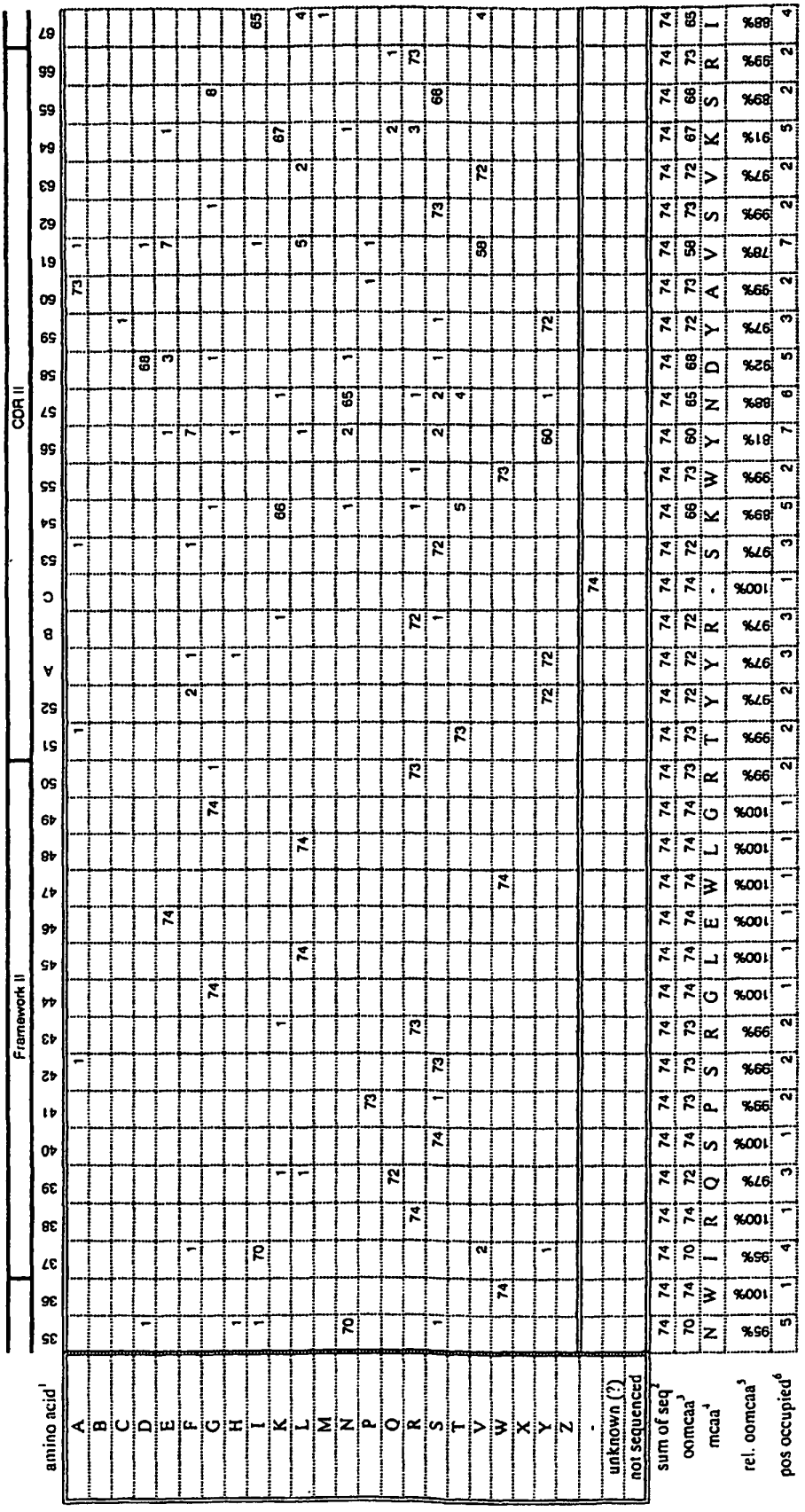


Table 6G: Analysis of V heavy chain subgroup 6

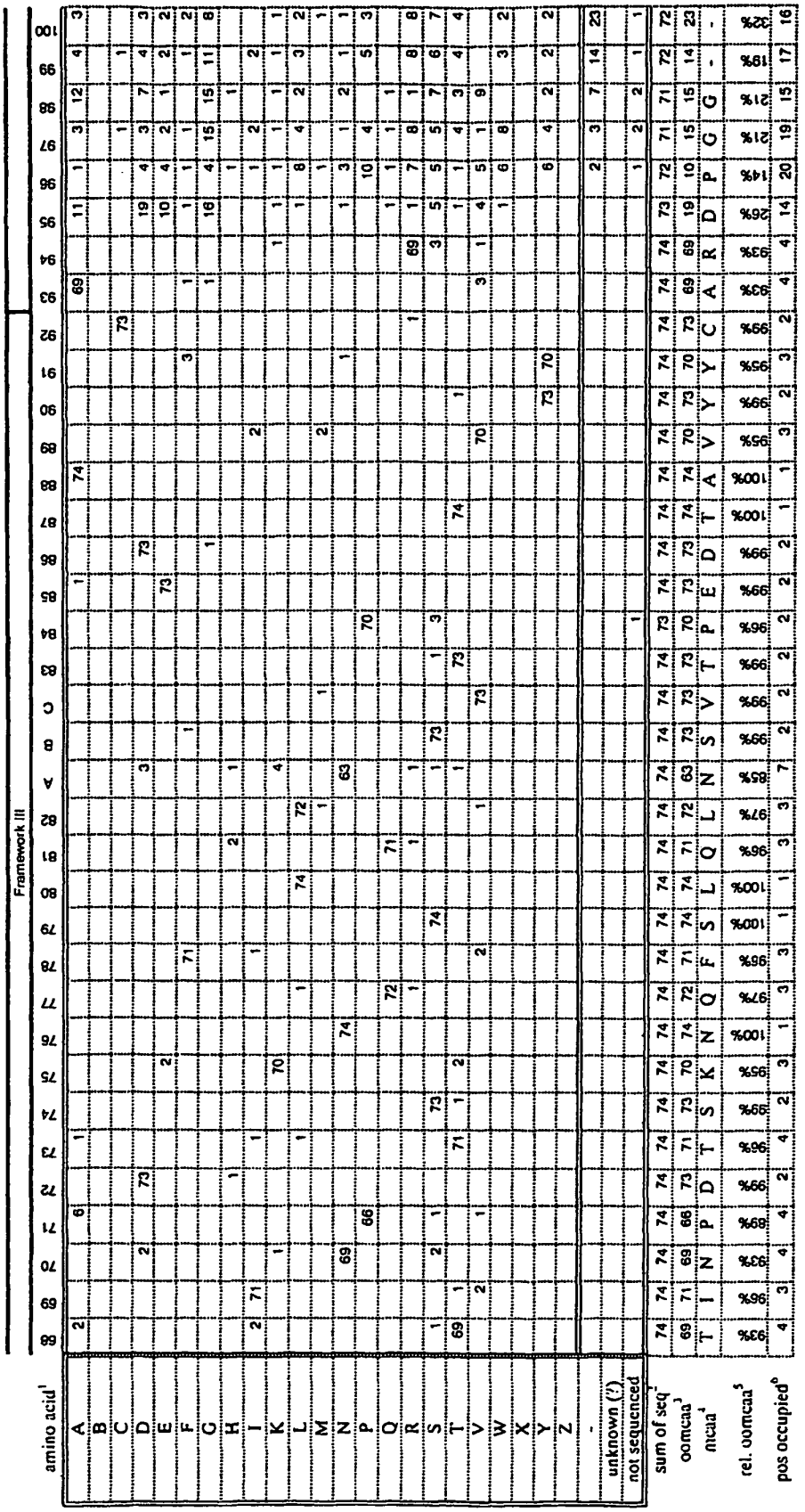
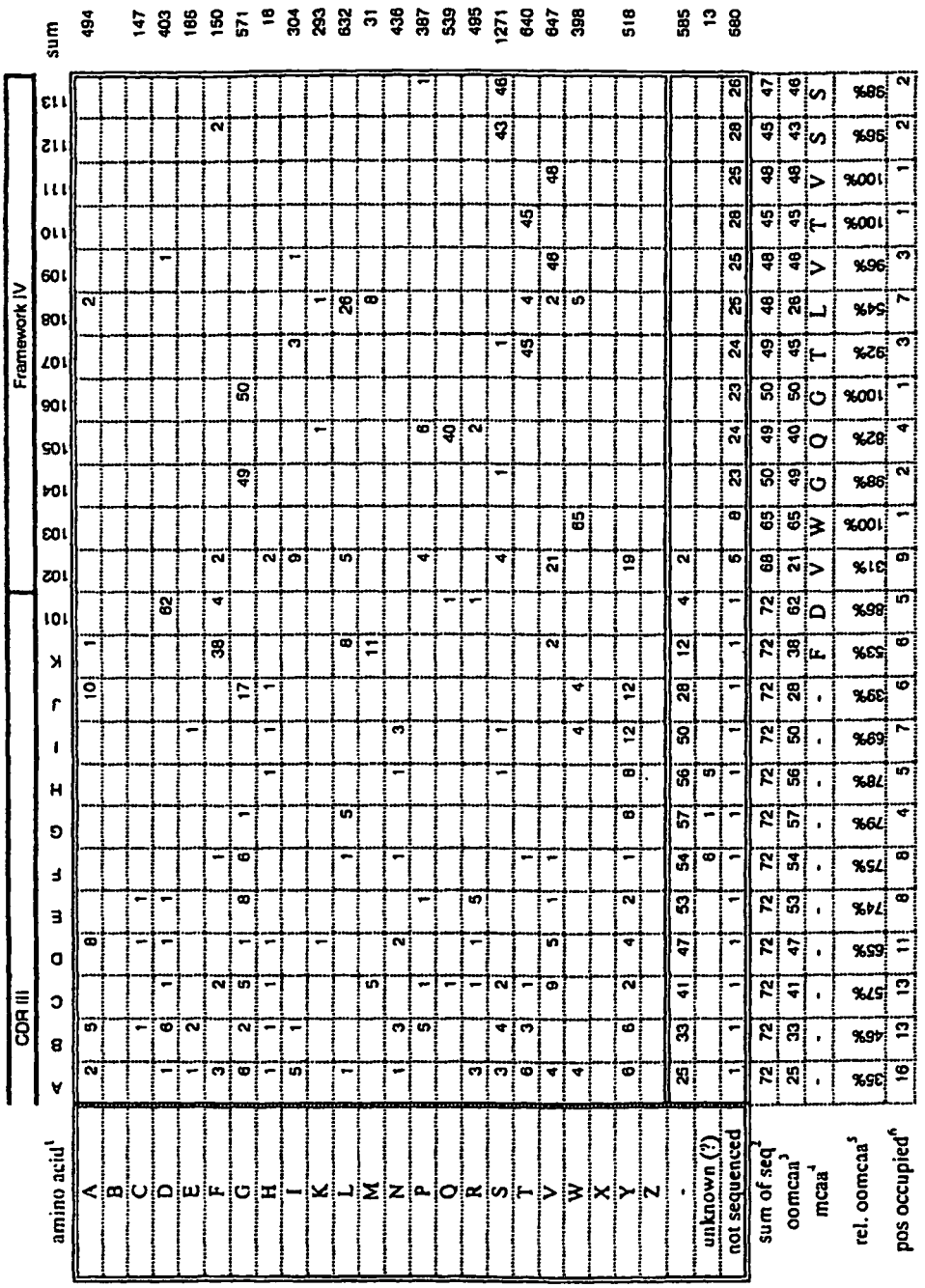


Table 6G: Analysis of V heavy chain subgroup 6



Appendix to Tables 1A-C

A. References of rearranged sequences

References of rearranged human kappa sequences used for alignment

[0146]

1 Alescio-Zonta, L. & Baglioni, C. (1970) Eur.J.Biochem., 15, 450-463.

2 Andrews, D.W. & Capra, J.D. (1981) Biochemistry, 20, 5816-5822.

3 Andris, J.S., Ehrlich, P.H., Ostberg, L. & Capra, J.D. (1992) J.Immunol., 149, 4053-4059.

4 Atkinson, P.M., Lampman, G.W., Furie, B.C., Naparstek, Y., Schwartz, R.S., Stollar, B.D. & Furie, B. (1985) J.Clin.Invest., 75, 1138-1143.

5 Aucouturier, P., Bauwens, M., Khamlichi, A.A., Denoroy, L, Spinelli, S., Touchard, G., Preud'homme, J.-L. & Cogne, M. (1993) J.Immunol., 150, 3561-3568.

6 Avila, M.A., Vazques, J., Danielsson, L., Fernandez De Cossio, M.E. & Borrebaeck, C.A.K. (1993) Gene, 127, 273-274.

7 Barbas, C.F., III, Crowe, Jr., J.E., Cababa, D., Jones, T.M., Zebedee, S.L., Murphy, B.R., Chanock, R.M. & Burton, D.R. (1992) Proc.Natl.Acad.Sci.USA, 89, 10164-10168.

8 Barbas, C.F., III, et al. (1993) J-Mol-Biol., 230, 812-23.

9 Bentley, D.L. & Rabbitts, T.H. (1980) Nature, 288, 730-733.

10 Bentley, D.L. & Rabbitts, T.H. (1983) Cell, 32, 181-189.

11 Bentley, D.L. (1984) Nature, 307, 77-80.

12 Bhat, N.M., Bieber, M.M., Chapman, C.J., Stevenson, F.K. & Teng, N.N.H. (1993) J.Immunol., 151, 5011-5021.

13 Blaison, G., Kuntz, J.-L. & Pasquali, J.-L. (1991) Eur.J.Immunol., 21, 1221-1227.

14 Braun, H., Leibold, W., Barnikol, H.U. & Hilschmann, N. (1971) Z.Physiol.Chem., 352, 647-651; (1972) Z.Physiol.Chem., 353, 1284-1306.

15 Capra, J.D. & Kehoe, J.M. (1975) Adv.Immunology, 20, 1-40. ; Andrews, D.W. & Capra. J.D. (1981) Proc.Nat.Acad.Sci.USA, 78, 3799-3803.

16 Capra, J.D. & Kehoe, J.M. (1975) Adv.Immunology, 20, 1-40. ; Ledford. D.K., Goni, F., Pizzolato, M., Franklin, E.C., Solomon, A. & Frangione, B. (1983) J.Immunol., 131, 1322-1325.

17 Chastagner, P., Theze, J. & Zouali, M. (1991) Gene, 101, 305-306.

18 Chen, P.P., Robbins, D.L., Jirik, F.R., Kipps, T.J. & Carson, D.A. (1987) J.Exp.Med, 166, 1900-1905.

19 Chen, P.P., Robbins, D.L., Jirik, F.R., Kipps, T.J. & Carson, D.A. (1987) J.Exp.Med, 166, 1900-1905; Liu, M.-F., Robbins, D.L., Crowley, J.J., Sinha, S., Kozin, F., Kipps, T.J., Carson, D.A. & Chen.P.P. (1989) J.Immunol., 142, 688-694.

20 Chersi, A. & Natali, P.G. (1978) Immunochemistry, 15, 585-589.

21 Co, M.S., Deschamps, M., Whitley, R.J. & Queen, C. (1991) Proc.Natl.Acad.Sci.USA, 88, 2869-2873.

22 Cuisinier, A.-M., Fumoux, F., Fougereau, M. & Tonnelle, C. (1992) Mol.Immunol., 29, 1363-1373.

23 Davidson, A., Manheimer-Lory, A., Aranow, C., Peterson, R., Hannigan, N. & Diamond, B. (1990) J.Clin.Invest., 85, 1401-1409.

24 Denomme, G.A., Mahmoudi, M., Edwards, J.Y., Massicotte, H., Cairns, E. & Bell, D.A. (1993) Hum.Antibod.Hybridomas, 4, 98-103.

25 Dersimonian, H., Mcadam, K.P.W.J., Mackworth-Young, C. & Stollar, B.D. (1989) J.Iinmunol., 142, 4027-4033.

26 Dreyer, W.J., Gray, W.R. & Hood, L. (1967) Cold Spring Harbor Symp. Quantitative Biol., 32, 353-367.

27 Ebeling, S.B., Schutte, M.E.M. & Logtenberg, T. (1993) Eur.J.Immunol., 23, 1405-1408.

28 Eulitz, M. & Kley, H.-P. (1977) Immunochem. 14, 289-297.

29 Eulitz, M. & Linke, R.P. (1982) Z.Physiol.Chem., 363, 1347-1358.

30 Eulitz, M., Breuer, M., Eblen, A., Weiss, D.T. & Solomon, A. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark Kluwer Academic

31 Eulitz, M., Gotze, D. & Hilschmann, N. (1972) Z.Physiol.Chem., 353, 487-491; Eulitz, M. & Hilschmann, N. (1974) Z.Physiol.Chem., 355, 842-866.

32 Eulitz, M., Kley, H.P. & Zeitler, H.J. (1979) Z.Physiol.Chem., 360, 725-734.

33 Ezaki, I., Kanda, H., Sakai, K., Fukui, N., Shingu, M., Nobunaga, M. & Watanabe, T. (1991) Arthritis And Rheumatism, 34, 343-350.

34 Felgenhauer, M., Kohl, J. & Ruker, F. (1990) Nucl.Acids Res., 18. 4927.

35 Ferri. G., Stoppini, M., Iadarola, P., Bellotti, V. & Merlini, G. (1989) Biochim.Biophys.Acta, 995, 103-108.

36 Gillies, S.D., Dorai, H., Wesolowski, J., Majeau, G., Young, D., Boyd, J., Gardner, J. & James, K. (1989) Bio/Tech., 7, 799-804.

37 Goni, F. & Frangione, B. (1983) Proc.Nat.Acad.Sci.USA, 80, 4837-4841.

38 Goni, F.R., Chen, P.P., Mcginnis, D., Arjonilla, M.L., Fernandez, J., Carson, D., Solomon, A., Mendez, E. & Frangione, B. (1989) J.Immunol., 142, 3158-3163.

39 Gorman. S.D., Clark, M.R., Rouledge, E.G., Cobbold, S.P. & Waldmann, H. (1991) Proc.Natl.Acad.Sci.USA, 88, 4181-4185.

40 Gottlieb, P.D., Cunningham, B.A., Rutishauser, U. & Edelman, G.M. (1970) Biochemistry, 9, 3155-3161.

41 Griffiths, A.D., Malmqvist, M., Marks, J.D., Bye, J.M., Embleton, M.J., Mccafferty, J., Baier, M., Holliger, K.P., Gorick, B.D., Hughes-Jones, N.C., Hoogenboom, H.R. & Winter, G. (1993) Embo J., 12, 725-734.

42 Hieter, P.A., Max, E.E., Seidman, J.G., Maizel, J.V., Jr. & Leder, P. (1980) Cell, 22, 197-207; Klobeck, H.G, Meindl, A., Combriato, G., Solomon, A. & Zachau, H.G. (1985) Nucl.Acids Res., 13, 6499-6513; Weir, L. & Leder, P. (1986)

43 Hilschmann, N. & Craig, L.C. (1965) Proc.Nat.Acad.Sci.USA, 53, 1403-1409; Hilschmann, N. (1967) Z.Pliysiol.Chem., 348, 1077-1080.

44 Hilschmann, N. & Craig, L.C. (1965) Proc.Nat.Acad.Sci.USA, 53, 1403-1409; Hilschmann, N. (1967) Z.Physiol.Chem., 348, 1718-1722; Hilschmann, N. (1969) Naturwissenschaften, 56, 195-205.

45 Hirabayashi, Y., Munakata, Y., Sasaki, T. & Sano, H. (1992) Nucl.Acids Res., 20, 2601.

46 Jaenichen, H.-R., Pech, M., Lindenmaier, W., Wildgruber, N. & Zachau, H.G. (1984) Nuc.Acids Res., 12, 5249-5263.

47 Jirik. F.R., Sorge, J., Fong, S., Heitzmann, J.G., Curd, J.G., Chen, P.P., Goldfien, R & Carson, D.A. (1986) Proc.Nat.Acad.Sci.USA, 83, 2195-2199.

48 Kaplan, A.P. & Metzger, H. (1969) Biochemistry, 8, 3944-3951. ; Klapper, D.G. & Capra, J.D. (1976) Ann.Immunol. (Inst.Pasteur), 127c, 261-271.

49 Kennedy, M.A. (1991) J.Exp.Med., 173, 1033-1036.

50 Kim, H.S. & Deutsch, H.F. (1988) Immunol., 64, 573-579.

51 Kipps, T.J., Tomhave, E., Chen, P.P. & Carson, D.A. (1988) J.Exp.Med., 167, 840-852.

52 Kipps, T.J., Tomhave, E., Chen, P.P. & Fox, R.I. (1989) J.Immunol., 142, 4261-4268.

53 Klapper, D.G. & Capra, J.D. (1976) Ann.Immunol.(Inst.Pasteur), 127c, 261-271.

54 Klein, U., Kuppers, R. & Rajewsky, K. (1993) Eur.J.Immunol., 23, 3272-3277.

55 Klobeck, H.G, Meindl, A., Combriato, G., Solomon, A. & Zachau, H.G. (1985) Nucl.Acids Res., 13,6499-6513.

56 Klobeck, H.G., Bornkammm, G.W., Combriato, G., Mocikat. R., Pohlenz, H.D. & Zachau, H.G. (1985) Nucl.Acids Res., 13, 6515-6529.

57 Klobeck, H.G., Combriato, G. & Zachau, H.G. (1984) Nuc.Acids Res., 12, 6995-7006.

58 Klobeck, H.G., Solomon, A. & Zachau, H.G. (1984) Nature, 309, 73-76.

59 Knight, G.B., Agnello, V., Bonagura, V., Barnes, J.L., Panka, D.J. & Zhang, Q-X. (1993) J.Exp.Med., 178, 1903-1911.

60 Kohler, H., Shimizu, A., Paul, C. & Putnam, F.W. (1970) Science, 169, 56-59. (Kaplan, A.P. & Metzger, H. (1969) Biochemistry, 8, 3944-3951.)

61 Kratzin, H., Yang, C.Y., Krusche, J.U. & Hilschmann, N. (1980) Z.Physiol.Chem., 361, 1591-1598.

62 Kunicki, T.J., Annis, D.S., Gorski, J. & Nugent, D.J. (1991) J.Autoimmunity, 4, 433-446.

63 Larrick, J.W., Wallace, E.F., Coloma, M.J., Bruderer, U., Lang, A.B. & Fry, K.E. (1992) Immunological Reviews, 130, 69-85.

64 Laure, C.J., Watanabe, S. & Hilschmann, N. (1973) Z.Physiol.Chem., 354, 1503-1504.

65 Ledford, D.K., Goni, F., Pizzolato, M., Franklin, E.C., Solomon, A. & Frangione, B. (1983) J.Immunol., 131, 1322-1325.

66 Ledford, D.K., Goni, F., Pizzolato, M., Franklin, E.C., Solomon, A. & Frangione, B. (1983) J.Immunol., 131, 1322-1325.

67 Ledford, D.K., Goni, F., Pizzolato, M., Franklin, E.C., Solomon, A. & Frangione, B. (1983) J.Immunol., 131, 1322-1325. Pons-Estel, B., Goni, F., Solomon, A. & Frangione, B. (1984) J.Exp.Med., 160, 893.

68 Levy. S., Mendel, E., Kon, S., Avnur, Z. & Levy, R. (1988) J.Exp.Med., 168, 475-489.

69 Liepnieks, J.J., Dwulet, F.E. & Benson, M.D. (1990) Mol.Immunol., 27, 481-485.

70 Manheimer-Lory, A., Katz, J.B., Pillinger, M., Ghossein, C., Smith, A. & Diamond, B. (1991) J.Exp.Med., 174, 1639-1652.

71 Mantovani, L., Wilder, R.L. & Casali, P. (1993) J.Inununol., 151, 473-488.

72 Mariette, X., Tsapis, A. & Brouet, J.-C. (1993) Eur.J.Immunol., 23, 846-851.

73 Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., Mccafferty, J., Griffiths, A.D. & Winter. G. (1991) J.Mol.Biol., 222, 581-597.

74 Marsh, P., Mills, F. & Gould, H. (1985) Nuc.Acids Res., 13, 6531-6544.

75 Middaugh, C.R. & Litman, G.W. (1987) J.Biol.Chem., 262, 3671-3673.

76 Milstein, C. & Deverson, E.V. (1971) Biochem.J., 123, 945-958.

77 Milstein, C. (1969) Febs Letters, 2, 301-304.

78 Milstein, C. (1969) Febs Letters, 2, 301-304.

79 Milstein, C.P. & Deverson, E.V. (1974) Eur.J.Biochem., 49, 377-391.

80 Moran, M.J., Andris, J.S., Matsumato, Y.-I., Capra, J.D. & Hersh, E.M. (1993) Mol.Immunol., 30, 1543-1551.

81 Nakatani, T., Nomura, N., Horigome, K., Ohtsuka, H. & Noguchi, H. (1989) Bio/Tech., 7,805-810.

82 Newkirk, M., Chen, P.P., Carson, D., Posnett, D. & Capra, J.D. (1986) Mol.Inimunol., 23, 239-244.

83 Newkirk, M.M., Gram, H., Heinrich, G.F., Ostberg, L., Capra, J.D. & Wasserman, R.L. (1988) J.Clin.Invest., 81, 1511-1518.

84 Newkirk, M.M., Mageed, R.A., Jefferis, R., Chen, P.P. & Capra, J.D. (1987) J.Exp.Med., 166, 550-564.

85 Olee, B.T., Lu, E.W., Huang, D.-F., Soto-Gil, R.W., Deftos, M., Kozin, F., Carson, D.A. & Chen, P.P. (1992) J.Exp.Med., 175, 831-842,

86 Palin, W. & Hilschmann, N. (1973) Z.Physiol.Chem., 354, 1651-1654, (1975) Z.Physiol.Chem., 356, 167-191.

87 Pascual, V., Victor, K., Lelsz, D., Spellerberg, M.B., Hamblin, T.J., Thompson, K.M., Randen, I., Natvig, J., Capra, J.D. & Stevenson, F.K. (1991) J.Immunol., 146,4385-4391.

88 Pascual, V., Victor, K., Randen, I., Thompson, K., Steinitz, M., Forre, O., Fu, S.-M., Natvig, J.B. & Capra, J.D. (1992) Scand.J.Immunol., 36, 349-362.

89 Pech, M. & Zachau, H.G. (1984) Nuc.Acids Res., 12, 9229-9236.

90 Pech, M., Jaenichen, H.-R., Pohlenz, H.-D., Neumaier, P.S., Klobeck, H.-G. & Zachau, H.G. (1984) J.Mol.Biol., 176, 189-204.

91 Pons-Estel, B., Goni, F., Solomon, A. & Frangione, B. (1984) J.Exp.Med., 160, 893-904.

92 Portolano, S., Mclachlan, S.M. & Rapoport, B. (1993) J.Immunol., 151, 2839-2851.

93 Portolano, S., Seto, P., Chazenbalk, G.D., Nagayama, Y., Mclachlan, S.M. & Rapoport, B. (1991) Biochem.Biophys.Res.Commun., 179, 372-377.

94 Pratt, L.F., Rassenti, L., Larrick, J., Robbins, B., Banks, P.M. & Kipps, T.J. (1989) J.Immunol., 143, 699-705.

95 Prelli, F., Tummolo, D., Solomon, A. & Frangione, B. (1986) J.Immunol., 136, 4169-4173.

96 Putnam, F.W., Whitley, E.J., Jr., Paul, C.& Davidson, J.N. (1973) Biochemistry, 12, 3763-3780.

97 Randen, I., Pascual, V., Victor, K., Thompson, K.M., Forre, O., Capra, J.D. & Natvig, J.B. (1993) Eur.J.Lnmunol., 23, 1220-1225.

98 Rassenti, L.Z., Pratt, L.F., Chen, P.P., Carson, D.A. & Kipps, T.J. (1991) J.Immunol., 147, 1060-1066.

99 Reidl, L.S., Friedman. D.F., Goldman, J., Hardy, RR., Jefferies, L.C. & Silberstein, L.E. (1991) J.Immunol., 147,3623-3631.

100 Riechmann, L., Clark, M., Waldmann, H. & Winter, G. (1988) Nature, 332, 323-327.

101 Riesen, W., Rudikoff, S., Oriol, R. & Potter, M. (1975) Biochemistry, 14, 1052-1057; Riesen, W.F., Braun, D.G. & Jaton, J.C. (1976) Proc.Nat.Acad.Sci.USA, 73, 2096-2100; Riesen, W.F. & Jaton, J.C. (1976) Biochemistry, 15, 3829.

102 Rodilla Sala, E., Kratzin, D.H., Pick, A.I. & Hilschmann, N. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark, Kluwer Academic

103 Schiechl, H. & Hilschmann, N. (1971) Z.Physiol.Chem., 352,111-115; (1972) Z.Physiol.Chem., 353, 345-370.

104 Schneider, M. & Hilschmann, N. (1974) Z.Physiol.Chem., 355, 1164-1168.

105 Shearman, C.W., Pollock, D., White, G., Hehir, K., Moore, G.P., Kanzy, E.J. & Kurrle, R. (1991) J.Immunol., 147, 4366-4373.

106 Shinoda, T. (1973) J.Biochem., 73, 433-446.

107 Shinoda, T. (1975) J.Biochein., 77, 1277-1296.

108 Shinoda, T., Takenawa, T., Hoshi, A. & Isobe, T. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig. O.Forre, G.Husby, A.Husebekk, B.Skogen. K.Sletten & P.WestermarlC, Kluwer Academic Publishers, Dordrecht/Boston/ London, Pp. 157.

109Silberstein, L.E., Litwin, S. & Carmack. C.E. (1989) J.Exp.Med., 169, 1631-1643.

110 Sims, M.J., Hassal, D.G., Brett, S., Rowan, W., Lockyer, M.J., Angel, A., Lewis, A.P., Hale, G., Waldmann, H. & Crowe, J.S. (1993) J.Immunol., 151, 2296-2308.

111 Spatz, L.A., Wong, K.K., Williams, M., Desai, R., Golier, J., Berman, J.E., Alt, F.W. & Latov, N. (1990) J.Immunol., 144, 2821-2828.

112 Stavnezer, J., Kekish, O., Batter, D., Grenier, J., Balazs, I., Henderson, E. & Zegers, B.J.M. (1985) Nucl.Acids Res., 13, 3495-3514.

113 Straubinger, B., Thiebe, R., Pech, M. & Zachau, H.G. (1988) Gene, 69, 209-214.

114 Suter, L., Barnikol, H.U., Watanabe, S. & Hilschmann, N. (1969) Z.Physiol.Chem., 350, 275-278; (1972) Z.Physiol.Chem., 353, 189-208.

115 Tempest, P.R., Bremner, P., Lambert, M., Taylor, G., Furze, J.M., Carr, F.J. & Harris, W.J. (1991) Bio/Tech., 9, 266-271.

116 Titani, K., Shinoda, T. & Putnam, F.W. (1969) J.Biol.Chem., 244, 3550-3560.

117 Toft, K.G., Olstad, O.K., Sletten, K. & Westermark, P. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P. Westermark, Kluwer Academic

118 Van Es, J.H., Aanstoot, H., Ginelig-Meyling, F.H.J., Derksen, R.H.W.M. & Logtenberg, T. (1992) J.Immunol., 149, 2234-2240.

119 Victor, K.D., Pascual, V., Lefvert, A.K. & Capra, J.D. (1992) Mol.Immunol., 29, 1501-1506.

120 Victor, K.D., Pascual, V., Williams, C.L., Lennon, V.A. & Capra, J.D. (1992) Eur.J.Immunol., 22, 2231-2236.

121 Victor, K.D., Randen, I., Thompson, K., Forre, O., Natvig, J.B., Fu, S.M. & Capra, J.D. (1991) J.Clin.Invest., 87, 1603-1613.

122 Wagner, S.D. & Luzzatto, L. (1993) Eur.J.Immunol., 23, 391-397.

123 Watanabe, S. & Hilschmann, N. (1970) Z.Physiol.Chem., 351, 1291-1295.

124 Weisbart, R.H., Wong, A.L., Noritake, D., Kacena, A., Chan, G., Ruland, C., Chin, E., Chen, I.S.Y. & Rosenblatt, J.D. (1991) J.Immunol., 147, 2795-2801.

125 Weng, N.-P., Yu-Lee, L.-Y., Sanz, I., Patten, B.M. & Marcus, D.M. (1992) J.Immunol., 149, 2518-2529.

126 Winkler, T.H., Fehr, H. & Kalden, J.R. (1992) Eur.J.Immunol., 22, 1719-1728.

References of rearranged human lambda sequences used for alignment

[0147]

1 Alexandre, D., Chuchana, P., Brockly, F., Blancher, A., Lefranc, G. & Lefranc, M.-P. (1989) Nuc.Acids Res., 17, 3975.

2 Anderson, M.L.M., Brown, L., Mckenzie, E., Kellow, J.E. & Young, B.D. (1985) Nuc.Acids Res., 13, 2931-2941.

3 Andris, J.S., Brodeur, B.R. & Capra, J.D. (1993) Mol.Immunol., 30, 1601-1616.

4 Andris, J.S., Ehrlich, P.H., Ostberg, L. & Capra, J.D. (1992) J.Immunol., 149, 4053-4059.

5 Baczko, K., Braun, D.G., Hess, M. & Hilschmann, N. (1970) Z.Physiol.Chem., 351, 763-767; Baczko, K., Braun, D.G. & Hilschmann, N. (1974) Z.Physiol.Chem., 355, 131-154.

6 Berinstein, N., Levy, S. & Levy, R. (1989) Science, 244, 337-339.

7 Bhat, N.M., Bieber, M.M., Chapman, C.J., Stevenson, F.K. & Teng, N.N.H. (1993) J.Immunol., 151, 5011-5021.

8 Cairns, E., Kwong, P.C., Misener, V., Ip, P., Bell, D.A. & Siminovitch, K.A. (1989) J.Immunol., 143, 685-691.

9 Carroll, W.L., Yu, M., Link, M.P. & Korsmeyer, S.J. (1989) J.Immunol., 143, 692-698.

10 Chen, B.L. & Poljak, R.J. (1974) Biochemistry, 13, 1295-1302.

11 Chen, B.L., Chiu, Y.Y.H., Humphrey, R.L. & Poljak, R.J. (1978) Biochim.Biophys.Acta, 537, 9-21.

12 Combriato, G. & Klobeck, H.G. (1991) Eur.J.Immunol., 21, 1513-1522.

13 Cuisinier, A.-M., Fumoux, F., Fougereau, M. & Tonnelle, C. (1992) Mol.Immunol., 29, 1363-1373.

14 Dwulet, F.E., Strako, K. & Benson, M.D. (1985) Scand.J.Immunol., 22, 653-660.

15 Elahna, P., Livneh, A., Manheimer-Lory, A.J. & Diamond, B. (1991) J.Immunol., 147, 2771-2776.

16 Engelhard, M., Hess, M. & Hilschmann, N. (1974) Z.Physiol.Chem., 355, 85-88; Engelhard, M. & Hilschmann, N. (1975) Z.Physiol.Chem., 356, 1413-1444.

17 Eulitz, M. (1974) Eur.J.Biochem., 50, 49-69.

18 Eulitz, M., Breuer, M. & Linke, R.P. (1987) Biol.Che.Hoppe-Seyler, 368, 863-870.

19 Eulitz, M., Murphy, C., Weiss, D.T. & Solomon, A. (1991) J.Immunol., 146, 3091-3096.

20 Fett, J.W. & Deutsch, H.F. (1974) Biochemistry, 13,4102-4114.

21 Fett, J.W. & Deutsch, H.F. (1976) Immunochem., 13, 149-155.; Jabusch, J.R. & Deutsch, H.F. (1982) Mol.Immunol., 19, 901-906.

22 Furey, W. Jr., Wang, B.C., Yoo, C.S. & Sax, M. (1983) J.Mol.Biol., 167, 661-692.

23 Fykse, E.-M., Sletten, K., Husby, G. & Cornwell, G.G., III (1988) Biochem.J., 256, 973-980.

24 Garver, F.A. & Hilschmann, N (1971) Febs Letters, 16. 128-132; (1972) Eur.J.Biochem., 26, 10-32.

25 Gawinowicz, M.A., Merlini, G., Birken, S., Osserman, E.F. & Kabat, E.A. (1991) J.Immunol., 147,915-920.

26 Ghiso, J., Solomon, A. & Frangione, B. (1986) J.Immunol., 136, 716-719.

27 Griffiths, A.D., Malmqvist, M., Marks, J.D., Bye, J.M., Embleton, M.J., Mccafferty, J., Baier, M., Holliger, K.P., Gorick, B.D., Hughes-Jones, N.C., Hoogenboom, H.R. & Winter, G. (1993) Embo J., 12, 725-734.

28 Gullasken, N., Idso, H., Nilsen, R., Sletten, K., Husby, G. & Cornwell, G.G. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark, Kluwer Academic

29 Harindranath, N., Goldfarb, I.S., Ikematsu, H., Burastero, S.E., Wilder, R.L., Notkins, A.L. & Casali, P. (1991) Int.Immunol., 3, 865-875.

30 Holm, E., Sletten, K. & Husby, G. (1986) Biochem.J., 239, 545-551.

31 Hughes-Jones, N.C., Bye, J.M., Beale, D. & Coadwell, J. (1990) Biochem.J., 268, 135-140.

32 Kametani, F., Yoshimura, K., Tonoike, H., Hoshi, A., Shinoda, T. & Isobe, T. (1985) Biochem.Biophys.Res.Commun., 126, 848-852.

33 Kiefer, C.R., Mcguire, B.S., Jr., Osserman, E.F. & Garver, F.A. (1983) J.Immunol., 131, 1871-1875.

34 Kiefer, C.R., Patton, H.M., Jr., Mcquire, B.S., Jr. & Garver, F.A. (1984) J.Immunol., 124.301-306.

35 Kishimoto, T., Okajima, H., Okumoto, T. & Taniguchi, M. (1989) Nud.Acids Res., 17, 4385.

36 Klafki, H.-W., Kratzin, H.D., Pick, A.I., Eckart, K. & Hilschmann, N. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark, Kluwer Academic

37 Kohler, H., Rudofsky, S. & Kluskens, L. (1975) J.Irmnunology, 114, 415-421.

38 Kojima, M., Odani, S. & Ikenaka, T. (1980) Mol.Immunol., 17, 1407-1414.

39 Komori, S., Yamasaki, N., Shigeta, M., Isojima, S. & Watanabe, T. (1988) Clin.Exp.immunol., 71, 508-516.

40 Kratzin, H.D., Palm, W., Stangel, M., Schmidt, W.E., Friedrich, J. & Hilschmann, N. (1989) Biol.Cheni.Hoppe-

Seyler, 370, 263-272.

41 Kratzin, H.D., Pick, A.I., Stangel, M. & Hilschmann, N. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark. Kluwer Academic Publishers, Dordrecht/ Boston/London, Pp.181.

42 Langer, B., Steinmetz-Kayne, M. & Hilschmann, N. (1968) Z.Physiol.Chem., 349. 945-951.

43 Larrick, J.W., Danielsson, L., Brenner, C.A., Wallace, E.F., Abrahamson, M., Fry, K.E. & Borrebaeck, C.A.K. (1989) Bio/Tech., 7, 934-938.

44 Levy, S., Mendel, E., Kon, S., Avnur, Z. & Levy, R. (1988) J.Exp.Med., 168, 475-489.

45 Lewis, A.P., Lemon, S.M., Barber, K.A., Murphy, P., Parry, N.R., Peakman, T.C., Sims, M.J., Worden, J. & Crowe, J.S. (1993) J.Immunol., 151, 2829-2838.

46 Liu. V.Y.S., Low, T.L.K., Infante, A. & Putnam, F.W. (1976) Science, 193, 1017-1020; Infante. A. & Putnam, F.W. (1979) J.Biol.Chem., 254, 9006-9016.

47 Lopez De Castro, J.A., Chiu, Y.Y.H. & Poljak, R.J. (1978) Biochemistry, 17, 1718-1723.

48 Mantovani, L., Wilder, R.L. & Casali, P. (1993) J.Immunol., 151, 473-488.

49 Marks, J.D., Hoogenboom, H.R., Bonnert, T.P., Mccafferty, J., Griffiths, A.D. & Winter, G. (1991) J.Mol.Biol., 222, 581-597.

50 Mihaesco, E., Roy, J.-P., Congy, N., Peran-Rivat, L. & Mihaesco, C. (1985) Eur.J.Biochem., 150, 349-357.

51 Milstein, C., Clegg, J.B. & Jarvis, J.M. (1968) Biocliem.J., 110, 631-652.

52 Moran. M.J., Andris, J.S., Matsumato, Y.-I., Capra, J.D. & Hersh, E.M. (1993) Mol.Immunol., 30, 1543-1551.

53 Nabeshima, Y. & Ikenaka, T. (1979) Mol.Immunol., 16, 439-444.

54 Olee, B.T., Lu, E.W., Huang, D.-F., Soto-Gil, R.W., Deftos, M., Kozin, F., Carson, D.A. & Chen, P.P. (1992) J.Exp.Med., 175, 831-842.

55 Pascual, V., Victor, K., Randen, I., Thompson, K., Steinitz, M., Forre, O., Fu, S.-M., Natvig, J.B. & Capra, J.D. (1992) Scand.J.Immunol., 36, 349-362.

56 Paul, E., Iliev, A.A., Livneh, A. & Diamond, B. (1992) J.Immunol., 149, 3588-3595.

57 Pick, A.I., Kratzin, H.D., Bamikol-Watanabe, S. & Hilschmann, N. (1990) In Amyloid And Amyloidosis, Eds. J.B.Natvig, O.Forre, G.Husby, A.Husebekk, B.Skogen, K.Sletten & P.Westermark, Kluwer Academic

58 Ponstingl, H. & Hilschmann, N. (1969) Z.Physioi.Chem., 350, 1148-1152; (1971) Z.Physiol.Chem., 352, 859-877.

59 Ponstingl, H., Hess, M. & Hilschmann, N. (1968) Z.Physiol.Chem., 349, 867-871; (1971) Z.Physiol.Chem., 352, 247-266.

60 Randen, L, Pascual, V., Victor, K., Thompson, K.M., Forre, O., Capra, J.D. & Natvig. J.B. (1993) Eur.J.Immunol., 23, 1220-1225.

61 Scholz, R. & Hilschmann, N. (1975) Z.Physiol.Chem., 356, 1333-1335.

62 Settmacher, U., Jahn, S., Siegel, P., Von Baehr, R. & Hansen, A. (1993) Mol.Immunol., 30, 953-954.

63 Shinoda, T., Titani, K. & Putnam, F.W. (1970) J.Biol.Chem., 245, 4475-4487.

64 Sletten, K., Husby, G. & Natvig, J.B. (1974) Scand.J.Immunol., 3, 833-836.; Sletten, K., Natvig, J.B., Husby, G. & Juul, J. (1981) Biochem.J., 195, 561-572.

65 Solomon, A., Frangione, B. & Franklin, E.C. (1982) J.Clin.invest., 70, 453-460.; Frangione, B., Moloshok, T. & Solomon, A. (1983) J.Immunol., 131, 2490-2493.

66 Takahashi, N., Takayasu, T., Isobe, T., Shinoda, T., Okuyama, T. & Shimizu, A. (1979) J.Biochem., 86, 1523-1535.

67 Takahashi, N., Takayasu, T., Shinoda, T., Ito, S., Okuyama, T. & Shimizu, A. (1980) Biomed.Res., 1, 321-333.

68 Takahashi, Y., Takahashi, N., Tetaert, D. & Putnam, F.W. (1983) Proc.Nat.Acad.Sci.USA, 80, 3686-3690.

69 Takayasu, T., Takahashi, N., Shinoda, T., Okuyama, T. & Tomioka, H. (1980) J.Biochem., 89.421-436.

70 Titani. K., Wikler, M., Shinoda, T. & Putnam, F.W. (1970) J. Biol.Chem., 245, 2171-2176.

71 Toft, K.G., Sletten, K. & Husby, G. (1985) Biol.Cliem.Hoppe-Seyler. 366, 617-625.

72 Tonoike, H., Kametani, F., Hoshi, A., Shinoda, T. & Isobe, T. (1985) Biochem.Biophys.Res.Commun., 126, 1228-1234.

73 Tonoike, H., Kametani, F., Hoshi, A., Shinoda, T. & Isobe, T. (1985) Febs Letters, 185, 139-141.

74 Tsujimoto, Y. & Croce, C.M. (1984) Nucl.Acids Res., 12. 8407-8414.

75 Tsunetsugu-Yokota, Y. Minekawa, T., Shigemoto, K., Shirasawa, T. & Takemori, T. (1992) Mol.Immunol., 29, 723-728.

76 Tveteraas, T., Sletten, K. & Westermark, P. (1985) Biochem.J., 232, 183-190.

77 Vasicek, T.J. & Leder, P. (1990) J.Exp.Med., 172, 609-620.

78 Victor, K.D., Randen, I., Thompson, K., Forre, O., Natvig, J.B., Fu, S.M. & Capra, J.D. (1991) J.Clin.Invest., 87, 1603-1613.

79 Weng, N.-P., Yu-Lee, L.-Y., Sanz, I., Patten, B.M. & Marcus, D.M. (1992) J.Immunol., 149, 2518-2529.

80 Wikler, M. & Putnam, F.W. (1970) J.Biol.Chem., 245, 4488-4507.

81 Winkler, T.H., Fehr, H. & Kalden, J.R. (1992) Eur.J.Immunol., 22, 1719-1728.

82 Yago, K., Zenita, K., Ohwaki, I., Harada, Y., Nozawa, S., Tsukazaki, K., Iwamori, M., Endo, N., Yasuda, N., Okuma, M. & Kannagi, R. (1993) MoLImmunol., 30, 1481-1489.

83 Yamasaki, N., Komori, S. & Watanabe, T. (1987) Mol.Immunol., 24, 981-985.

84 Zhu, D., Kim, H.S. & Deutsch, H.F. (1983) Mol.Immunol., 20,1107-1116.

85 Zhu. D., Zhang, H., Zhu, N. & Luo, X. (1986) Scientia Sinica, 29, 746-755.

References of rearranged human heavy chain sequences used for alignment

[0148]

1 Adderson, E.E., Azmi, F.H., Wilson, P.M., Shackelford, P.G. & Carroll, W.L. (1993) J.Immunol., 151, 800-809.

2 Adderson, E.E., Shackelford, P.G., Quinn, A. & Carroll, W.L. (1991) J.Immunol., 147, 1667-1674.

3 Akahori, Y., Kurosawa, Y., Kamachi, Y., Torii, S. & Matsuoka, H. (1990) J.Clin.Invest., 85, 1722-1727.

4 Andris, J.S., Brodeur, B.R. & Capra, J.D. (1993) Mol.Immunol., 30, 1601-1616.

5 Andris, J.S., Ehrlich, P.H., Ostberg. L. & Capra, J.D. (1992) J.Immunol., 149,4053-4059.

6 Andris, J.S., Johnson, S., Zolla-Pazner. S. & Capra, J.D. (1991) Proc.Natl.Acad.Sci.USA, 88, 7783-7787.

7 Anker, R., Conley, M.E. & Pollok, B.A. (1989) J.Exp.Med., 169, 2109-2119.

8 Atkinson, P.M., Lampman, G.W., Furie, B.C., Naparstek, Y., Schwartz, R.S., Stollar, B.D. & Furie, B. (1985) J.Clin.Invest., 75, 1138-1143.;Lampman, G.W., Furie, B., Schwartz, R.S., Stollar, B.D. & Furie, B.C. (1989)

9 Avila, M.A., Vazques, J., Danielsson, L., Fernandez De Cossio, M.E. & Borrebaeck, C.A.K. (1993) Gene, 127, 273-274.

10 Bakkus, M.H.C., Heirman, C., Van Riet, I., Van Camp, B. & Thielemans, K. (1992) Blood, 80, 2326-2335.

11 Barbas, C.F., III, Crowe, Jr., J.E., Cababa, D., Jones, T.M., Zebedee, S.L., Murphy, B.R., Chanock, R.M. & Burton, D.R. (1992) Proc.Natl.Acad.Sci.USA, 89, 10164-10168.

12 Barbas, C.F., III, Collet, T.A., Amberg, W., Roben, P., Binley, J.M., Hoekstra, D., Cababa, D., Jones, T.M., Williamson, R.A., Pilkington, G.R., Haigwood, N.L., Cabezas, E., Satterthwait, A.C., Sanz, I. & Burton, D.R. (1993) J.Mol.Biol., 230, 812-823.

13 Berman, J.E., Humphries, C.G., Barth, J., Alt, F.W. & Tucker, P.W. (1991) J.Exp.Med., 173, 1529-1535.

14 Berman, J.E., Mellis, S.J., Pollock. R., Smith, C.L., Suh, H., Heinke, B., Kowal, C., Surti, U., Chess, L., Cantor, C.R & Alt, F.W. (1988) Embo J., 7, 727-738.

15 Bhat, N.M., Bieber, M.M., Chapman, C.J., Stevenson, F.K. & Teng. N.N.H. (1993) J.Immunol., 151, 5011-5021.

16 Bird, J., Galili, N., Link, M., Stites, D. & Sklar, J. (1988) J.Exp.Med., 168, 229-245.

17 Cai, J., Humphries, C., Richardson, A. & Tucker, P.W. (1992) J.Exp.Med., 176, 1073-1081.

18 Cairns, E., Kwong, P.C., Misener, V., Ip, P., Bell, D.A. & Siminovitch, K.A. (1989) J.Immunol., 143, 685-691.

19 Capra, J.D. & Hopper, J.E. (1976) Immunochemistry, 13, 995-999; Hopper, J.E., Noyes, C., Heinrikson, R. & Kessel, J.W. (1976) J.Immunol., 116, 743-746.

20 Capra, J.D. & Kehoe, J.M. (1974) Proc.Nat.Acad.Sci.USA, 71, 845-848.

21 Carroll, W.L., Yu, M., Link, M.P. & Korsmeyer, S.J. (1989) J.Immunol., 143, 692-698.

22 Chen, P.P., Liu, M.-F., Glass, C.A., Sinha, S., Kipps, T.J. & Carson, D.A. (1989) Arthritis & Rheumatism, 32, 72-76; Kipps, T.J., Tomhave, E., Pratt, L.F., Duffy, S., Chen, P.P. & Carson. D.A. (1989) Proc.Natl.Acad.Sci.USA, 86, 5913-5917.

23 Chiu, Y.Y.H., Lopez De Castro, J.A. & Poljak, R.J. (1979) Biochemistry, 18, 553-560.

24 Cleary, M.L., Meeker, T.C., Levy, S., Lee, E., Trela, M., Sklar, J. & Levy, R. (1986) Cell, 44, 97-106.

25 Cuisinier, A.-M., Fumoux, F., Fougereau, M. & Tonnelle, C. (1992) Mol.Immunol., 29, 1363-1373.

26 Cuisinier, A.-M., Gauthier, L., Boubli, L., Fougereau, M. & Tonnelle, C. (1993) Eur.J.Immunol., 23. 110-118.

27 Cunningham. B.A., Gottlieb.P.D., Pflumm, M.N. & Edelman, G.M. (1971) Progress In Immunology (B.Amos, Ed.),

Academic Press, N.Y., Pp.3-24.

28 Cunningham, B.A., Rutishauser, U., Gall, W.E., Gottlieb, P.D., Waxdal, M.J. & Edelman, G.M. (1970) Biochemistry, 9, 3161-3170.

29 Deane, M. & Norton, J.D. (1990) Eur.J.Immunol., 20. 2209-2217.

31 Dersimonian, H., Schwartz, R.S., Barrett, K.J. & Stollar, B.D. (1987) J.Immunol., 139, 2496-2501.

32 Dersimonian, H., Schwartz, R.S., Barrett, K.J. & Stollar, B.D. (1987) J.Immunol., 139, 2496-2501; Chen, P.P., Liu, M.-F., Sinha, S. & Carson, D.A. (1988) Arth.Rheum., 31, 1429-1431.

33 Desai, R., Spatz, L., Matsuda, T., Ilyas, A.A., Berman, J.E., Alt, F.W., Kabat, E.A. & Latov, N. (1990) J.Neuroimmunol., 26, 35-41.

34 Ezaki, I., Kanda, H., Sakai, K., Fukui, N., Shingu, M., Nobunaga, M. & Watanabe, T. (1991) Arthritis And Rheumatism. 34. 343-350.

35 Felgenhauer, M., Kohl, J. & Ruker, F. (1990) Nucl.Acids Res., 18, 4927.

36 Florent, G., Lehman, D. & Putnam, F.W. (1974) Biochemistry, 13, 2482-2498.

37 Friedlander, R.M., Nussenzweig, M.C. & Leder, P. (1990) Nucl.Acids Res., 18, 4278.

38 Gawinowicz, M.A., Merlini, G., Birken, S., Osserman, E.F. & Kabat, E.A. (1991) J.Immunol., 147,915-920.

39 Gillies, S.D., Dorai, H., Wesolowski, J., Majeau, G., Young, D., Boyd, J., Gardner, J. & James, K. (1989) Bio/Tech., 7, 799-804.

40 Goni, F. & Frangione, B. (1983) Proc.Nat.Acad.Sci.USA, 80, 4837-4841.

41 Gorman, S.D., Clark, M.R., Routledge, E.G., Cobbold, S.P. & Waldmann, H. (1991) Proc.Natl.Acad.Sci.USA, 88,4181-4185.

42 Griffiths, A.D., Malmqvist, M., Marks, J.D., Bye, J.M., Embleton, M.J., Mccafferty, J., Baier, M., Holliger, K.P., Gorick, B.D., Hughes-Jones, N.C., Hoogenboom, H.R. & Winter, G. (1993) Embo J., 12, 725-734.

43 Grillot-Courvalin, C., Brouet, J.-C., Piller, F., Rassenti, L.Z., Labaume, S., Silverman, G.J., Silberstein, L. & Kipps, T.J. (1992) Eur.J.Immunol., 22, 1781-1788.

44 Guillaume, T., Rubinstein, D.B., Young, F., Tucker, L., Logtenberg, T., Schwartz, R.S. & Barrett, K.L. (1990) J.Immunol., 145, 1934-1945;Young, F., Tucker, L., Rubinstein, D., Guillaume, T., Andre-Schwartz, J., Barrett, K.J., Schwartz, R.S. & Logtenberg, T. (1990)

45 Harindranath, N., Goldfarb, I.S., Ikematsu, H., Burastero, S.E., Wilder, R.L., Notkins, A.L. & Casali, P. (1991) Int.Immunol., 3, 865-875.

46 Hillson, J.L., Oppliger, I.R., Sasso, E.H., Milner, E.C.B. & Wener, M.H. (1992) J.Immunol., 149, 3741-3752.

47 Hirabayashi, Y., Munakata, Y., Sasaki, T. & Sano, H. (1992) Nucl.Acids Res., 20, 2601.

48 Hoch, S. & Schwaber, J. (1987) J.Immunol., 139, 1689-1693.

49 Huang, C., Stewart, A.K., Schwartz, R.S. & Stollar, B.D. (1992) J.Clin.Invest., 89, 1331-1343.

50 Hughes-Jones, N.C., Bye, J.M., Beale, D. & Coadwell, J. (1990) Biochem.J., 268, 135-140.

51 Ikematsu, H., Harindranath, N., Ueki, Y., Notkins, A.L. & Casali, P. (1993) J.Immunol., 150, 1325-1337.

52 Ikematsu, H., Kasaian, M.T., Schettino, E.W. & Casali, P. (1993) J.Immunol., 151, 3604-3616.

53 Kelly, P.J., Pascual, V., Capra, J.D. & Lipsky, P.E. (1992) J.Immunol., 148, 1294-1301.

54 Kipps, T.J. & Duffy, S.F. (1991) J.Clin.Invest., 87. 2087-2096.

55 Kipps, T.J., Tomhave, E., Pratt, L.F., DufFy, S., Chen, P.P. & Carson, D.A. (1989) Proc.Natl.Acad.Sci.USA, 86, 5913-5917.

56 Kishimoto, T., Okajima, H., Okumoto, T. & Taniguchi, M. (1989) Nucl.Acids Res., 17, 4385.

57 Knight, G.B., Agnello, V., Bonagura, V., Barnes, J.L., Panka, D.J. & Zhang, Q.-X. (1993) J.Exp.Med., 178, 1903-1911.

58 Kohler, H., Shimizu, A., Paul, C., Moore, V. & Putnam, F.W. (1970) Nature, 227, 1318-1320; Florent, G., Lehman, D. & Putnam, F.W. (1974) Biochemistry, 13, 2482-2498

59 Komori, S., Yamasaki, N., Shigeta, M., Isojima, S. & Watanabe, T. (1988) Clin.Exp.Immunol., 71, 508-516.

60 Kon. S., Levy, S. & Levy, R. (1987) Proc.Natl.Acad.Sci.USA, 84, 5053-5057.

61 Kratzin, H., Altevogt, P., Ruban, E., Kortt, A., Staroscik, K. & Hilschmann, N. (1975) Z.Physiol.Chem., 356, 1337-1342; Kratzin, H., Altevogt, P., Kortt, A., Ruban, E. & Hilschmann, N. (1978) Z.Physiol.Chem., 359, 1717-1745.

62 Kudo, A., Ishihara, T., Nishimura, Y. & Watanabe, T. (1985) Gene, 33, 181-189.

63 Kunicki, T.J., Annis, D.S., Gorski, J. & Nugent, D.J. (1991) J.Autoimmunity, 4, 433-446.

64 Larrick, J.W., Wallace, E.F., Coloma, M.J., Bruderer, U., Lang, A.B. & Fry, K.E. (1992) Immunological Reviews, 130, 69-85.

65 Lehman, D.W. & Putnam, F.W. (1980) Proc.Nat.Acad.Sci.USA, 77, 3239-3243.

66 Lewis, A.P., Lemon, S.M., Barber, K.A., Murphy, P., Parry, N.R., Peakman, T.C., Sims, M.J., Worden, J. & Crowe, J.S. (1993) J.Immunol., 151, 2829-2838.

67 Liu, V.Y.S., Low, T.L.K., Infante, A. & Putnam, F.W. (1976) Science, 193, 1017-1020.

68 Logtenberg, T., Young, F.M., Van Es, J., Gmelig-Meyling, F.H.J., Berman, J.E. & Alt, F.W. (1989) J. Autoimmunity, 2, 203-213.

69 Logtenberg, T., Young, F.M., Van Es, J.H., Gnielig-Meyling, F.H.J. & Alt, F.W. (1989) J.Exp.Med., 170, 1347-1355.

70 Manfteimer-Lory, A., Katz. J.B., Pillinger, M., Ghossein, C., Smith, A. & Diamond, B. (1991) J.Exp.Med., 174, 1639-1652.

71 Mantovani, L., Wilder, R.L. & Casali, P. (1993) J.Immunol., 151, 473-488.

72 Mariette, X., Tsapis, A. & Brouet, J.-C. (1993) Eur.J.Immunol., 23, 846-851.

73 Marks, J.D., Hoogenboom, H.R. Bonnert, T.P., Mccafferty, J., Griffiths, A.D. & Winter, G. (1991) J.Mol.Biol., 222, 581-597.

74 Meeker, T.C., Grimaldi, J., O'rourke, R., Loeb, J.Juliusson, G. & Einhorn, S. (1988) J.Immol., 141, 3994-3998.

75 Milili, M., Fougereau, M., Guglielmi, P. & Schiff, C. (1991) Mol.Immunol., 28, 753-761.

76 Moran, M.J., Andris, J.S., Matsumato, Y.-I., Capra, J.D. & Hersh, E.M. (1993) Mol.Immunol., 30, 1543-1551.

77 Mortari, F., Wang, J.-Y. & Schroeder, Jr., H.W. (1993) J.Inununol., 150, 1348-1357.

78 Newkirk, M.M., Gram, H., Heinrich, G.F., Ostberg, L., Capra, J.D. & Wasserman, R.L. (1988) J.Clin.Invest., 81, 1511-1518.

79 Newkirk, M.M., Mageed, R.A., Jefferis, R., Chen, P.P. & Capra, J.D. (1987) J.Exp.Med., 166, 550-564.

80 Nickerson, K.G., Berman, J., Glickman, E., Chess, L. & Alt, F.W. (1989) J.Exp.Med., 169, 1391-1403.

81 Olee, B.T., Lu, E.W., Huang, D.-F., Soto-Gil, R.W., Deftos, M., Kozin, F., Carson, D.A. & Chen, P.P. (1992) J.Exp.Med., 175, 831-842.

82 Pascual, V., Randen, I., Thompson, K., Sioud, M.Forre, O., Natvig, J. & Capra, J.D. (1990) J.Clin.Invest., 86, 1320-1328.

83 Pascual, V., Randen, I., Thompson, K., Sioud, M.Forre, O., Natvig, J. & Capra, J.D. (1990) J.Clin.Invest., 86, 1320-1328; Randen, I., Brown, D., Thompson, K.M., Hughes-Jones. N., Pascual, V., Victor, K., Capra, J.D., Forre, O. & Natvig, J.B. (1992)

84 Pascual, V., Victor, K., Lelsz, D., Spellerberg, M.B., Hamblin, T.J., Thompson, K.M., Randen, I., Natvig, J., Capra, J.D. & Stevenson, F.K. (1991) J.Immunol., 146, 4385-4391.

85 Pascual, V., Victor, K., Randen, L, Thompson, K., Steinitz, M., Forre, O., Fu, S.-M., Natvig, J.B. & Capra, J.D. (1992) Scand.J.Immunol., 36, 349-362.

86 Pascual, V., Victor, K., Spellerberg, M., Hamblin, T.J., Stevenson, F.K. & Capra, J.D. (1992) J.Immunol., 149, 2337-2344.

87 Ponstingl, H., Schwarz, J., Reichel, W. & Hilschmann, N. (1970) Z. Physiol. Chem., 351, 1591-1594.; Ponstingl, H. & Hilschmann, N. (1976) Z.Physiol.Chem., 357, 1571-1604.

88 Portolano, S., Mclachlan, S.M. & Rapoport, B. (1993) J.Iminunol., 151, 2839-2851.

89 Portolano, S., Seto, P., Chazenbalk, G.D., Nagayama, Y., Mclachlan, S.M. & Rapoport, B. (1991) Biochem.Biophys.Res.Commun., 179,372-377.

90 Pratt, L.F., Szubin, R., Carson, D.A. & Kipps, T.J. (1991) J.Iminunol., 147, 2041-2046.

91 Press, E.M. & Hogg, N.M. (1970) Biochem.J., 117, 641-660.

92 Putnam, F.W., Shimizu, A., Paul., C., Shinoda, T. & Kohler, H. (1971) Ann.N.Y.Acad.Sci., 190, 83-103.

93 Putnam, F.W., Takahashi, N., Tetaert, D., Debuire, B. & Lin, L.C. (1981) Proc.Nat.Acad.Sci.USA, 78, 6168-6172.; Takahashi, N., Tetaert, D., Debuire, B., Lin, L. & Putnam, F.W. (1982) Proc.Nat.Acad.Sci.USA, 79, 2850-2854.

94 Raaphorst, F.M., Timmers, E., Kenter, M.J.H., Van Tol, M.J.D., Vossen, J.M. & Schuurman, R.K.B. (1992) Eur.J.Immunol., 22, 247-251.

95 Rabbitts. T.H., Bentley, D.L., Dunnick, W., Forster, A., Matthyssens, G. & Milstein, C. (1980) Cold Spring Harb.Symp.Quanti.Biol., 45, 867-878; Matthyssens, G. & Rabbitts, T.H. (1980) Proc.Nat.Acad.Sci.USA, 77, 6561-6565.

96 Randen, I., Pascual, V., Victor, K., Thompson, K.M., Forre, O., Capra, J.D. & Natvig, J.B. (1993) Eur.J.Immunol., 23, 1220-1225.

97 Rassenti, L.Z. & Kipps, T.J. (1993) J.Exp.Med., 177, 1039-1046.

98 Reidl, L.S., Friedman, D.F., Goldman, J., Hardy, R.R., Jefferies, L.C. & Silberstein, L.E. (1991) J.Immunol., 147,

3623-3631.

99 Roudier, J., Silverman, G.J., Chen, P.P., Carson, D.A. & Kipps, T.J. (1990) J.Immunol., 144, 1526-1530.

100 Sanz, I., Casali, P., Thomas, J.W., Notkins, A.L. & Capra, J.D. (1989) J.Immunol., 142, 4054-4061.

101 Sanz, I., Dang, H., Takei, M., Talal, N. & Capra, J.D. (1989) J.Immunol., 142, 883-887.

102 Schmidt, W.E., Jung, H-.D., Palm, W. & Hilschmann, N. (1983) Z.Physiol.Chem., 364, 713-747.

103 Schroeder. H.W., Jr. & Wang, J.Y. (1990) Proc.Natl.Acad.Sci.USA, 87, 6146-6150.

104 Schroeder, H.W., Jr., Hillson, J.L. & Perlmutter, R.M. (1987) Science, 238, 791-793.

105 Schroeder, H.W., Jr., Hillson, J.L. & Perlmutter, R.M. (1987) Science, 238, 791-793; Chen. P.P., Liu, M.-F., Glass. C.A., Sinha, S., Kipps, T.J. & Carson, D.A. (1989) Arthritis & Rheumatism, 32, 72-76.

106 Schroeder, H.W., Jr., Hillson, J.L. & Perlmutter, R.M. (1987) Science, 238, 791-793; Chen, P.P., Liu, M.-F., Sinha, S. & Carson, D.A. (1988) Arth.Rheum., 31, 1429-1431.

107 Schutte, M.E., Ebeling, S.B., Akkermans, K.E., Gmelig-Meyling, F.H. & Logtenberg, T. (1991) Eur.J.Immunol., 21, 1115-1121.

108 Schutte, M.E., Ebeling, S.B., Akkermans, K.E., Gmelig-Meyling, F.H.J. & Logtenberg, T. (1991) Eur.J.Immunol., 21, 1115-1121.

109 Settmacher, U., Jalin, S., Siegel, P., Von Baehr, R. & Hansen, A. (1993) Mol.Immunol., 30, 953-954.

110 Shen, A., Humphries, C., Tucker, P. & Blattner, F. (1987) Proc.Natl.Acad.Sci.USA, 84, 8563-8567.

111 Shimizu, A., Nussenzweig, M.C., Mizuta, T.-R., Leder, P. & Honjo, T. (1989) Proc. Natl.Acad. Sci. USA, 86, 8020-8023.

112 Shin, E.K., Matsuda, F., Fujikura, J., Akamizu, T., Sugawa, H., Mori, T. & Honjo, T. (1993) Eur.J.Immunol., 23, 2365-2367.

113 Silberstein, L.E., Litwin, S. & Carmack, C.E. (1989) J.Exp.Med., 169, 1631-1643.

114 Singal, D.P., Frame. B., Joseph. S., Blajchman, M.A. & Leber, B.F. (1993) Immunogenet., 38, 242.

115 Spatz, L.A., Wong, K.K., Williams, M., Desai, R., Golier, J., Berman, J.E., Alt, F.W. & Latov, N. (1990) J.Immunol., 144, 2821-2828.

116 Steiner, L.A., Garcia-Pardo, A. & Margolies, M.N. (1979) Biochemistry, 18, 4068-4080.

117 Stewart, A.K., Huang, C., Stollar, B.D. & Schwartz, R.S. (1993) J.Exp.Med., 177,409-418.

118 Thomas, J.W. (1993) J.Immunol., 150, 1375-1382.

119 Torano, A. & Putnam, F.W. (1978) Proc.Nat.Acad.Sci.USA, 75, 966-969.

120 Van Der Heijden, R.W.J., Bunschoten, H., Pascual, V., Uytdehaag, F.G.C.M., Osterhaus, A.D.M.E. & Capra, J.D. (1990) J.Immunol., 144, 2835-2839.

121 Van Der Stoep, N., Van Der Linden, J. & Logtenberg, T. (1993) J.Exp.Med., 177, 99-107.

122 Van Es, J.H., Gmelig-Meyling, F.H.J. & Logtenberg, T. (1992) Eur.J.Immunol., 22, 2761-2764.

123 Varade, W.S., Marin, E., Kittelberger, A.M. & Insel, R.A. (1993) J.Immunol., 150, 4985-4995.

124 Victor, K.D., Pascual, V., Lefvert, A.K. & Capra, J.D. (1992) Mol.Immunol., 29, 1501-1506.

125 Victor, K.D., Pascual, V., Williams, C.L., Lennon, V.A. & Capra, J.D. (1992) Eur.J.Immunol., 22, 2231-2236.

126 Watanabe, S., Barnikol, H.U., Horn, J., Bertram, J. & Hilschmann, N. (1973) Z.Physiol.Chem., 354, 1505-1509.

127 Weng, N.-P., Yu-Lee, L.-Y., Sanz, L, Patten, B.M. & Marcus, D.M. (1992) J.Intmunol., 149, 2518-2529.

128 White, M.B., Word, C.J., Humphries, C.G., Blattner, F.R. & Tucker, P.W. (1990) Mol.Cell.Biol., 10, 3690-3699.

129 Winkler, T.H., Fehr, H. & Kalden, J.R. (1992) Eur.J.Immunol., 22,1719-1728.

130 Yago, K., Zenita, K., Ohwaki, I., Harada, Y., Nozawa, S., Tsukazaki, K., Iwamori, M., Endo, N., Yasuda, N., Okuma, M. & Kannagi, R. (1993) Mol.Immunol., 30, 1481-1489.

131 Zelenetz, A.D., Chen, T.T. & Levy, R. (1992) J.Exp.Med., 176, 1137-1148.

B. References of germline sequences

References of human germline kappa sequences

[0149]

1 Cox, J.P.L., Tomlinson, I.M. & Winter, G. (1994) Eur.J.Immunol., 24, 827-836.

2 Huber, C., Et Al. (1993) Eur.J.Immunol., 23, 2868.

3 Klobeck, H.G., Bornkammm, G.W., Combriato, G., Mocikat, R., Pohlenz, H.D. & Zachau, H.G. (1985) Nucl.Acids Res., 13, 6515-6529.

4 Lautner-Rieske. A., Huber, C., Meindl, A., Pargent, W., Schäble, K.F., Thiebe, R., Zocher, I. & Zachau, H.G. (1992) Eur.J.Immunol. 22, 1023.

5 Lorenz, W., SchAble, K.F., Thiebe, R., Stavnezer, J. & Zachau, H.G. (1988) Mol.Immunol., 25, 479.

6 Pargent, W., Meindl, A., Thiebe, R., Mitzel, S. & Zachau, H.G. (1991) Eur.J.Immunol., 21, 1821-1827.

7 Pech, M. & Zachau, H.G. (1984) Nuc.Acids Res., 12, 9229-9236.

8 Pech, M., Jaenichen, H.-R., Pohlenz, H.-D., Neumaier, P.S., Klobeck, H.-G. & Zachau, H.G. (1984) J.Mol.Biol., 176, 189-204.

9 Scott, M.G., Crimmins, D.L., Mccourt, D.W., Chung, G., Schable, K.F., Thiebe, R., Quenzel, E.-M., Zachau, H.G. & Nahm, M.H. (1991) J.Immunol., 147, 4007-4013.

10 Stavnezer, J., Kekish, O., Batter, D., Grenier, J., Balazs, I., Henderson, E. & Zegers, B.J.M. (1985) Nucl.Acids Res., 13, 3495-3514.

11 Straubinger, B., Huber, E., Lorenz, W., Osterholzer, E., Pargent, W., Pech, M., Pohfenz, H.-D., Zimmer, F.-J. & Zachau, H.G. (1988) J.Mol.Biol., 199, 23-34.

12 Straubinger, B., Thiebe, R., Huber, C., Osterholzer, E. & Zachau, H.G. (1988) Biol. Chem. Hoppe-Seyer, 369, 601-607.

References of human germline lambda sequences

[0150]

1 Williams, S.C. & Winter, G. (1993) Eur.J.Immunol., 23, 1456-1461.

2 Siminovitch, K.A., Misener, V., Kwong, P.C., Song, Q.-L. & Chen, P.P. (1989) J.Clin.Invest., 84, 1675-1678.

3 Brockly, F., Alexandre, D., Chuchana, P., Huck, S., Lefranc, G. & Lefranc, M.-P. (1989) Nuc.Acids.Res., 17, 3976.

4 Daley, M.D., Peng, H.-Q., Misener, V., Liu, X.-Y., Chen, P.P. & Siminovitch, K.A. (1992) Mol.Immunol., 29, 1515-1518.

5 Deftos, M., Soto-Gil, R., Quan, M., Olee, T. & Chen, P.P. (1994) Scand. J. Immunol., 39,95.

6 Stiernholm, N.B.J., Kuzniar, B. & Berinstein, N.L. (1994) J. Immunol., 152, 4969-4975.

7 Combriato, G. & Klobeck, H.G. (1991) Eur.J.Immunol., 21, 1513-1522.

8 Anderson, M.L.M., Szajnert, M.F., Kaplan, J.C., Mccoll, L. & Young, B.D. (1984) Nuc.Acids Res., 12, 6647-6661.

References of human germline heavy chain sequences

[0151]

1 Adderson, E.E., Azmi, F.H., Wilson, P.M., Shackelford, P.G. & Carroll, W.L. (1993) J.Immunol., 151, 800-809.

2 Andris. J.S., Brodeur, B.R. & Capra, J.D. (1993) Mol.Immunol., 30, 1601-1616.

3 Berman, J.E., Mellis, S.J., Pollock, R., Smith, C.L., Suh, H., Heinke, B., Kowal, C., Surti, U., Chess, L., Cantor, C.R & Alt, F.W. (1988) Embo J., 7, 727-738.

4 Buluwela, L. & Rabbitts, T.H. (1988) Eur.J.Immunol., 18, 1843-1845.; Buluwela, L., Albertson, D.G., Sherrittgton, P., Rabbitts, P.H., Spurr, N. & Rabbitts, T.H. (1988) Embo J., 7, 2003-2010.

5 Chen, P.P., Liu, M.-F., Sinha, S. & Carson, D.A. (1988) Arth.Rheum., 31, 1429-1431.

6 Chen, P.P., Liu, M.-F., Glass, C.A., Sinha, S., Kipps, T.J. & Carson, D.A. (1989) Arthritis & Rheumatism, 32, 72-76.

7 Cook, G.P. et al. (1994) Nature Genetics 7, 162-168.

8 Haino, M. et al., (1994). J. Biol. Chem. 269, 2619-2626

9 Humphries, C.G., Shen, A., Kuziel, W.A., Capra, J.D., Blattner, F.R. & Tucker, P.W. (1988) Nature, 331, 446-449.

10 Kodaira, M., Kinashi, T., Umemura, I., Matsuda, F., Noma, T., One. Y. & Honjo, T. (1986) J.Mol.Biol., 190. 529-541.

11 Lee, K.H., Matsuda, F., Kinashi, T., Kodaira, M. & Honjo, T. (1987) J.Mol.Biol., 195, 761-768.

12 Matsuda, F., Lee, K.H., Nakai, S., Sato, T., Kodaira, M., Zong, S.Q., Ohno, H., Fukuhara, S. & Honjo, T. (1988) Embo J., 7, 1047-1051.

13 Matsuda, F., Shin, E.K., Hirabayashi, Y., Nagaoka, H., Yoshida, M.C., Zong, S.Q. & Honjo, T. (1990) Embo J., 9, 2501-2506.

14 Matsuda, F., Shin, E.K., Nagaoka, H., Matsumura, R., Haino, M., Fukita, Y., Taka-Ishi, S., Imai, T., Riley, J.H., Anand, R. Et, Al. (1993) Nature Genet. 3, 88-94

15 Nagaoka, H., Ozawa, K., Matsuda, F., Hayashida, H., Matsumura, R., Haino, M., Shin, E.K., Fukita, Y., Imai, T., Anand, R., Yokoyama, K., Eki, T., Soeda, E. & Honjo, T. (1993). (Temporal)

16 Rechavi, G., Bienz, B., Ram, D., Ben-Neriah, Y., Cohen, J.B., Zakut, R. & Givol, D. (1982) Proc.Nat.Acad.Sci.USA, 79, 4405-4409.

17 Sanz, I., Kelly, P., Williams, C., Scholl, S., Tucker, P. & Capra, J.D. (1989) Embo J., 8, 3741-3748.

18 Shin. E.K., Matsuda. F., Fujikura, J., Akamizu, T., Sugawa, H., Mori, T. & Honjo, T. (1993) Eur.J.Immunol., 23. 2365-2367.

19 Tomlinson, Im., Walter, G., Marks. Jd., Llewelyn, Mb. & Winter. G. (1992) J.MoLBiol. 227. 776-798.

20 Van Der Maarel, S., Van Dijk, K.W., Alexander, C.M., Sasso, E.H., Bull, A. & Milner, E.C.B. (1993) J.Immunol., 150, 2858-2868.

21 Van Dijk, K.W., Mortari, F., Kirkham, P.M., Schroeder, Jr., H.W. & Milner, E.C.B. (1993) Eur.J.Immunol., 23, 832-839.

22 Van Es, J.H., Aanstoot, H., Gmelig-Meyling, F.H.J., Derksen, R.H.W.M. & Logtenberg, T. (1992) J.Immunol., 149, 223 4-2240.

23 Weng, N.-P., Snyder, J.G., Yu-Lee, L.-Y. & Marcus, D.M. (1992) Eur.J.Immunot., 22, 1075-1082.

24 Winkler, T.H., Fehr, H. & Kalden, J.R. (1992) Eur.J.Immunol., 22, 1719-1728.

25 Olee, T., Yang, P.M., Siminovitch, K.A., Olsen, N.J., Hillson, J.L., Wu, J., Kozin, F., Carson, D.A.&Chen, P.P. (1991) J. Clin. Invest. 88, 193-203.

26 Chen, P.P.& Yang, P.M. (1990) Scand. J. Immunol. 31, 593-599.

27 Tomlinson, M., Walter, G., Cook&Winter, G. (Unpublished)

SEQUENCE LISTING

[0152]

(1) GENERAL INFORMATION:

(i) APPLICANT:

(A) NAME: Morphosys Gesellschaft fuer Proteinoptimierung mbH
(B) STREET: Frankfurter Ring 193a
(C) CITY: Muenchen
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 80807

(A) NAME: Knappik, Achim
(B) STREET: Killerstr. 16
(C) CITY: Graefelfing
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 82166

(A) NAME: Pack, Peter
(B) STREET: Franz-Wolter-Str. 4
(C) CITY: Muenchen
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 81925

(A) NAME: Ilag, Vic
(B) STREET: Knorrstr. 85
(C) CITY: Muenchen
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 80807

(A) NAME: Ge, Liming
(B) STREET: Nestroystr. 17
(C) CITY: Muenchen
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 81373

(A) NAME: Moroney, Simon
(B) STREET: Osterwaldstr. 44
(C) CITY: Muenchen
(E) COUNTRY: DE
(F) POSTAL CODE (ZIP): 80805

(A) NAME: Plueckthun, Andreas
(B) STREET: Moehrlistr. 97
(C) CITY: Zuerich
(E) COUNTRY: CH
(F) POSTAL CODE (ZIP): 8006

(ii) TITLE OF INVENTION: Protein/(Poly)peptide libraries

(iii) NUMBER OF SEQUENCES: 371

(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)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: WO PCT/EP96/03647

(vi) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: EP 95 11 3021.0
(B) FILING DATE: 18-AUG-1995

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein

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

Ala Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
1 5 10 15

Gly Gly Gly Ser
20

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

TCAGCGGGTG GCGGTTCTGG CGGCGGTGGG AGCGGTGGCG GTGGTTCTGG CGGTGGTGGT 60

TCCGATATCG GTCCACGTAC GG 82

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

AATTCCGTAC GTGGACCGAT ATCGGAACCA CCACCGCCAG AACCACCGCC ACCGCTCCCA 60

CCGCCGCCAG AACCGCCACC CGC 83

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc feature
(B) LOCATION:28..45
(D) OTHER INFORMATION:/product= "6 random codons by trinucleotide mutagenesis (19aa, no Cys)"

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

GATACGGCCG TGTATTATTG CGCGCGTNNK NNKNNKNNKN NKNNKGATTA TTGGGGCCAA 60

GGCACCCTG 69

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 84 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:28..57
(D) OTHER INFORMATION:/product= "10 random codons by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:58..60
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (TTT/ATG)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:64..66
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (GTT/TAT)"

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

GATACGGCCG TGTATTATTG CGCGCGTNNK NNKNNKNNKN NKNNKNNKNN KNNKNNKWTK 60

GATKWTTGGG GCCAAGGCAC CCTG 84

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21 base pairs

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
GATACGGCCG TGTATTATTG C 21

(2) INFORMATION FOR SEQ ID NO: 7:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
CAGGGTGCCT TGGCCCC 17

(2) INFORMATION FOR SEQ ID NO: 8:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
GCAGAAGGCG AACGTCC 17

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 80 base pairs
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:39..41
(D) OTHER INFORMATION:/product= "random codon (mixture of GCT, CGT, CAT, TCT, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:42..53
(D) OTHER INFORMATION:/product= "random codons by trinucleotide mutagenesis (19 aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:57..59
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19 aa, no Cys)"

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

TGGAAGCTGA AGACGTGGGC GTGTATTATT GCCAGCAGBV TNNKNNKNNK NNKCCGNNKT 60

TTGGCCAGGG TACGAAAGTT 80

(2) INFORMATION FOR SEQ ID NO: 10:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
AACTTTCGTA CCCTGGCC 18

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:21..23
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:27..35
(D) OTHER INFORMATION:/product= "random codons by trinucleotide mutagenesis (19 aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:36..41
(D) OTHER INFORMATION:/product= "random codons by mixed monomers (A/G A/C/G T)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:42..44
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:48..50
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

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

AGGGTCTCGA GTGGGTGAGC NNKATTNNKN NKNNKRVTRV TNNKACCNNK TATGCGGATA 60

GCGTGAAAGG CCGTTTTACC ATTTCACGTG ATAATTCGAA AAACACCA 108

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:21..23
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:27..32
(D) OTHER INFORMATION:/product= "random codons by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:33..38
(D) OTHER INFORMATION:/product= "random codons by mixed monomers (A/G A/C/G T)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:39..41
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:45..47
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

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

AGGGTCTCGA GTGGGTGAGC NNKATTNNKN NKRVTRVTNN KACCNNKTAT GCGGATAGCG 60

TGAAAGGCCG TTTTACCATT TCACGTGATA ATTCGAAAAA CACCA 105

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:
TGGTGTTTTT CGAATTATCA 20

(2) INFORMATION FOR SEQ ID NO: 14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

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

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Ser Ile Ser Asn Tyr
            20                  25                  30

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

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln Tyr Tyr Ser Thr Pro Leu
                85                  90                  95

Thr Phe Gly Gly Gly Thr Lys Val Glu Ile Lys Arg
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

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

Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
            20                  25                  30

Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln Ser
        35                  40                  45

Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg Ala Ser Gly Val Pro
    50                  55                  60

Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65                  70                  75                  80

Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Met Gln Ala
                85                  90                  95

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

Arg
```

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
            20                  25                  30

Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu
        35                  40                  45
Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Ile Pro Asp Arg Phe Ser
    50                  55                  60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Arg Leu Glu
65                  70                  75                  80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln Tyr Gly Asn Ser Pro
                85                  90                  95

Tyr Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

```
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 Val Leu Tyr Ser
            20                  25                  30

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

Lys Arg
```

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

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

Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Asn Asn
            20                  25                  30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
        35                  40                  45

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

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Thr Gly Leu Gln
65                  70                  75                  80

Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Thr Trp Asp Asp Ser Leu
                85                  90                  95

Ser Gly Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 112 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

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

Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr
            20                  25                  30

Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu
        35                  40                  45

Met Ile Tyr Asp Val Ser Lys Arg Pro Ser Gly Val Ser Asn Arg Phe
    50                  55                  60

Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu
65                  70                  75                  80

Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Ser Ser Tyr Ala Gly Ser
                85                  90                  95

Ser Thr Val Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

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

Ala Arg Ile Thr Cys Ser Gly Asp Ser Leu Gly Ser Lys Tyr Ala Ser
            20                  25                  30
```

```
Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Ile Tyr Asp
        35                  40                  45

Asp Asn Lys Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser Asn
    50                  55                  60

Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Gly Val Gln Ala Glu Asp
65                  70                  75                  80

Glu Ala Asp Tyr Tyr Cys Gln Ser Trp Asp Ser Ser Gly Asn Val Val
                85                  90                  95

Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly Gln
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

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

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45

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

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr 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 Ala Pro Gly Tyr Cys Ser Gly Phe Asp Tyr Trp Gly Gln Gly
            100                 105                 110

Thr Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

```
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 Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
            20                  25                  30

Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45

Gly Trp Ile Asn Pro Asn Ser Gly Asn Thr Asn Tyr Ala Gln Lys Phe
    50                  55                  60

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

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

Ala Arg Asp Gly Asp Gly Gly Phe Asp Tyr Trp Gly Gln Gly Thr Leu
            100                 105                 110

Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

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

Thr Leu Thr Leu Thr Cys Thr Phe Ser Gly Phe Ser Leu Ser Thr Ser
            20                  25                  30

Gly Met Gly Val Ser Trp Ile Arg Gln Pro Pro Gly Lys Ala Leu Glu
        35                  40                  45

Trp Leu Ala His Ile Asp Trp Asp Asp Asp Lys Tyr Tyr Ser Thr Ser
    50                  55                  60
```

```
65                  70                  75                  80

Val Leu Thr Met Thr Asn Met Asp Pro Val Asp Thr Ala Thr Tyr Tyr
                85                  90                  95

Cys Ala Arg Ile His Asn Ile Gly Glu Ala Phe Asp Val Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 24:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

```
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1               5                   10                  15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45

Ser Val Ile Ser Tyr Asp Gly Gly Asn Thr Tyr Tyr Ala Asp Ser Val
    50                  55                  60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65                  70                  75                  80

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

Ala Arg Asp Arg Gly Gly Ser Gly Asp Tyr Trp Gly Gln Gly Thr Leu
            100                 105                 110

Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

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

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr
            20                  25                  30

Tyr Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
        35                  40                  45

Gly Glu Ile Tyr His Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
    50                  55                  60

Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
65                  70                  75                  80

Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
                85                  90                  95

Arg Gly Arg Gly Gly Gly Gly Val Phe Asp Tyr Trp Gly Gln Gly Thr
            100                 105                 110

Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

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

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr
            20                  25                  30

Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
        35                  40                  45

Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe
    50                  55                  60

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

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
                85                  90                  95
```

```
Ala Arg Leu Gly Gly Gly Gly Tyr Tyr Phe Asp Tyr Trp Gly Gln Gly
            100                 105                 110

Thr Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

```
Gln Val Gln Leu Gln Gln Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
1               5                   10                  15

Thr Leu Ser Leu Thr Cys Ala Ile Ser Gly Asp Ser Val Ser Ser Asn
            20                  25                  30

Ser Ala Ala Trp Asn Trp Ile Arg Gln Ser Pro Ser Arg Gly Leu Glu
        35                  40                  45

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

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

Gln Phe Ser Leu Gln Leu Asn Ser Val Thr Pro Glu Asp Thr Ala Val
                85                  90                  95

Tyr Tyr Cys Ala Arg Asp Pro Gly Gly Phe Asp Val Trp Gly Gln Gly
            100                 105                 110

Thr Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

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

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Tyr
            20                  25                  30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
        35                  40                  45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
                85                  90                  95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

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

Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
            20                  25                  30

Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln Ser
        35                  40                  45

Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg Ala Ser Gly Val Pro
    50                  55                  60

Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65                  70                  75                  80

Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Gln Gln His
                85                  90                  95

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

Arg Thr
```

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
            20                  25                  30

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

Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Val Pro Ala Arg Phe Ser
    50                  55                  60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu
65                  70                  75                  80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
                85                  90                  95

Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

```
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 Arg Ser Ser Gln Ser Val Leu Tyr Ser
            20                  25                  30

Ser Asn Asn 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 Ser 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

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

Lys Arg Thr
        115
```

(2) INFORMATION FOR SEQ ID NO: 32:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

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

Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn
            20                  25                  30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
        35                  40                  45

Ile Tyr Asp Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser
    50                  55                  60

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Thr Gly Leu Gln
65                  70                  75                  80

Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
                85                  90                  95

Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 33:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

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

Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr
            20                  25                  30

Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu
        35                  40                  45

Met Ile Tyr Asp Val Ser Asn Arg Pro Ser Gly Val Ser Asn Arg Phe
    50                  55                  60

Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu
65                  70                  75                  80

Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr
                85                  90                  95

Pro Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 34:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

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

Thr Ala Arg Ile Ser Cys Ser Gly Asp Ala Leu Gly Asp Lys Tyr Ala
            20                  25                  30

Ser Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Ile Tyr
        35                  40                  45

Asp Asp Ser Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
    50                  55                  60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Gly Thr Gln Ala Glu
65                  70                  75                  80

Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro Val
                85                  90                  95

Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 35:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

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

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45

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

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr 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 Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 36:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

```
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 Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
            20                  25                  30

Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45
Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe
    50                  55                  60

Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile 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 Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

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

Thr Leu Thr Leu Thr Cys Thr Phe Ser Gly Phe Ser Leu Ser Thr Ser
            20                  25                  30

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

Trp Leu Ala Leu Ile Asp Trp Asp Asp Asp Lys Tyr Tyr Ser Thr Ser
    50                  55                  60

Leu Lys Thr Arg Leu Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Val
65                  70                  75                  80

Val Leu Thr Met Thr Asn Met Asp Pro Val Asp Thr Ala Thr Tyr Tyr
                85                  90                  95

Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly
            100                 105                 110

Gln Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

```
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1               5                  10                  15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
    50                  55                  60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65                  70                  75                  80

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

Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

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

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr
            20                  25                  30
```

```
        35                  40                  45

Gly Tyr Ile Tyr Tyr Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
    50                  55                  60

Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
65                  70                  75                  80

Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
                85                  90                  95

Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln Gly
            100                 105                 110

Thr Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

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

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr
            20                  25                  30

Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
        35                  40                  45

Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe
    50                  55                  60

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

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
                85                  90                  95

Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids
(B) TYPE: amino acid

(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

```
Gln Val Gln Leu Gln Gln Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
1               5                  10                  15

Thr Leu Ser Leu Thr Cys Ala Ile Ser Gly Asp Ser Val Ser Ser Asn
            20                  25                  30

Ser Ala Ala Trp Asn Trp Ile Arg Gln Ser Pro Gly Arg Gly Leu Glu
        35                  40                  45

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

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

Gln Phe Ser Leu Gln Leu Asn Ser Val Thr Pro Glu Asp Thr Ala Val
                85                  90                  95

Tyr Tyr Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr
            100                 105                 110

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: 327 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..327
(D) OTHER INFORMATION:/product= "V kappa 1"

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

```
GAT ATC CAG ATG ACC CAG AGC CCG TCT AGC CTG AGC GCG AGC GTG GGT        48
Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser Val Gly
  1               5                  10                  15

GAT CGT GTG ACC ATT ACC TGC AGA GCG AGC CAG GGC ATT AGC AGC TAT        96
```

```
Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Tyr
            20                  25                  30

CTG GCG TGG TAC CAG CAG AAA CCA GGT AAA GCA CCG AAA CTA TTA ATT      144
Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
        35                  40                  45

TAT GCA GCC AGC AGC TTG CAA AGC GGG GTC CCG TCC CGT TTT AGC GGC      192
Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

TCT GGA TCC GGC ACT GAT TTT ACC CTG ACC ATT AGC AGC CTG CAA CCT      240
Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

GAA GAC TTT GCG ACC TAT TAT TGC CAG CAG CAT TAT ACC ACC CCG CCG      288
Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
                85                  90                  95

ACC TTT GGC CAG GGT ACG AAA GTT GAA ATT AAA CGT ACG                  327
Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

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

Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Gln Gly Ile Ser Ser Tyr
            20                  25                  30

Leu Ala Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu Leu Ile
        35                  40                  45

Tyr Ala Ala Ser Ser Leu Gln Ser Gly Val Pro Ser Arg Phe Ser Gly
    50                  55                  60

Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Gln Pro
65                  70                  75                  80

Glu Asp Phe Ala Thr Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro
                85                  90                  95

Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 342 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc * "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..342
(D) OTHER INFORMATION:/product- "V kappa 2"

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

```
GAT ATC GTG ATG ACC CAG AGC CCA CTG AGC CTG CCA GTG ACT CCG GGC          48
Asp Ile Val Met Thr Gln Ser Pro Leu Ser Leu Pro Val Thr Pro Gly
110                 115                 120                 125

GAG CCT GCG AGC ATT AGC TGC AGA AGC AGC CAA AGC CTG CTG CAT AGC          96
Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
                130                 135                 140

AAC GGC TAT AAC TAT CTG GAT TGG TAC CTT CAA AAA CCA GGT CAA AGC         144
Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln Ser
            145                 150                 155

CCG CAG CTA TTA ATT TAT CTG GGC AGC AAC CGT GCC AGT GGG GTC CCG         192
        160                 165                 170

GAT CGT TTT AGC GGC TCT GGA TCC GGC ACC GAT TTT ACC CTG AAA ATT         240
Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
    175                 180                 185

AGC CGT GTG GAA GCT GAA GAC GTG GGC GTG TAT TAT TGC CAG CAG CAT         288
Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Gln Gln His
190                 195                 200                 205

TAT ACC ACC CCG CCG ACC TTT GGC CAG GGT ACG AAA GTT GAA ATT AAA         336
Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys
                210                 215                 220

CGT ACG                                                                342
Arg Thr
```

(2) INFORMATION FOR SEQ ID NO: 45:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

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

Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln Ser Leu Leu His Ser
            20                  25                  30

Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln Lys Pro Gly Gln Ser
        35                  40                  45

Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg Ala Ser Gly Val Pro
    50                  55                  60

Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Lys Ile
65                  70                  75                  80

Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr Tyr Cys Gln Gln His
                85                  90                  95

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

Arg Thr
```

(2) INFORMATION FOR SEQ ID NO: 46:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..330
(D) OTHER INFORMATION:/product= "V kappa 3"

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

```
GAT ATC GTG CTG ACC CAG AGC CCG GCG ACC CTG AGC CTG TCT CCG GGC        48
Asp Ile Val Leu Thr Gln Ser Pro Ala Thr Leu Ser Leu Ser Pro Gly
115                 120                 125                 130

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
                135                 140                 145

TAT CTG GCG TGG TAC CAG CAG AAA CCA GGT CAA GCA CCG CGT CTA TTA        144
Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Arg Leu Leu
            150                 155                 160

ATT TAT GGC GCG AGC AGC CGT GCA ACT GGG GTC CCG GCG CGT TTT AGC        192
Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Val Pro Ala Arg Phe Ser
        165                 170                 175

GGC TCT GGA TCC GGC ACG GAT TTT ACC CTG ACC ATT AGC AGC CTG GAA        240
Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu
    180                 185                 190

CCT GAA GAC TTT GCG GTG TAT TAT TGC CAG CAG CAT TAT ACC ACC CCG        288
Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
195                 200                 205                 210

CCG ACC TTT GGC CAG GGT ACG AAA GTT GAA ATT AAA CGT ACG                330
Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
                215                 220
```

(2) INFORMATION FOR SEQ ID NO: 47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:

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

Glu Arg Ala Thr Leu Ser Cys Arg Ala Ser Gln Ser Val Ser Ser Ser
            20                  25                  30

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

Ile Tyr Gly Ala Ser Ser Arg Ala Thr Gly Val Pro Ala Arg Phe Ser
    50                  55                  60

Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr Ile Ser Ser Leu Glu
65                  70                  75                  80

Pro Glu Asp Phe Ala Val Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
                85                  90                  95

Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile Lys Arg Thr
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 345 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..345
(D) OTHER INFORMATION:/product= "V kappa 4"

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

```
GAT ATC GTG ATG ACC CAG AGC CCG GAT AGC CTG GCG GTG AGC CTG GGC          48
Asp Ile Val Met Thr Gln Ser Pro Asp Ser Leu Ala Val Ser Leu Gly
                115                 120                 125

GAA CGT GCG ACC ATT AAC TGC AGA AGC AGC CAG AGC GTG CTG TAT AGC          96
Glu Arg Ala Thr Ile Asn Cys Arg Ser Ser Gln Ser Val Leu Tyr Ser
            130                 135                 140

AGC AAC AAC AAA AAC TAT CTG GCG TGG TAC CAG CAG AAA CCA GGT CAG         144
Ser Asn Asn Lys Asn Tyr Leu Ala Trp Tyr Gln Gln Lys Pro Gly Gln
        145                 150                 155

CCG CCG AAA CTA TTA ATT TAT TGG GCA TCC ACC CGT GAA AGC GGG GTC         192
Pro Pro Lys Leu Leu Ile Tyr Trp Ala Ser Thr Arg Glu Ser Gly Val
    160                 165                 170

CCG GAT CGT TTT AGC GGC TCT GGA TCC GGC ACT GAT TTT ACC CTG ACC         240
Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp Phe Thr Leu Thr
175                 180                 185                 190

ATT TCG TCC CTG CAA GCT GAA GAC GTG GCG GTG TAT TAT TGC CAG CAG         288
Ile Ser Ser Leu Gln Ala Glu Asp Val Ala Val Tyr Tyr Cys Gln Gln
                195                 200                 205

CAT TAT ACC ACC CCG CCG ACC TTT GGC CAG GGT ACG AAA GTT GAA ATT         336
His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr Lys Val Glu Ile
            210                 215                 220

AAA CGT ACG                                                             345
Lys Arg Thr
        225
```

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 115 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

```
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 Arg Ser Ser Gln Ser Val Leu Tyr Ser
            20                  25                  30

Ser Asn Asn 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 Ser 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

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

Lys Arg Thr
        115
```

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 327 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..327
(D) OTHER INFORMATION:/Product= "V lambda 1"

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

```
CAG AGC GTG CTG ACC CAG CCG CCT TCA GTG AGT GGC GCA CCA GGT CAG        48
Gln Ser Val Leu Thr Gln Pro Pro Ser Val Ser Gly Ala Pro Gly Gln
                120                 125                 130

CGT GTG ACC ATC TCG TGT AGC GGC AGC AGC AGC AAC ATT GGC AGC AAC        96
Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn
            135                 140                 145

TAT GTG AGC TGG TAC CAG CAG TTG CCC GGG ACG GCG CCG AAA CTG CTG       144
Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
        150                 155                 160

ATT TAT GAT AAC AAC CAG CGT CCC TCA GGC GTG CCG GAT CGT TTT AGC       192
Ile Tyr Asp Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser
    165                 170                 175

GGA TCC AAA AGC GGC ACC AGC GCG AGC CTT GCG ATT ACG GGC CTG CAA       240
Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Thr Gly Leu Gln
180                 185                 190                 195

AGC GAA GAC GAA GCG GAT TAT TAT TGC CAG CAG CAT TAT ACC ACC CCG       288
Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
                200                 205                 210

CCT GTG TTT GGC GGC GGC ACG AAG TTA ACC GTT CTT GGC                   327
Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            215                 220
```

(2) INFORMATION FOR SEQ ID NO: 51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 109 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

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

Arg Val Thr Ile Ser Cys Ser Gly Ser Ser Ser Asn Ile Gly Ser Asn
             20                  25                  30

Tyr Val Ser Trp Tyr Gln Gln Leu Pro Gly Thr Ala Pro Lys Leu Leu
         35                  40                  45

Ile Tyr Asp Asn Asn Gln Arg Pro Ser Gly Val Pro Asp Arg Phe Ser
     50                  55                  60

Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala Ile Thr Gly Leu Gln
 65                  70                  75                  80

Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro
                 85                  90                  95

Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 52:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 330 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(D) OTHER INFORMATION:/product= "V lambda 2"

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

```
CAG AGC GCA CTG ACC CAG CCA GCT TCA GTG AGC GGC TCA CCA GGT CAG        48
Gln Ser Ala Leu Thr Gln Pro Ala Ser Val Ser Gly Ser Pro Gly Gln
110                 115                 120                 125

AGC ATT ACC ATC TCG TGT ACG GGT ACT AGC AGC GAT GTG GGC GGC TAT        96
Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr
                130                 135                 140

AAC TAT GTG AGC TGG TAC CAG CAG CAT CCC GGG AAG GCG CCG AAA CTG       144
Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu
            145                 150                 155

ATG ATT TAT GAT GTG AGC AAC CGT CCC TCA GGC GTG AGC AAC CGT TTT       192
Met Ile Tyr Asp Val Ser Asn Arg Pro Ser Gly Val Ser Asn Arg Phe

        160                 165                 170

AGC GGA TCC AAA AGC GGC AAC ACC GCG AGC CTG ACC ATT AGC GGC CTG       240
Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu
    175                 180                 185

CAA GCG GAA GAC GAA GCG GAT TAT TAT TGC CAG CAG CAT TAT ACC ACC       288
Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr
190                 195                 200                 205

CCG CCT GTG TTT GGC GGC GGC ACG AAG TTA ACC GTT CTT GGC               330
Pro Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
                210                 215
```

(2) INFORMATION FOR SEQ ID NO: 53:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 110 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

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

Ser Ile Thr Ile Ser Cys Thr Gly Thr Ser Ser Asp Val Gly Gly Tyr
            20                  25                  30

Asn Tyr Val Ser Trp Tyr Gln Gln His Pro Gly Lys Ala Pro Lys Leu
        35                  40                  45

Met Ile Tyr Asp Val Ser Asn Arg Pro Ser Gly Val Ser Asn Arg Phe
    50                  55                  60

Ser Gly Ser Lys Ser Gly Asn Thr Ala Ser Leu Thr Ile Ser Gly Leu
65                  70                  75                  80

Gln Ala Glu Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr
                85                  90                  95

Pro Pro Val Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            100                 105                 110
```

(2) INFORMATION FOR SEQ ID NO: 54:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..321
(D) OTHER INFORMATION:/product= "V lambda 3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

```
AGC TAT GAA CTG ACC CAG CCG CCT TCA GTG AGC GTT GCA CCA GGT CAG        48
Ser Tyr Glu Leu Thr Gln Pro Pro Ser Val Ser Val Ala Pro Gly Gln
                115                 120                 125

ACC GCG CGT ATC TCG TGT AGC GGC GAT GCG CTG GGC GAT AAA TAC GCG        96
Thr Ala Arg Ile Ser Cys Ser Gly Asp Ala Leu Gly Asp Lys Tyr Ala
            130                 135                 140

AGC TGG TAC CAG CAG AAA CCC GGG CAG GCG CCA GTT CTG GTG ATT TAT       144
Ser Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Ile Tyr
        145                 150                 155

GAT GAT TCT GAC CGT CCC TCA GGC ATC CCG GAA CGC TTT AGC GGA TCC       192
Asp Asp Ser Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
    160                 165                 170

AAC AGC GGC AAC ACC GCG ACC CTG ACC ATT AGC GGC ACT CAG GCG GAA       240
Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Gly Thr Gln Ala Glu
175                 180                 185                 190

GAC GAA GCG GAT TAT TAT TGC CAG CAG CAT TAT ACC ACC CCG CCT GTG       288
Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro Val
                195                 200                 205

TTT GGC GGC GGC ACG AAG TTA ACC GTT CTT GGC                          321
Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
            210                 215
```

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 107 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

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

Thr Ala Arg Ile Ser Cys Ser Gly Asp Ala Leu Gly Asp Lys Tyr Ala
             20                  25                  30

Ser Trp Tyr Gln Gln Lys Pro Gly Gln Ala Pro Val Leu Val Ile Tyr
         35                  40                  45

Asp Asp Ser Asp Arg Pro Ser Gly Ile Pro Glu Arg Phe Ser Gly Ser
     50                  55                  60

Asn Ser Gly Asn Thr Ala Thr Leu Thr Ile Ser Gly Thr Gln Ala Glu
 65                  70                  75                  80
Asp Glu Ala Asp Tyr Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro Val
                 85                  90                  95
```

```
Phe Gly Gly Gly Thr Lys Leu Thr Val Leu Gly
100                 105
```

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..360
(D) OTHER INFORMATION:/product= "VH1A"

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

```
CAG GTG CAA TTG GTT CAG TCT GGC GCG GAA GTG AAA AAA CCG GGC AGC        48
Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ser
        110                 115                 120

AGC GTG AAA GTG AGC TGC AAA GCC TCC GGA GGC ACT TTT AGC AGC TAT        96
Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
    125                 130                 135

GCG ATT AGC TGG GTG CGC CAA GCC CCT GGG CAG GGT CTC GAG TGG ATG       144
Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
140                 145                 150                 155

GGC GGC ATT ATT CCG ATT TTT GGC ACG GCG AAC TAC GCG CAG AAG TTT       192
Gly Gly Ile Ile Pro Ile Phe Gly Thr Ala Asn Tyr Ala Gln Lys Phe
                160                 165                 170

CAG GGC CGG GTG ACC ATT ACC GCG GAT GAA AGC ACC AGC ACC GCG TAT       240
Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr Ser Thr Ala Tyr
            175                 180                 185

ATG GAA CTG AGC AGC CTG CGT AGC GAA GAT ACG GCC GTG TAT TAT TGC       288
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
        190                 195                 200

GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC CAA       336
Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
    205                 210                 215

GGC ACC CTG GTG ACG GTT AGC TCA G                                    361
Gly Thr Leu Val Thr Val Ser Ser
220                 225
```

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

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

Ser Val Lys Val Ser Cys Lys Ala Ser Gly Gly Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Ile Ser Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45

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

Gln Gly Arg Val Thr Ile Thr Ala Asp Glu Ser Thr 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 Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..360
(D) OTHER INFORMATION:/product= "VH1B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

```
CAG GTG CAA TTG GTT CAG AGC GGC GCG GAA GTG AAA AAA CCG GGC GCG          48
Gln Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Ala
                125                 130                 135

AGC GTG AAA GTG AGC TGC AAA GCC TCC GGA TAT ACC TTT ACC AGC TAT          96
Ser Val Lys Val Ser Cys Lys Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
            140                 145                 150

TAT ATG CAC TGG GTC CGC CAA GCC CCT GGG CAG GGT CTC GAG TGG ATG         144
```

```
Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        155                 160                 165

GGC TGG ATT AAC CCG AAT AGC GGC GGC ACG AAC TAC GCG CAG AAG TTT         192
Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe
    170                 175                 180

CAG GGC CGG GTG ACC ATG ACC CGT GAT ACC AGC ATT AGC ACC GCG TAT         240
Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile Ser Thr Ala Tyr
185                 190                 195                 200

ATG GAA CTG AGC AGC CTG CGT AGC GAA GAT ACG GCC GTG TAT TAT TGC         288
Met Glu Leu Ser Ser Leu Arg Ser Glu Asp Thr Ala Val Tyr Tyr Cys
                205                 210                 215

GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC CAA         336
Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            220                 225                 230

GGC ACC CTG GTG ACG GTT AGC TCA G                                      361
Gly Thr Leu Val Thr Val Ser Ser
        235                 240
```

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

```
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 Ala Ser Gly Tyr Thr Phe Thr Ser Tyr
            20                  25                  30

Tyr Met His Trp Val Arg Gln Ala Pro Gly Gln Gly Leu Glu Trp Met
        35                  40                  45

Gly Trp Ile Asn Pro Asn Ser Gly Gly Thr Asn Tyr Ala Gln Lys Phe
    50                  55                  60

Gln Gly Arg Val Thr Met Thr Arg Asp Thr Ser Ile 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 Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 364 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..363
(D) OTHER INFORMATION:/product= "VH2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

```
CAG GTG CAA TTG AAA GAA AGC GGC CCG GCC CTG GTG AAA CCG ACC CAA        48
Gln Val Gln Leu Lys Glu Ser Gly Pro Ala Leu Val Lys Pro Thr Gln
                125                 130                 135

ACC CTG ACC CTG ACC TGT ACC TTT TCC GGA TTT AGC CTG TCC ACG TCT        96
Thr Leu Thr Leu Thr Cys Thr Phe Ser Gly Phe Ser Leu Ser Thr Ser
            140                 145                 150

GGC GTT GGC GTG GGC TGG ATT CGC CAG CCG CCT GGG AAA GCC CTC GAG       144
Gly Val Gly Val Gly Trp Ile Arg Gln Pro Pro Gly Lys Ala Leu Glu
        155                 160                 165

TGG CTG GCT CTG ATT GAT TGG GAT GAT GAT AAG TAT TAT AGC ACC AGC       192
Trp Leu Ala Leu Ile Asp Trp Asp Asp Asp Lys Tyr Tyr Ser Thr Ser
    170                 175                 180

CTG AAA ACG CGT CTG ACC ATT AGC AAA GAT ACT TCG AAA AAT CAG GTG       240
Leu Lys Thr Arg Leu Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Val
185                 190                 195                 200

GTG CTG ACT ATG ACC AAC ATG GAC CCG GTG GAT ACG GCC ACC TAT TAT       288
Val Leu Thr Met Thr Asn Met Asp Pro Val Asp Thr Ala Thr Tyr Tyr
                205                 210                 215

TGC GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC       336
Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly
            220                 225                 230

CAA GGC ACC CTG GTG ACG GTT AGC TCA G                                 364
Gln Gly Thr Leu Val Thr Val Ser Ser
        235                 240
```

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 121 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

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

Thr Leu Thr Leu Thr Cys Thr Phe Ser Gly Phe Ser Leu Ser Thr Ser
            20                  25                  30

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

Trp Leu Ala Leu Ile Asp Trp Asp Asp Asp Lys Tyr Tyr Ser Thr Ser
    50                  55                  60

Leu Lys Thr Arg Leu Thr Ile Ser Lys Asp Thr Ser Lys Asn Gln Val
65                  70                  75                  80

Val Leu Thr Met Thr Asn Met Asp Pro Val Asp Thr Ala Thr Tyr Tyr
                85                  90                  95

Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly
            100                 105                 110

Gln Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..360
(D) OTHER INFORMATION:/product= "VH3"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

```
GAA GTG CAA TTG GTG GAA AGC GGC GGC GGC CTG GTG CAA CCG GGC GGC        48
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
            125                 130                 135

AGC CTG CGT CTG AGC TGC GCG GCC TCC GGA TTT ACC TTT AGC AGC TAT        96
Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
        140                 145                 150

GCG ATG AGC TGG GTG CGC CAA GCC CCT GGG AAG GGT CTC GAG TGG GTG       144
Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
    155                 160                 165

AGC GCG ATT AGC GGT AGC GGC GGC AGC ACC TAT TAT GCG GAT AGC GTG       192
Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
170                 175                 180                 185

AAA GGC CGT TTT ACC ATT TCA CGT GAT AAT TCG AAA AAC ACC CTG TAT       240
Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
                190                 195                 200

CTG CAA ATG AAC AGC CTG CGT GCG GAA GAT ACG GCC GTG TAT TAT TGC       288
Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
            205                 210                 215

GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC CAA       336
Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
        220                 225                 230

GGC ACC CTG GTG ACG GTT AGC TCA G                                    361
Gly Thr Leu Val Thr Val Ser Ser
    235                 240
```

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

```
Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
 1               5                  10                  15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Ser Tyr
            20                  25                  30

Ala Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45

Ser Ala Ile Ser Gly Ser Gly Gly Ser Thr Tyr Tyr Ala Asp Ser Val
    50                  55                  60

Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65                  70                  75                  80

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

Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 64:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 358 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..357
(D) OTHER INFORMATION:/product= "VH4"

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

```
CAG GTG CAA TTG CAA GAA AGT GGT CCG GGC CTG GTG AAA CCG AGC GAA        48
Gln Val Gln Leu Gln Glu Ser Gly Pro Gly Leu Val Lys Pro Ser Glu
                125                 130                 135

ACC CTG AGC CTG ACC TGC ACC GTT TCC GGA GGC AGC ATT AGC AGC TAT        96
Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr
            140                 145                 150

TAT TGG AGC TGG ATT CGC CAG CCG CCT GGG AAG GGT CTC GAG TGG ATT       144
Tyr Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
        155                 160                 165

GGC TAT ATT TAT TAT AGC GGC AGC ACC AAC TAT AAT CCG AGC CTG AAA       192
Gly Tyr Ile Tyr Tyr Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
    170                 175                 180

AGC CGG GTG ACC ATT AGC GTT GAT ACT TCG AAA AAC CAG TTT AGC CTG       240
Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
185                 190                 195                 200

AAA CTG AGC AGC GTG ACG GCG GCG GAT ACG GCC GTG TAT TAT TGC GCG       288
Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
                205                 210                 215

CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC CAA GGC       336
Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln Gly
            220                 225                 230

ACC CTG GTG ACG GTT AGC TCA G                                        358
Thr Leu Val Thr Val Ser Ser
        235
```

(2) INFORMATION FOR SEQ ID NO: 65:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 119 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

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

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

Thr Leu Ser Leu Thr Cys Thr Val Ser Gly Gly Ser Ile Ser Ser Tyr
             20                  25                  30
```

```
Tyr Trp Ser Trp Ile Arg Gln Pro Pro Gly Lys Gly Leu Glu Trp Ile
         35                  40                  45

Gly Tyr Ile Tyr Tyr Ser Gly Ser Thr Asn Tyr Asn Pro Ser Leu Lys
     50                  55                  60

Ser Arg Val Thr Ile Ser Val Asp Thr Ser Lys Asn Gln Phe Ser Leu
 65                  70                  75                  80

Lys Leu Ser Ser Val Thr Ala Ala Asp Thr Ala Val Tyr Tyr Cys Ala
                 85                  90                  95

Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln Gly
            100                 105                 110

Thr Leu Val Thr Val Ser Ser
        115
```

(2) INFORMATION FOR SEQ ID NO: 66:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 361 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..360
(D) OTHER INFORMATION:/product= "VH5"

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

```
GAA GTG CAA TTG GTT CAG AGC GGC GCG GAA GTG AAA AAA CCG GGC GAA        48
Glu Val Gln Leu Val Gln Ser Gly Ala Glu Val Lys Lys Pro Gly Glu
120                 125                 130                 135

AGC CTG AAA ATT AGC TGC AAA GGT TCC GGA TAT TCC TTT ACG AGC TAT        96
Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr
                140                 145                 150

TGG ATT GGC TGG GTG CGC CAG ATG CCT GGG AAG GGT CTC GAG TGG ATG       144
Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
            155                 160                 165

GGC ATT ATT TAT CCG GGC GAT AGC GAT ACC CGT TAT TCT CCG AGC TTT       192
Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe
        170                 175                 180

CAG GGC CAG GTG ACC ATT AGC GCG GAT AAA AGC ATT AGC ACC GCG TAT       240
Gln Gly Gln Val Thr Ile Ser Ala Asp Lys Ser Ile Ser Thr Ala Tyr
    185                 190                 195

CTT CAA TGG AGC AGC CTG AAA GCG AGC GAT ACG GCC ATG TAT TAT TGC       288

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
200                 205                 210                 215

GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT TGG GGC CAA       336
Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln

GGC ACC CTG GTG ACG GTT AGC TCA G                                     361
Gly Thr Leu Val Thr Val Ser Ser
            235
```

(2) INFORMATION FOR SEQ ID NO: 67:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 67:

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

Ser Leu Lys Ile Ser Cys Lys Gly Ser Gly Tyr Ser Phe Thr Ser Tyr
            20                  25                  30

Trp Ile Gly Trp Val Arg Gln Met Pro Gly Lys Gly Leu Glu Trp Met
        35                  40                  45

Gly Ile Ile Tyr Pro Gly Asp Ser Asp Thr Arg Tyr Ser Pro Ser Phe
    50                  55                  60

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

Leu Gln Trp Ser Ser Leu Lys Ala Ser Asp Thr Ala Met Tyr Tyr Cys
                85                  90                  95

Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp Gly Gln
            100                 105                 110

Gly Thr Leu Val Thr Val Ser Ser
        115                 120
```

(2) INFORMATION FOR SEQ ID NO: 68:

(i) SEQUENCE CHARACTERISTICS:

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..369
(D) OTHER INFORMATION:/product= "VH6"

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

```
CAG GTG CAA TTG CAA CAG TCT GGT CCG GGC CTG GTG AAA CCG AGC CAA        48
Gln Val Gln Leu Gln Gln Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
                125                 130                 135

ACC CTG AGC CTG ACC TGT GCG ATT TCC GGA GAT AGC GTG AGC AGC AAC        96
Thr Leu Ser Leu Thr Cys Ala Ile Ser Gly Asp Ser Val Ser Ser Asn
            140                 145                 150

AGC GCG GCG TGG AAC TGG ATT CGC CAG TCT CCT GGG CGT GGC CTC GAG       144
Ser Ala Ala Trp Asn Trp Ile Arg Gln Ser Pro Gly Arg Gly Leu Glu
        155                 160                 165

TGG CTG GGC CGT ACC TAT TAT CGT AGC AAA TGG TAT AAC GAT TAT GCG       192
Trp Leu Gly Arg Thr Tyr Tyr Arg Ser Lys Trp Tyr Asn Asp Tyr Ala
    170                 175                 180

GTG AGC GTG AAA AGC CGG ATT ACC ATC AAC CCG GAT ACT TCG AAA AAC       240
Val Ser Val Lys Ser Arg Ile Thr Ile Asn Pro Asp Thr Ser Lys Asn
185                 190                 195                 200

CAG TTT AGC CTG CAA CTG AAC AGC GTG ACC CCG GAA GAT ACG GCC GTG       288
Gln Phe Ser Leu Gln Leu Asn Ser Val Thr Pro Glu Asp Thr Ala Val
                205                 210                 215

TAT TAT TGC GCG CGT TGG GGC GGC GAT GGC TTT TAT GCG ATG GAT TAT       336
Tyr Tyr Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr
            220                 225                 230

TGG GGC CAA GGC ACC CTG GTG ACG GTT AGC TCA G                        370
Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser
        235                 240
```

(2) INFORMATION FOR SEQ ID NO: 69:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 69:

```
Gln Val Gln Leu Gln Gln Ser Gly Pro Gly Leu Val Lys Pro Ser Gln
  1               5                  10                  15

Thr Leu Ser Leu Thr Cys Ala Ile Ser Gly Asp Ser Val Ser Ser Asn
             20                  25                  30

Ser Ala Ala Trp Asn Trp Ile Arg Gln Ser Pro Gly Arg Gly Leu Glu
         35                  40                  45

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

Val Ser Val Lys Ser Arg Ile Thr Ile Asn Pro Asp Thr Ser Lys Asn
```

```
65                  70                  75                  80

Gln Phe Ser Leu Gln Leu Asn Ser Val Thr Pro Glu Asp Thr Ala Val
                85                  90                  95

Tyr Tyr Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr
            100                 105                 110

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

(2) INFORMATION FOR SEQ ID NO: 70:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 70:
GAATGCATAC GCTGATATCC AGATGACCCA GAGCCCGTCT AGCCTGAGC 49

(2) INFORMATION FOR SEQ ID NO: 71:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 71:
CGCTCTGCAG GTAATGGTCA CACGATCACC CACGCTCGCG CTCAGGCTAG ACGGGC 56

(2) INFORMATION FOR SEQ ID NO: 72:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GACCATTACC TGCAGAGCGA GCCAGGGCAT TAGCAGCTAT CTGGCGTGGT ACCAGCAG 58

(2) INFORMATION FOR SEQ ID NO: 73:

(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 = "synthetic oligonucleotide"

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

CTTTGCAAGC TGCTGGCTGC ATAAATTAAT AGTTTCGGTG CTTTACCTGG TTTCTGCTGG 60

TACCACGCCA G 71

(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 = "synthetic oligonucleotide"

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

CAGCCAGCAG CTTGCAAAGC GGGGTCCCGT CCCGTTTTAG CGGCTCTGGA TCCGGCACTG 60

ATTTTAC 67

(2) INFORMATION FOR SEQ ID NO: 75:

(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 = "synthetic oligonucleotide"

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

GATAATAGGT CGCAAAGTCT TCAGGTTGCA GGCTGCTAAT GGTCAGGGTA AAATCAGTGC 60

CGGATCC 67

(2) INFORMATION FOR SEQ ID NO: 76:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 76:
CGATATCGTG ATGACCCAGA GCCCACTGAG CCTGCCAGTG ACTCCGGGCG AGCC 54

(2) INFORMATION FOR SEQ ID NO: 77:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCCGTTGCTA TGCAGCAGGC TTTGGCTGCT TCTGCAGCTA ATGCTCGCAG GCTCGCCCGG 60

AGTCAC 66

(2) INFORMATION FOR SEQ ID NO: 78:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CTGCTGCATA GCAACGGCTA TAACTATCTG GATTGGTACC TTCAAAAACC AGGTCAAAGC 60

CC 62

(2) INFORMATION FOR SEQ ID NO: 79:

(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 = "synthetic oligonucleotide"

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

CGATCCGGGA CCCCACTGGC ACGGTTGCTG CCCAGATAAA TTAATAGCTG CGGGCTTTGA 60

CCTGGTTTTT G 71

(2) INFORMATION FOR SEQ ID NO: 80:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

AGTGGGGTCC CGGATCGTTT TAGCGGCTCT GGATCCGGCA CCGATTTTAC CCTGAAAATT 60

AGCCGTGTG 69

(2) INFORMATION FOR SEQ ID NO: 81:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 54 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 81:
CCATGCAATA ATACACGCCC ACGTCTTCAG CTTCCACACG GCTAATTTTC AGGG 54

(2) INFORMATION FOR SEQ ID NO: 82:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 82:
GAATGCATAC GCTGATATCG TGCTGACCCA GAGCCCGG 38

(2) INFORMATION FOR SEQ ID NO: 83:

(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 = "synthetic oligonucleotide"

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

CGCTCTGCAG CTCAGGGTCG CACGTTCGCC CGGAGACAGG CTCAGGGTCG CCGGGCTCTG 60

GGTCAGC 67

(2) INFORMATION FOR SEQ ID NO: 84:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 84:
CCCTGAGCTG CAGAGCGAGC CAGAGCGTGA GCAGCAGCTA TCTGGCGTGG TACCAG 56

(2) INFORMATION FOR SEQ ID NO: 85:

(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 = "synthetic oligonucleotide"

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

GCACGGCTGC TCGCGCCATA AATTAATAGA CGCGGTGCTT GACCTGGTTT CTGCTGGTAC 60

CACGCCAGAT AG 72

(2) INFORMATION FOR SEQ ID NO: 86:

(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 = "synthetic oligonucleotide"

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

GCGCGAGCAG CCGTGCAACT GGGGTCCCGG CGCGTTTTAG CGGCTCTGGA TCCGGCACGG 60

ATTTTAC 67

(2) INFORMATION FOR SEQ ID NO: 87:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GATAATACAC CGCAAAGTCT TCAGGTTCCA GGCTGCTAAT GGTCAGGGTA AAATCCGTGC 60

CGGATC 66

(2) INFORMATION FOR SEQ ID NO: 88:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 88:
GAATGCATAC GCTGATATCG TGATGACCCA GAGCCCGGAT AGCCTGGCG 49

(2) INFORMATION FOR SEQ ID NO: 89:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 89:
GCTTCTGCAG TTAATGGTCG CACGTTCGCC CAGGCTCACC GCCAGGCTAT CCGGGC 56

(2) INFORMATION FOR SEQ ID NO: 90:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CGACCATTAA CTGCAGAAGC AGCCAGAGCG TGCTGTATAG CAGCAACAAC AAAAACTATC 60

TGGCGTGGTA CCAG 74

(2) INFORMATION FOR SEQ ID NO: 91:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 91:

GATGCCCAAT AAATTAATAG TTTCGGCGGC TGACCTGGTT TCTGCTGGTA CCACGCCAGA 60

TAG 63

(2) INFORMATION FOR SEQ ID NO: 92:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

AAACTATTAA TTTATTGGGC ATCCACCCGT GAAAGCGGGG TCCCGGATCG TTTTAGCGGC 60

TCTGGATCCG GCAC 74

(2) INFORMATION FOR SEQ ID NO: 93:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GATAATACAC CGCCACGTCT TCAGCTTGCA GGGACGAAAT GGTCAGGGTA AAATCAGTGC 60

CGGATCCAGA GCC 73

(2) INFORMATION FOR SEQ ID NO: 94:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 94:
GAATGCATAC GCTCAGAGCG TGCTGACCCA GCCGCCTTCA GTGAGTGG 48

(2) INFORMATION FOR SEQ ID NO: 95:

(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 = "synthetic oligonucleotide"

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

CAATGTTGCT GCTGCTGCCG CTACACGAGA TGGTCACACG CTGACCTGGT GCGCCACTCA 60

CTGAAGGCGG C 71

(2) INFORMATION FOR SEQ ID NO: 96:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 96:
GGCAGCAGCA GCAACATTGG CAGCAACTAT GTGAGCTGGT ACCAGCAGTT GCCCGGGAC 59

(2) INFORMATION FOR SEQ ID NO: 97:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CCGGCACGCC TGAGGGACGC TGGTTGTTAT CATAAATCAG CAGTTTCGGC GCCGTCCCGG 60

GCAACTGC 68

(2) INFORMATION FOR SEQ ID NO: 98:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 98:
CCCTCAGGCG TGCCGGATCG TTTTAGCGGA TCCAAAAGCG GCACCAGCGC GAGCCTTGCG 60

(2) INFORMATION FOR SEQ ID NO: 99:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 99:
CCGCTTCGTC TTCGCTTTGC AGGCCCGTAA TCGCAAGGCT CGCGCTGG 48

(2) INFORMATION FOR SEQ ID NO: 100:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 100:
GAATGCATAC GCTCAGAGCG CACTGACCCA GCCAGCTTCA GTGAGCGGC 49

(2) INFORMATION FOR SEQ ID NO: 101:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CGCTGCTAGT ACCCGTACAC GAGATGGTAA TGCTCTGACC TGGTGAGCCG CTCACTGAAG 60

CTGG 64

(2) INFORMATION FOR SEQ ID NO: 102:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GTACGGGTAC TAGCAGCGAT GTGGGCGGCT ATAACTATGT GAGCTGGTAC CAGCAGCATC 60

CCGG 64

(2) INFORMATION FOR SEQ ID NO: 103:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CGCCTGAGGG ACGGTTGCTC ACATCATAAA TCATCAGTTT CGGCGCCTTC CCGGGATGCT 60

GCTGGTAC 68

(2) INFORMATION FOR SEQ ID NO: 104:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 104:
CAACCGTCCC TCAGGCGTGA GCAACCGTTT TAGCGGATCC AAAAGCGGCA ACACCGCGAG 60

(2) INFORMATION FOR SEQ ID NO: 105:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 105:
CCGCTTCGTC TTCCGCTTGC AGGCCGCTAA TGGTCAGGCT CGCGGTGTTG CCG 53

(2) INFORMATION FOR SEQ ID NO: 106:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 106:
GAATGCATAC GCTAGCTATG AACTGACCCA GCCGCCTTCA GTGAGCG 4 7

(2) INFORMATION FOR SEQ ID NO: 107:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CGCCCAGCGC ATCGCCGCTA CACGAGATAC GCGCGGTCTG ACCTGGTGCA ACGCTCACTG 60

AAGGCGGC 68

(2) INFORMATION FOR SEQ ID NO: 108:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 108:
GGCGATGCGC TGGGCGATAA ATACGCGAGC TGGTACCAGC AGAAACCCGG GCAGGCGC 58

(2) INFORMATION FOR SEQ ID NO: 109:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGTTCCGGG ATGCCTGAGG GACGGTCAGA ATCATCATAA ATCACCAGAA CTGGCGCCTG 60

CCCGGGTTTC 70

(2) INFORMATION FOR SEQ ID NO: 110:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CAGGCATCCC GGAACGCTTT AGCGGATCCA ACAGCGGCAA CACCGCGACC CTGACCATTA 60

GCGG 64

(2) INFORMATION FOR SEQ ID NO: 111:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 111:
CCGCTTCGTC TTCCGCCTGA GTGCCGCTAA TGGTCAGGGT C 41

(2) INFORMATION FOR SEQ ID NO: 112:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 112:
GCTCTTCACC CCTGTTACCA AAGCCCAGGT GCAATTG 37

(2) INFORMATION FOR SEQ ID NO: 113:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 113:

GGCTTTGCAG CTCACTTTCA CGCTGCTGCC CGGTTTTTTC ACTTCCGCGC CAGACTGAAC 60

CAATTGCACC TGGGCTTTG 79

(2) INFORMATION FOR SEQ ID NO: 114:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GAAAGTGAGC TGCAAAGCCT CCGGAGGCAC TTTTAGCAGC TATGCGATTA GCTGGGTGCG 60

CCAAGCCCCT GGGCAGGGTC 80

(2) INFORMATION FOR SEQ ID NO: 115:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 115:

GCCCTGAAAC TTCTGCGCGT AGTTCGCCGT GCCAAAAATC GGAATAATGC CGCCCATCCA 60

CTCGAGACCC TGCCCAGGGG C 81

(2) INFORMATION FOR SEQ ID NO: 116:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGCAGAAGT TTCAGGGCCG GGTGACCATT ACCGCGGATG AAAGCACCAG CACCGCGTAT 60

ATGGAACTGA GCAGCCTGCG 80

(2) INFORMATION FOR SEQ ID NO: 117:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGCGCAATA ATACACGGCC GTATCTTCGC TACGCAGGCT GCTCAGTTCC 50

(2) INFORMATION FOR SEQ ID NO: 118:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GGCTTTGCAG CTCACTTTCA CGCTCGCGCC CGGTTTTTTC ACTTCCGCGC CGCTCTGAAC 60

CAATTGCACC TGGGCTTTG 79

(2) INFORMATION FOR SEQ ID NO: 119:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GAAAGTGAGC TGCAAAGCCT CCGGATATAC CTTTACCAGC TATTATATGC ACTGGGTCCG 60

CCAAGCCCCT GGGCAGGGTC 80

(2) INFORMATION FOR SEQ ID NO: 120:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCCCTGAAAC TTCTGCGCGT AGTTCGTGCC GCCGCTATTC GGGTTAATCC AGCCCATCCA 60

CTCGAGACCC TGCCCAGGGG C 81

(2) INFORMATION FOR SEQ ID NO: 121:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGCAGAAGT TTCAGGGCCG GGTGACCATG ACCCGTGATA CCAGCATTAG CACCGCGTAT 60

ATGGAACTGA GCAGCCTGCG 80

(2) INFORMATION FOR SEQ ID NO: 122:

(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 = "synthetic oligonucleotide"

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

GGTACAGGTC AGGGTCAGGG TTTGGGTCGG TTTCACCAGG GCCGGGCCGC TTTCTTTCAA 60

TTGCACCTGG GCTTTG 76

(2) INFORMATION FOR SEQ ID NO: 123:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 85 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 123:

CTGACCCTGA CCTGTACCTT TTCCGGATTT AGCCTGTCCA CGTCTGGCGT TGGCGTGGGC 60

TGGATTCGCC AGCCGCCTGG GAAAG 85

(2) INFORMATION FOR SEQ ID NO: 124:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGTTTTCAG GCTGGTGCTA TAATACTTAT CATCATCCCA ATCAATCAGA GCCAGCCACT 60

CGAGGGCTTT CCCAGGCGGC TGG 83

(2) INFORMATION FOR SEQ ID NO: 125:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCACCAGCCT GAAAACGCGT CTGACCATTA GCAAAGATAC TTCGAAAAAT CAGGTGGTGC 60

TGACTATGAC CAACATGG 78

(2) INFORMATION FOR SEQ ID NO: 126:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 126:
GCGCGCAATA ATAGGTGGCC GTATCCACCG GGTCCATGTT GGTCATAGTC AGC 53

(2) INFORMATION FOR SEQ ID NO: 127:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 51 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 127:
CGAAGTGCAA TTGGTGGAAA GCGGCGGCGG CCTGGTGCAA CCGGGCGGCA G 51

(2) INFORMATION FOR SEQ ID NO: 128:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CATAGCTGCT AAAGGTAAAT CCGGAGGCCG CGCAGCTCAG ACGCAGGCTG CCGCCCGGTT 60

GCAC 64

(2) INFORMATION FOR SEQ ID NO: 129:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GATTTACCTT TAGCAGCTAT GCGATGAGCT GGGTGCGCCA AGCCCCTGGG AAGGGTCTCG 60

AGTGGGTGAG 70

(2) INFORMATION FOR SEQ ID NO: 130:

(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 = "synthetic oligonucleotide"

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

GGCCTTTCAC GCTATCCGCA TAATAGGTGC TGCCGCCGCT ACCGCTAATC GCGCTCACCC 60

ACTCGAGACC C 71

(2) INFORMATION FOR SEQ ID NO: 131:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CGGATAGCGT GAAAGGCCGT TTTACCATTT CACGTGATAA TTCGAAAAAC ACCCTGTATC 60

TGCAAATGAA CAG 73

(2) INFORMATION FOR SEQ ID NO: 132:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 62 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CACGCGCGCA ATAATACACG GCCGTATCTT CCGCACGCAG GCTGTTCATT TGCAGATACA 60

GG 62

(2) INFORMATION FOR SEQ ID NO: 133:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GGTCAGGCTC AGGGTTTCGC TCGGTTTCAC CAGGCCCGGA CCACTTTCTT GCAATTGCAC 60

CTGGGCTTTG 70

(2) INFORMATION FOR SEQ ID NO: 134:

(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 = "synthetic oligonucleotide"

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

GAAACCCTGA GCCTGACCTG CACCGTTTCC GGAGGCAGCA TTAGCAGCTA TTATTGGAGC 60

TGGATTCGCC AGCCGC 76

(2) INFORMATION FOR SEQ ID NO: 135:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GATTATAGTT GGTGCTGCCG CTATAATAAA TATAGCCAAT CCACTCGAGA CCCTTCCCAG 60

GCGGCTGGCG AATCCAG 77

(2) INFORMATION FOR SEQ ID NO: 136:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc - "synthetic oligonucleotide"

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

CGGCAGCACC AACTATAATC CGAGCCTGAA AAGCCGGGTG ACCATTAGCG TTGATACTTC 60

GAAAAACCAG TTTAGCCTG 79

(2) INFORMATION FOR SEQ ID NO: 137:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGCGCAATA ATACACGGCC GTATCCGCCG CCGTCACGCT GCTCAGTTTC AGGCTAAACT 60

GGTTTTTCG 69

(2) INFORMATION FOR SEQ ID NO: 138:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 138:
GCTCTTCACC CCTGTTACCA AAGCCGAAGT GCAATTG 37

(2) INFORMATION FOR SEQ ID NO: 139:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

```
CCTTTGCAGC TAATTTTCAG GCTTTCGCCC GGTTTTTTCA CTTCCGCGCC GCTCTGAACC      60

AATTGCACTT CGGCTTTGG                                                   79
```

(2) INFORMATION FOR SEQ ID NO: 140:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

```
CCTGAAAATT AGCTGCAAAG GTTCCGGATA TTCCTTTACG AGCTATTGGA TTGGCTGGGT      60

GCGCCAGATG CCTGG                                                       75
```

(2) INFORMATION FOR SEQ ID NO: 141:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 141:

```
CGGAGAATAA CGGGTATCGC TATCGCCCGG ATAAATAATG CCCATCCACT CGAGACCCTT   60

CCCAGGCATC TGGCGCAC                                                78
```

(2) INFORMATION FOR SEQ ID NO: 142:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 77 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 142:

```
CGATACCCGT TATTCTCCGA GCTTTCAGGG CCAGGTGACC ATTAGCGCGG ATAAAAGCAT   60

TAGCACCGCG TATCTTC                                                 77
```

(2) INFORMATION FOR SEQ ID NO: 143:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

```
GCGCGCAATA ATACATGGCC GTATCGCTCG CTTTCAGGCT GCTCCATTGA AGATACGCGG   60

TGCTAATG                                                           68
```

(2) INFORMATION FOR SEQ ID NO: 144:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 81 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 144:

```
GAAATCGCAC AGGTCAGGCT CAGGGTTTGG CTCGGTTTCA CCAGGCCCGG ACCAGACTGT   60
TGCAATTGCA CCTGGGCTTT G                                             81
```

(2) INFORMATION FOR SEQ ID NO: 145:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 79 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

```
GCCTGACCTG TGCGATTTCC GGAGATAGCG TGAGCAGCAA CAGCGCGGCG TGGAACTGGA   60
TTCGCCAGTC TCCTGGGCG                                                79
```

(2) INFORMATION FOR SEQ ID NO: 146:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

```
CACCGCATAA TCGTTATACC ATTTGCTACG ATAATAGGTA CGGCCCAGCC ACTCGAGGCC   60
ACGCCCAGGA GACTGGCG                                                 78
```

(2) INFORMATION FOR SEQ ID NO: 147:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GGTATAACGA TTATGCGGTG AGCGTGAAAA GCCGGATTAC CATCAACCCG GATACTTCGA 60

AAAACCAGTT TAGCCTGC 78

(2) INFORMATION FOR SEQ ID NO: 148:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 68 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GCGCGCAATA ATACACGGCC GTATCTTCCG GGGTCACGCT GTTCAGTTGC AGGCTAAACT 60

GGTTTTTC 68

(2) INFORMATION FOR SEQ ID NO: 149:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 149:

GGCTGAAGAC GTGGGCGTGT ATTATTGCCA GCAGCATTAT ACCACCCCGC CGACCTTTGG 60

CCAGGGTAC 69

(2) INFORMATION FOR SEQ ID NO: 150:

(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 = "synthetic oligonucleotide"

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

GCGGAAAAAT AAACACGCTC GGAGCAGCCA CCGTACGTTT AATTTCAACT TTCGTACCCT 60

GGCCAAAGGT C 71

(2) INFORMATION FOR SEQ ID NO: 151:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GAGCGTGTTT ATTTTTCCGC CGAGCGATGA ACAACTGAAA AGCGGCACGG CGAGCGTGGT 60

GTGCCTGCTG 70

(2) INFORMATION FOR SEQ ID NO: 152:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 152:

CAGCGCGTTG TCTACTTTCC ACTGAACTTT CGCTTCACGC GGATAAAAGT TGTTCAGCAG 60

GCACACCACG C 71

(2) INFORMATION FOR SEQ ID NO: 153:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GAAAGTAGAC AACGCGCTGC AAAGCGGCAA CAGCCAGGAA AGCGTGACCG AACAGGATAG 60

CAAAGATAG 69

(2) INFORMATION FOR SEQ ID NO: 154:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 154:

GTTTTTCATA ATCCGCTTTG CTCAGGGTCA GGGTGCTGCT CAGAGAATAG GTGCTATCTT 60

TGCTATCCTG TTCG 74

(2) INFORMATION FOR SEQ ID NO: 155:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 155:

GCAAAGCGGA TTATGAAAAA CATAAAGTGT ATGCGTGCGA AGTGACCCAT CAAGGTCTGA 60

GCAGCCCGGT G 71

(2) INFORMATION FOR SEQ ID NO: 156:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 156:
GGCATGCTTA TCAGGCCTCG CCACGATTAA AAGATTTAGT CACCGGGCTG CTCAGAC 57

(2) INFORMATION FOR SEQ ID NO: 157:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 157:
GGCGTCTAGA GGCCAAGGCA CCCTGGTGAC GGTTAGCTCA GCGTCGAC 48

(2) INFORMATION FOR SEQ ID NO: 158:

(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 = "synthetic oligonucleotide"

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

GTGCTTTTGC TGCTCGGAGC CAGCGGAAAC ACGCTTGGAC CTTTGGTCGA CGCTGAGCTA 60

ACC 63

(2) INFORMATION FOR SEQ ID NO: 159:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 66 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

CTCCGAGCAG CAAAAGCACC AGCGGCGGCA CGGCTGCCCT GGGCTGCCTG GTTAAAGATT 60

ATTTCC 66

(2) INFORMATION FOR SEQ ID NO: 160:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 160:

CTGGTCAGCG CCCCGCTGTT CCAGCTCACG GTGACTGGTT CCGGGAAATA ATCTTTAACC 60

AGGCA 65

(2) INFORMATION FOR SEQ ID NO: 161:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 60 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 161:
AGCGGGGCGC TGACCAGCGG CGTGCATACC TTTCCGGCGG TGCTGCAAAG CAGCGGCCTG 60

(2) INFORMATION FOR SEQ ID NO: 162:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GTGCCTAAGC TGCTGCTCGG CACGGTCACA ACGCTGCTCA GGCTATACAG GCCGCTGCTT 60

TGCAG 65

(2) INFORMATION FOR SEQ ID NO: 163:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 61 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

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

GAGCAGCAGC TTAGGCACTC AGACCTATAT TTGCAACGTG AACCATAAAC CGAGCAACAC 60

C 61

(2) INFORMATION FOR SEQ ID NO: 164:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 164:
GCGCGAATTC GCTTTTCGGT TCCACTTTTT TATCCACTTT GGTGTTGCTC GGTTTATGG 59

(2) INFORMATION FOR SEQ ID NO: 165:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 333 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:7..321
(D) OTHER INFORMATION:/product= "C kappa"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 165:

```
CGTACG GTG GCT GCT CCG AGC GTG TTT ATT TTT CCG CCG AGC GAT GAA          48
       Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu
           125                 130                 135

CAA CTG AAA AGC GGC ACG GCG AGC GTG GTG TGC CTG CTG AAC AAC TTT          96
Gln Leu Lys Ser Gly Thr Ala Ser Val Val Cys Leu Leu Asn Asn Phe
        140                 145                 150

TAT CCG CGT GAA GCG AAA GTT CAG TGG AAA GTA GAC AAC GCG CTG CAA         144
Tyr Pro Arg Glu Ala Lys Val Gln Trp Lys Val Asp Asn Ala Leu Gln
    155                 160                 165

AGC GGC AAC AGC CAG GAA AGC GTG ACC GAA CAG GAT AGC AAA GAT AGC         192
Ser Gly Asn Ser Gln Glu Ser Val Thr Glu Gln Asp Ser Lys Asp Ser
170                 175                 180                 185

ACC TAT TCT CTG AGC AGC ACC CTG ACC CTG AGC AAA GCG GAT TAT GAA         240
Thr Tyr Ser Leu Ser Ser Thr Leu Thr Leu Ser Lys Ala Asp Tyr Glu
                190                 195                 200

AAA CAT AAA GTG TAT GCG TGC GAA GTG ACC CAT CAA GGT CTG AGC AGC         288
Lys His Lys Val Tyr Ala Cys Glu Val Thr His Gln Gly Leu Ser Ser
            205                 210                 215

CCG GTG ACT AAA TCT TTT AAT CGT GGC GAG GCC TGATAAGCAT GC               333
Pro Val Thr Lys Ser Phe Asn Arg Gly Glu Ala
        220                 225
```

(2) INFORMATION FOR SEQ ID NO: 166:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 166:

```
Val Ala Ala Pro Ser Val Phe Ile Phe Pro Pro Ser Asp Glu Gln Leu
  1               5                  10                  15

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

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

         35                  40                  45

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

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

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

Thr Lys Ser Phe Asn Arg Gly Glu Ala
            100                 105
```

(2) INFORMATION FOR SEQ ID NO: 167:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 327 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:6..317
(D) OTHER INFORMATION:/product= "CH1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 167:

```
GCTCA GCG TCG ACC AAA GGT CCA AGC GTG TTT CCG CTG GCT CCG AGC          47
      Ala Ser Thr Lys Gly Pro Ser Val Phe Pro Leu Ala Pro Ser
                    110                 115

AGC AAA AGC ACC AGC GGC GGC ACG GCT GCC CTG GGC TGC CTG GTT AAA        95
Ser Lys Ser Thr Ser Gly Gly Thr Ala Ala Leu Gly Cys Leu Val Lys
120                 125                 130                 135

GAT TAT TTC CCG GAA CCA GTC ACC GTG AGC TGG AAC AGC GGG GCG CTG        143
Asp Tyr Phe Pro Glu Pro Val Thr Val Ser Trp Asn Ser Gly Ala Leu
                140                 145                 150

ACC AGC GGC GTG CAT ACC TTT CCG GCG GTG CTG CAA AGC AGC GGC CTG        191
Thr Ser Gly Val His Thr Phe Pro Ala Val Leu Gln Ser Ser Gly Leu
            155                 160                 165

TAT AGC CTG AGC AGC GTT GTG ACC GTG CCG AGC AGC AGC TTA GGC ACT        239
Tyr Ser Leu Ser Ser Val Val Thr Val Pro Ser Ser Ser Leu Gly Thr
        170                 175                 180

CAG ACC TAT ATT TGC AAC GTG AAC CAT AAA CCG AGC AAC ACC AAA GTG        287
Gln Thr Tyr Ile Cys Asn Val Asn His Lys Pro Ser Asn Thr Lys Val
    185                 190                 195

GAT AAA AAA GTG GAA CCG AAA AGC GAA TTC TGATAAGCTT                    327
Asp Lys Lys Val Glu Pro Lys Ser Glu Phe

200                 205
```

(2) INFORMATION FOR SEQ ID NO: 168:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 168:

```
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 Glu Phe
            100
```

(2) INFORMATION FOR SEQ ID NO: 169:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 408 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:85..396
(D) OTHER INFORMATION:/product- "C lambda"

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

```
GAAGACGAAG CGGATTATTA TTGCCAGCAG CATTATACCA CCCCGCCTGT GTTTGGCGGC      60

GGCACGAAGT TAACCGTTCT TGGC CAG CCG AAA GCC GCA CCG AGT GTG ACG        111
                             Gln Pro Lys Ala Ala Pro Ser Val Thr
                             105                 110
```

```
CTG TTT CCG CCG AGC AGC GAA GAA TTG CAG GCG AAC AAA GCG ACC CTG        159
Leu Phe Pro Pro Ser Ser Glu Glu Leu Gln Ala Asn Lys Ala Thr Leu
    115                 120                 125

GTG TGC CTG ATT AGC GAC TTT TAT CCG GGA GCC GTG ACA GTG GCC TGG        207
Val Cys Leu Ile Ser Asp Phe Tyr Pro Gly Ala Val Thr Val Ala Trp
130                 135                 140                 145

AAG GCA GAT AGC AGC CCC GTC AAG GCG GGA GTG GAG ACC ACC ACA CCC        255
Lys Ala Asp Ser Ser Pro Val Lys Ala Gly Val Glu Thr Thr Thr Pro
                150                 155                 160

TCC AAA CAA AGC AAC AAC AAG TAC GCG GCC AGC AGC TAT CTG AGC CTG        303
Ser Lys Gln Ser Asn Asn Lys Tyr Ala Ala Ser Ser Tyr Leu Ser Leu
            165                 170                 175

ACG CCT GAG CAG TGG AAG TCC CAC AGA AGC TAC AGC TGC CAG GTC ACG        351
Thr Pro Glu Gln Trp Lys Ser His Arg Ser Tyr Ser Cys Gln Val Thr
        180                 185                 190

CAT GAG GGG AGC ACC GTG GAA AAA ACC GTT GCG CCG ACT GAG GCC            396
His Glu Gly Ser Thr Val Glu Lys Thr Val Ala Pro Thr Glu Ala
    195                 200                 205

TGATAAGCAT GC                                                          408
```

(2) INFORMATION FOR SEQ ID NO: 170:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 104 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 170:

```
Gln Pro Lys Ala Ala Pro Ser Val Thr Leu Phe Pro Pro Ser Ser Glu
  1               5                  10                  15

Glu Leu Gln Ala Asn Lys Ala Thr Leu Val Cys Leu Ile Ser Asp Phe
             20                  25                  30

Tyr Pro Gly Ala Val Thr Val Ala Trp Lys Ala Asp Ser Ser Pro Val
         35                  40                  45

Lys Ala Gly Val Glu Thr Thr Thr Pro Ser Lys Gln Ser Asn Asn Lys
     50                  55                  60

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

His Arg Ser Tyr Ser Cys Gln Val Thr His Glu Gly Ser Thr Val Glu
                 85                  90                  95

Lys Thr Val Ala Pro Thr Glu Ala
            100
```

(2) INFORMATION FOR SEQ ID NO: 171:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 78 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 171:

```
GAAGACAAGC GGATTATTAT TGCCAGCAGC ATTATACCAC CCCGCCTGTG TTTGGCGGCG      60

GCACGAAGTT AACCGTTC                                                    78
```

(2) INFORMATION FOR SEQ ID NO: 172:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 172:

```
CAATTCTTCG CTGCTCGGCG GAAACAGCGT CACACTCGGT GCGGCTTTCG GCTGGCCAAG      60

AACGGTTAAC TTCGTGCCGC                                                  80
```

(2) INFORMATION FOR SEQ ID NO: 173:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 173:

```
CGCCGAGCAG CGAAGAATTG CAGGCGAACA AAGCGACCCT GGTGTGCCTG ATTAGCGACT      60

TTTATCCGGG AGCCGTGACA                                                  80
```

(2) INFORMATION FOR SEQ ID NO: 174:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 174:

```
TGTTTGGAGG GTGTGGTGGT CTCCACTCCC GCCTTGACGG GGCTGCTATC TGCCTTCCAG      60

GCCACTGTCA CGGCTCCCGG                                                  80
```

(2) INFORMATION FOR SEQ ID NO: 175:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 94 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 175:

```
CCACACCCTC CAAACAAAGC AACAACAAGT ACGCGGCCAG CAGCTATCTG AGCCTGACGC      60

CTGAGCAGTG GAAGTCCCAC AGAAGCTACA GCTG                                  94
```

(2) INFORMATION FOR SEQ ID NO: 176:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 176:

GCATGCTTAT CAGGCCTCAG TCGGCGCAAC GGTTTTTTCC ACGGTGCTCC CCTCATGCGT 60

GACCTGGCAG CTGTAGCTTC 80

(2) INFORMATION FOR SEQ ID NO: 177:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 843 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..843
(D) OTHER INFORMATION:/product= "VH3-Vk2"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 177:

```
ATG AAA CAA AGC ACT ATT GCA CTG GCA CTC TTA CCG TTG CTC TTC ACC        48
Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr
105                 110                 115                 120

CCT GTT ACC AAA GCC GAC TAC AAA GAT GAA GTG CAA TTG GTG GAA AGC        96
Pro Val Thr Lys Ala Asp Tyr Lys Asp Glu Val Gln Leu Val Glu Ser
                125                 130                 135

GGC GGC GGC CTG GTG CAA CCG GGC GGC AGC CTG CGT CTG AGC TGC GCG       144
Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala
            140                 145                 150
GCC TCC GGA TTT ACC TTT AGC AGC TAT GCG ATG AGC TGG GTG CGC CAA       192
Ala Ser Gly Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln
        155                 160                 165

GCC CCT GGG AAG GGT CTC GAG TGG GTG AGC GCG ATT AGC GGT AGC GGC       240
Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Ala Ile Ser Gly Ser Gly
    170                 175                 180

GGC AGC ACC TAT TAT GCG GAT AGC GTG AAA GGC CGT TTT ACC ATT TCA       288
Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
185                 190                 195                 200

CGT GAT AAT TCG AAA AAC ACC CTG TAT CTG CAA ATG AAC AGC CTG CGT       336
Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
                205                 210                 215

GCG GAA GAT ACG GCC GTG TAT TAT TGC GCG CGT TGG GGC GGC GAT GGC       384
Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Trp Gly Gly Asp Gly
            220                 225                 230

TTT TAT GCG ATG GAT TAT TGG GGC CAA GGC ACC CTG GTG ACG GTT AGC       432
Phe Tyr Ala Met Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser
        235                 240                 245

TCA GCG GGT GGC GGT TCT GGC GGC GGT GGG AGC GGT GGC GGT GGT TCT       480
Ser Ala Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
    250                 255                 260

GGC GGT GGT GGT TCC GAT ATC GTG ATG ACC CAG AGC CCA CTG AGC CTG       528
Gly Gly Gly Gly Ser Asp Ile Val Met Thr Gln Ser Pro Leu Ser Leu
265                 270                 275                 280
```

```
CCA GTG ACT CCG GGC GAG CCT GCG AGC ATT AGC TGC AGA AGC AGC CAA       576
Pro Val Thr Pro Gly Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln
                285                 290                 295

AGC CTG CTG CAT AGC AAC GGC TAT AAC TAT CTG GAT TGG TAC CTT CAA       624
Ser Leu Leu His Ser Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln
            300                 305                 310

AAA CCA GGT CAA AGC CCG CAG CTA TTA ATT TAT CTG GGC AGC AAC CGT       672
Lys Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg
        315                 320                 325

GCC AGT GGG GTC CCG GAT CGT TTT AGC GGC TCT GGA TCC GGC ACC GAT       720
Ala Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
    330                 335                 340

TTT ACC CTG AAA ATT AGC CGT GTG GAA GCT GAA GAC GTG GGC GTG TAT       768
Phe Thr Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr
345                 350                 355                 360

TAT TGC CAG CAG CAT TAT ACC ACC CCG CCG ACC TTT GGC CAG GGT ACG       816
Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr
                365                 370                 375

AAA GTT GAA ATT AAA CGT ACG GAA TTC                                  843
Lys Val Glu Ile Lys Arg Thr Glu Phe
            380                 385
```

(2) INFORMATION FOR SEQ ID NO: 178:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 281 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 178:

```
Met Lys Gln Ser Thr Ile Ala Leu Ala Leu Leu Pro Leu Leu Phe Thr
 1               5                  10                  15

Pro Val Thr Lys Ala Asp Tyr Lys Asp Glu Val Gln Leu Val Glu Ser
            20                  25                  30

Gly Gly Gly Leu Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala
        35                  40                  45

Ala Ser Gly Phe Thr Phe Ser Ser Tyr Ala Met Ser Trp Val Arg Gln
    50                  55                  60

Ala Pro Gly Lys Gly Leu Glu Trp Val Ser Ala Ile Ser Gly Ser Gly
65                  70                  75                  80

Gly Ser Thr Tyr Tyr Ala Asp Ser Val Lys Gly Arg Phe Thr Ile Ser
                85                  90                  95

Arg Asp Asn Ser Lys Asn Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg
            100                 105                 110

Ala Glu Asp Thr Ala Val Tyr Tyr Cys Ala Arg Trp Gly Gly Asp Gly
        115                 120                 125

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

Ser Ala Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser
145                 150                 155                 160

Gly Gly Gly Gly Ser Asp Ile Val Met Thr Gln Ser Pro Leu Ser Leu
                165                 170                 175

Pro Val Thr Pro Gly Glu Pro Ala Ser Ile Ser Cys Arg Ser Ser Gln
            180                 185                 190

Ser Leu Leu His Ser Asn Gly Tyr Asn Tyr Leu Asp Trp Tyr Leu Gln
        195                 200                 205

Lys Pro Gly Gln Ser Pro Gln Leu Leu Ile Tyr Leu Gly Ser Asn Arg
    210                 215                 220

Ala Ser Gly Val Pro Asp Arg Phe Ser Gly Ser Gly Ser Gly Thr Asp
225                 230                 235                 240

Phe Thr Leu Lys Ile Ser Arg Val Glu Ala Glu Asp Val Gly Val Tyr
                245                 250                 255

Tyr Cys Gln Gln His Tyr Thr Thr Pro Pro Thr Phe Gly Gln Gly Thr
            260                 265                 270

Lys Val Glu Ile Lys Arg Thr Glu Phe
        275                 280
```

(2) INFORMATION FOR SEQ ID NO: 179:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15 amino acids

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 179:

```
Cys Ala Arg Trp Gly Gly Asp Gly Phe Tyr Ala Met Asp Tyr Trp
1               5                  10                  15
```

(2) INFORMATION FOR SEQ ID NO: 180:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 180:

```
Cys Ala Arg Phe Gly Lys Met Asn Tyr Asp Tyr Trp
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 181:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 181:

```
Cys Ala Arg His Arg Thr Glu Trp His Asp Tyr Trp
1               5                  10
```

(2) INFORMATION FOR SEQ ID NO: 182:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 182:

```
Cys Ala Arg Val Arg Glu Leu Tyr His Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 183:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 183:

```
Cys Ala Arg Lys Phe Leu Lys Ala Arg Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 184:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 184:

```
Cys Ala Arg Trp Asn Thr Thr Gly Tyr Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 185:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ 10 NO: 185:

```
Cys Ala Arg Ile Asn Glu Ala Gln Pro Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 186:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 186:

```
Cys Ala Arg Thr Ala Ile Thr Arg Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 187:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 187:

```
Cys Ala Arg Trp Tyr Asn Arg Asn Ser Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 188:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 188:

Cys Ala Arg Ser Val Gly Asp Ser Lys Asp Tyr Trp
1 5 10

(2) INFORMATION FOR SEQ ID NO: 189:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 189:

Cys Ala Arg Ser Lys Thr Phe Ala Ala Asp Tyr Trp
1 5 10

(2) INFORMATION FOR SEQ ID NO: 190:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

Cys Ala Arg Val Ala Pro Gln Tyr Asp Asp Tyr Trp
1 5 10

(2) INFORMATION FOR SEQ ID NO: 191:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 191:

Cys Ala Arg Met Gln Ser Glu Trp Met Asp Tyr Trp
1 5 10

(2) INFORMATION FOR SEQ ID NO: 192:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 192:

```
Cys Ala Arg Tyr Phe Val His Phe Leu Tyr Thr Met Val Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 193:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 193:

```
Cys Ala Arg Met Ala Leu Arg Ala Ser Gly Lys Tyr Ile Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 194:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 194:

```
Cys Ala Arg Lys Asn Gln Met Val Phe His Ala Arg Lys Phe Asp Val

Trp
```

(2) INFORMATION FOR SEQ ID NO: 195:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 195:

```
Cys Ala Arg Thr Gln Ser Phe Trp Glu Gln Gln Lys Val Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 196:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 196:

```
Cys Ala Arg Tyr Pro Tyr Arg Ser Asn Phe Phe Met Pro Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 197:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION:3..4
(D) OTHER INFORMATION:/product= "see Figure 10C"
/label= R*G
/note= "* denotes codon with one-base deletion, causes shift of reading fr..."

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 197:

Cys Ala Arg Gly Ser Gly Ser Glu His Trp Ser Ile Phe Asp Val Trp
1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 198:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 198:

Cys Ala Arg Arg Asn Pro Trp Asn Val Asn Tyr Leu His Phe Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 199:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 199:

Cys Ala Arg Met Lys Pro Met Leu Asn Arg Asp Gly Thr Met Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 200:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 200:

```
Cys Ala Arg Lys Gly Ser Glu Phe Leu Glu Thr Asp Val Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 201:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 201:

```
Cys Ala Arg Ser Trp Thr Asn Asp Lys Pro Asn Phe Ile Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 202:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 202:

```
Cys Ala Arg Tyr Ala Gly Thr Thr Phe Lys Gln Gly Pro Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 203:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 203:

```
Cys Ala Arg Lys Arg Met Met Gln Asn Pro Arg Phe Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 204:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 204:

```
Cys Ala Arg Arg Ser Lys Gln Lys Arg Lys Met Arg Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 205:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 205:

```
Cys Ala Arg Arg Asn Gly Lys Arg His Leu Arg His Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 206:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 206:

```
Cys Ala Arg Arg Lys Met Arg Lys Arg Ile Lys Arg Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 207:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 207:

```
Cys Ala Arg Tyr Arg Lys Ile Met Lys Trp Lys Asn Ser Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 208:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 208:

```
Cys Ala Arg Leu Ile Glu Val His Pro Ser Phe Asp Gln Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 209:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 209:

```
Cys Ala Arg Arg Lys Pro Met Phe Leu Lys Lys Ala Val Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 210:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 210:

```
Cys Ala Arg Arg Lys Phe His Arg Tyr Ser Thr Val Lys Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 211:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 211:

```
Cys Ala Arg Arg Lys Thr Met Arg Ser Arg Val Lys Tyr Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 212:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 212:

```
Cys Ala Arg Lys Lys Arg Ser Trp Arg Arg Met Asp Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 213:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 213:

```
Cys Ala Arg Arg Asn Pro Arg Arg Gly Arg Met Asn Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 214:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 214:

```
Cys Ala Arg Lys Gly Lys Lys Lys Phe Ala Arg Pro Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 215:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 215:

```
Cys Ala Arg Arg Met Val His Lys Gly Lys Arg Lys Ile Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 216:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 216:

```
Cys Ala Arg Arg Lys His Ile Thr Tyr Pro Arg Lys Gln Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 217:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 217:

```
Cys Ala Arg Arg Trp Thr Lys Arg Arg Ser Phe Ala Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 218:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 218:

```
Cys Ala Arg Lys Lys Leu Lys Gln Tyr Thr Phe Ser Arg Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 219:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 219:

```
Cys Ala Arg Thr Arg Pro Trp Gln Ala Thr Arg Lys Gly Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 220:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 220:

```
Cys Ala Arg Asn Gln Trp Glu Phe Lys Asn Arg Arg Lys Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO:221:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 221:

```
Cys Ala Arg Lys Arg Trp Met Trp Pro Ile Gly Lys Arg Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 222:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 222:

```
Cys Ala Arg Tyr Ser Leu Trp Arg Leu Asp Glu Tyr Phe Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 223:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 223:

```
Cys Ala Arg Val Pro Trp Gly Asp Phe Trp Ser Trp His Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 224:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 224:

```
Cys Ala Arg Asn Gly Leu Glu Pro Arg His Arg Lys Met Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 225:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 225:

```
Cys Ala Arg Ile Met Lys Ala Pro Pro Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 226:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 226:

```
Cys Ala Arg Arg Lys Thr Trp His Trp Phe Tyr Lys Arg Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 227:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 227:

```
Cys Ala Arg Trp Lys Asp Met Trp Ser Gln Val Tyr Val Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 228:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 228:

```
Cys Ala Arg Asn Lys Gln Gln Met Arg Phe Arg Arg Phe Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 229:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 229:

```
Cys Ala Arg Asn Met Leu Ala Leu Ser Arg Gly Lys Glu Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 230:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 230:

```
Cys Ala Arg Asn Met Arg Leu Met Arg Met Arg Lys Asn Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 231:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 231:

```
Cys Ala Arg Tyr Ile Lys Gln Ala Lys Arg Lys Leu Ala Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 232:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 232:

```
Cys Ala Arg Tyr Asn Arg His Ala Trp Gln Lys Met Gln Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 233:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 233:

Cys Ala Arg Tyr Val Lys Tyr Ala Arg Asn Lys Met Gln Phe Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 234:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 234:

Cys Ala Arg Tyr Lys Arg Gly Ala Trp Met Lys Thr Met Phe Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 235:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 235:

Cys Ala Arg Arg Lys Pro Leu Arg Arg Ile Met Lys Trp Phe Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 236:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 236:

```
Cys Ala Arg Tyr Arg Lys Arg Ala Ser Arg Gln Met Gln Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 237:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 237:

```
Cys Ala Arg Gln Arg Tyr Arg Ser Lys Ile Lys Gly His Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 238:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 238:

```
Cys Ala Arg Trp Arg Asp Phe Asn Ser Tyr Asp Pro Met Asp Tyr Trp
1               5                   10                  15
```

(2) INFORMATION FOR SEQ ID NO: 239:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 239:

Cys Ala Arg Met Ala Asp Leu Asp Asn Tyr Trp Val Gln Phe Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 240:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 240:

Cys Ala Arg Leu Gln Ala Tyr Leu Lys Pro His His Trp Met Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 241:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 241:

Cys Ala Arg Arg Leu Ile Glu Gln Ala Arg Asp His Val Met Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 242:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 242:

Cys Ala Arg Ser Trp His Asn Ser Gln Phe Thr Gln Ser Phe Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 243:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 243:

Cys Ala Arg Val Asp His Phe Gln Thr Glu Asn Glu Trp Met Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 244:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 244:

Cys Ala Arg Asp Trp Pro Thr Leu Ile Phe Trp Tyr Trp Phe Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 245:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 245:

```
Cys Ala Arg Gly Phe Gly Phe Thr Glu Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 246:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 246:

```
Cys Ala Arg Gln Phe Asp Glu Asp Ser Phe Val Arg Arg Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 247:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 247:

```
Cys Ala Arg Ile Leu Lys Glu Ser Ser Lys Ser Arg Gln Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 248:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 248:

```
Cys Ala Arg Glu Gln Asp Glu Tyr Gly Ala Ile Arg Ile Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 249:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 249:

```
Cys Ala Arg Asn His Phe Glu Ala Ser Trp Pro Arg Arg Gln Met Asp
1               5                   10                  15

Val Trp
```

(2) INFORMATION FOR SEQ ID NO: 250:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 250:

```
Cys Ala Arg Glu Asn Glu Trp Val Asp Met Ile Leu Asp Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 251:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 251:

```
Cys Ala Arg Gln Tyr Ser Glu Thr Arg Trp Val Arg Lys Phe Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 252:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 252:

```
Cys Ala Arg Gln Phe Lys Glu Ser Lys Thr Arg Arg Lys Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 253:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 253:

```
Cys Ala Arg Lys Lys Thr Gln Tyr Val His Asp Trp Arg Met Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 254:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 254:

```
Cys Ala Arg Arg Trp Arg Glu Thr Lys Ser Lys Arg Phe Phe Asp Val
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 255:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 255:

```
Cys Ala Arg Asp Tyr Ile Met Glu Phe Asp Tyr Trp
1               5                   10
```

(2) INFORMATION FOR SEQ ID NO: 256:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 256:

```
Cys Ala Arg Gln Phe Glu Glu Thr Lys Gln Arg Arg Leu Met Asp Tyr
1               5                   10                  15

Trp
```

(2) INFORMATION FOR SEQ ID NO: 257:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 257:

Cys Ala Arg Asp Gln Gly Phe Tyr Ala Ile Asp Tyr Val Met Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 258:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 258:

Cys Ala Arg Val Phe Thr Tyr Met Tyr Asn Tyr Phe Arg Phe Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 259:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 259:

Cys Ala Arg Val Phe Phe Glu Gln Met Glu Val Val Arg Met Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 260:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 260:

Cys Ala Arg Glu Lys Glu Tyr Arg Leu Ser Trp Ser Gln Met Asp Tyr

1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 261:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 261:

Cys Ala Arg Tyr Pro Ser Arg Trp Ala Pro Asn Trp Tyr Met Asp Tyr
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 262:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: protein
(v) FRAGMENT TYPE: internal
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 262:

Cys Ala Arg Asp Gly Gly Phe Lys Pro Leu Thr His Phe Phe Asp Val
1 5 10 15

Trp

(2) INFORMATION FOR SEQ ID NO: 263:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 143 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 263:

```
ACATGTAAGC TTCCCCCCCC CCTTAATTAA CCCCCCCCCC TGTACACCCC CCCCCCGCTA      60
GCCCCCCCCC CCAGATCTCC CCCCCCCCGA CGTCCCCCCT CTAGACCCCC CCCCCGCATG     120
CCCCCCCCCC CGAATTCGAC GTC                                             143
```

(2) INFORMATION FOR SEQ ID NO: 264:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1947 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic vector"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:132..989
(D) OTHER INFORMATION:/product= "Amp resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 264:

```
CAGGTGGCAC TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC      60

ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA     120

AAAGGAAGAG T ATG AGT ATT CAA CAT TTC CGT GTC GCC CTT ATT CCC TTT       170
             Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe
                         285                 290

TTT GCG GCA TTT TGC CTT CCT GTT TTT GCT CAC CCA GAA ACG CTG GTG       218
Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val
295                 300                 305                 310

AAA GTA AAA GAT GCT GAA GAT CAG TTG GGT GCA CGA GTG GGT TAC ATC       266
Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile
                315                 320                 325

GAA CTG GAT CTC AAC AGC GGT AAG ATC CTT GAG AGT TTT CGC CCC GAA       314
Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu
            330                 335                 340

GAA CGT TTT CCA ATG ATG AGC ACT TTT AAA GTT CTG CTA TGT GGC GCG       362
Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala
        345                 350                 355

GTA TTA TCC CGT ATT GAC GCC GGG CAA GAG CAA CTC GGT CGC CGC ATA       410
Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile
    360                 365                 370
```

```
CAC TAT TCT CAG AAT GAC TTG GTT GAG TAC TCA CCA GTC ACA GAA AAG        458
His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys
375                 380                 385                 390

CAT CTT ACG GAT GGC ATG ACA GTA AGA GAA TTA TGC AGT GCT GCC ATA        506
His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile
                395                 400                 405

ACC ATG AGT GAT AAC ACT GCG GCC AAC TTA CTT CTG ACA ACG ATC GGA        554
Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly
            410                 415                 420

GGA CCG AAG GAG CTA ACC GCT TTT TTG CAC AAC ATG GGG GAT CAT GTA        602
Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val
        425                 430                 435

ACT CGC CTT GAT CGT TGG GAA CCG GAG CTG AAT GAA GCC ATA CCA AAC        650
Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn
    440                 445                 450

GAC GAG CGT GAC ACC ACG ATG CCT GTA GCA ATG GCA ACA ACG TTG CGC        698
Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg
455                 460                 465                 470

AAA CTA TTA ACT GGC GAA CTA CTT ACT CTA GCT TCC CGG CAA CAA TTA        746
Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu
                475                 480                 485

ATA GAC TGG ATG GAG GCG GAT AAA GTT GCA GGA CCA CTT CTG CGC TCG        794
Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser
            490                 495                 500

GCC CTT CCG GCT GGC TGG TTT ATT GCT GAT AAA TCT GGA GCC GGT GAG        842
Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu
        505                 510                 515

CGT GGG TCT CGC GGT ATC ATT GCA GCA CTG GGG CCA GAT GGT AAG CCC        890
Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro
    520                 525                 530

TCC CGT ATC GTA GTT ATC TAC ACG ACG GGG AGT CAG GCA ACT ATG GAT        938
Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp
535                 540                 545                 550

GAA CGA AAT AGA CAG ATC GCT GAG ATA GGT GCC TCA CTG ATT AAG CAT        986
Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His
                555                 560                 565

TGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG ATTTAAAACT             1039
Trp

TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT     1099

CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC     1159

TTCTTGAGAT CCTTTTTTTC TGCGCGTAAT CTGCTGCTTG CAAACAAAAA AACCACCGCT     1219

ACCAGCGGTG GTTTGTTTGC CGGATCAAGA GCTACCAACT CTTTTTCCGA AGGTAACTGG     1279
```

```
CTTCAGCAGA GCGCAGATAC CAAATACTGT CCTTCTAGTG TAGCCGTAGT TAGGCCACCA    1339

CTTCAAGAAC TCTGTAGCAC CGCCTACATA CCTCGCTCTG CTAATCCTGT TACCAGTGGC    1399

TGCTGCCAGT GGCGATAAGT CGTGTCTTAC CGGGTTGGAC TCAAGACGAT AGTTACCGGA    1459

TAAGGCGCAG CGGTCGGGCT GAACGGGGGG TTCGTGCACA CAGCCCAGCT TGGAGCGAAC    1519

GACCTACACC GAACTGAGAT ACCTACAGCG TGAGCTATGA GAAAGCGCCA CGCTTCCCGA    1579

AGGGAGAAAG GCGGACAGGT ATCCGGTAAG CGGCAGGGTC GGAACAGGAG AGCGCACGAG    1639

GGAGCTTCCA GGGGGAAACG CCTGGTATCT TTATAGTCCT GTCGGGTTTC GCCACCTCTG    1699

ACTTGAGCGT CGATTTTTGT GATGCTCGTC AGGGGGGCGG AGCCTATGGA AAAACGCCAG    1759

CAACGCGGCC TTTTTACGGT TCCTGGCCTT TTGCTGGCCT TTTGCTCACA TGTAAGCTTC    1819

CCCCCCCCCT TAATTAACCC CCCCCCCTGT ACACCCCCCC CCCGCTAGCC CCCCCCCCCA    1879

GATCTCCCCC CCCCCGACGT CCCCCCTCTA GACCCCCCCC CCGCATGCCC CCCCCCCCGA    1939

ATTCACGT                                                            1947
```

(2) INFORMATION FOR SEQ ID NO: 265:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 286 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 265:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
        275                 280                 285
```

(2) INFORMATION FOR SEQ ID NO: 266:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 142 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 266:

```
GACGTCTTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC      60

CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG     120

ACCATGATTA CGAATTTCTA GA                                              142
```

(2) INFORMATION FOR SEQ ID NO: 267:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 520 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..510
(D) OTHER INFORMATION:/product= "gIIIp ss with myc-tag, amber codon"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 267:

```
GAA TTC GAG CAG AAG CTG ATC TCT GAG GAG GAT CTG TAG GGT GGT GGC        48
Glu Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  *  Gly Gly Gly
            290                 295                 300

TCT GGT TCC GGT GAT TTT GAT TAT GAA AAG ATG GCA AAC GCT AAT AAG        96
Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys
        305                 310                 315

GGG GCT ATG ACC GAA AAT GCC GAT GAA AAC GCG CTA CAG TCT GAC GCT       144
Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala
    320                 325                 330

AAA GGC AAA CTT GAT TCT GTC GCT ACT GAT TAC GGT GCT GCT ATC GAT       192
Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp
335                 340                 345                 350

GGT TTC ATT GGT GAC GTT TCC GGC CTT GCT AAT GGT AAT GGT GCT ACT       240
Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr
                355                 360                 365

GGT GAT TTT GCT GGC TCT AAT TCC CAA ATG GCT CAA GTC GGT GAC GGT       288
Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly
            370                 375                 380

Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu
        385                 390                 395

CCT CAA TCG GTT GAA TGT CGC CCT TTT GTC TTT GGC GCT GGT AAA CCA       384
Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe Gly Ala Gly Lys Pro
    400                 405                 410

TAT GAA TTT TCT ATT GAT TGT GAC AAA ATA AAC TTA TTC CGT GGT GTC       432
Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val
415                 420                 425                 430

TTT GCG TTT CTT TTA TAT GTT GCC ACC TTT ATG TAT GTA TTT TCT ACG       480
Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr
                435                 440                 445

TTT GCT AAC ATA CTG CGT AAT AAG GAG TCT TGATAAGCTT                    520
Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
            450                 455
```

(2) INFORMATION FOR SEQ ID NO: 268:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 170 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 268:

```
Glu Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  *  Gly Gly Gly
 1               5                  10                  15

Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys
            20                  25                  30

Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala
        35                  40                  45

Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp
    50                  55                  60

Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr
65                  70                  75                  80

Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly
                85                  90                  95

Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu
            100                 105                 110

Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe Gly Ala Gly Lys Pro
        115                 120                 125

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

Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr
145                 150                 155                 160

Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
                165                 170
```

(2) INFORMATION FOR SEQ ID NO: 269:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 123 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 269:

```
GGGGGGGGGG AAGCTTGACC TGTGAAGTGA AAAATGGCGC AGATTGTGCG ACATTTTTTT      60

TGTCTGCCGT TTAATTAAAG GGGGGGGGGG GCCGGCCTGG GGGGGGGTGT ACAGGGGGGG     120

GGG                                                                   123
```

(2) INFORMATION FOR SEQ ID NO: 270:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 470 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 270:

```
GCTAGCACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG      60
TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC     120
TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGCATCCCT TTAGGGTTCC     180
GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT GGTTCTCGTA     240
GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA     300
ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG     360
ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA AAATGAGCTG ATTTAACAAA     420
AATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT TTCATGTACA                 470
```

(2) INFORMATION FOR SEQ ID NO: 271:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 733 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 271:

```
AGATCTGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA      60
AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC     120
AAAAAAACCA CCGCTACCAG CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT     180
TCCGAAGGTA ACTGGCTACA GCAGAGCGCA GATACCAAAT ACTGTTCTTC TAGTGTAGCC     240
GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT     300
CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT TGGACTCAAG     360
```

```
ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC     420
CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG     480
CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC     540
AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG     600
GTTTCGCCAC CTCTGACTTG AGCGTCGATT TTTGTGATGC TCGTCAGGGG GGCGGAGCCT     660
ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCTG GCCTTTTGCT GGCCTTTTGC     720
TCACATGGCT AGC                                                       733
```

(2) INFORMATION FOR SEQ ID NO: 272:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 813 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:102..758
(D) OTHER INFORMATION:/product= "cat resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 272:

```
GGGACGTCGG GTGAGGTTCC AACTTTCACC ATAATGAAAT AAGATCACTA CCGGGCGTAT      60

TTTTTGAGTT ATCGAGATTT TCAGGAGCTA AGGAAGCTAA A ATG GAG AAA AAA          113
                                                  Met Glu Lys Lys

ATC ACT GGA TAT ACC ACC GTT GAT ATA TCC CAA TGG CAT CGT AAA GAA        161
Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp His Arg Lys Glu
175                 180                 185                 190

CAT TTT GAG GCA TTT CAG TCA GTT GCT CAA TGT ACC TAT AAC CAG ACC        209
His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr Tyr Asn Gln Thr
                195                 200                 205

GTT CAG CTG GAT ATT ACG GCC TTT TTA AAG ACC GTA AAG AAA AAT AAG        257
Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val Lys Lys Asn Lys
            210                 215                 220

CAC AAG TTT TAT CCG GCC TTT ATT CAC ATT CTT GCC CGC CTG ATG AAT        305
His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala Arg Leu Met Asn
        225                 230                 235

GCT CAC CCG GAG TTC CGT ATG GCA ATG AAA GAC GGT GAG CTG GTG ATA        353
Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly Glu Leu Val Ile
    240                 245                 250
```

```
TGG GAT AGT GTT CAC CCT TGT TAC ACC GTT TTC CAT GAG CAA ACT GAA        401
Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His Glu Gln Thr Glu
255                 260                 265                 270

ACG TTT TCA TCG CTC TGG AGT GAA TAC CAC GAC GAT TTC CGG CAG TTT        449
Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp Phe Arg Gln Phe
                275                 280                 285

CTA CAC ATA TAT TCG CAA GAT GTG GCG TGT TAC GGT GAA AAC CTG GCC        497
Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly Glu Asn Leu Ala
            290                 295                 300

TAT TTC CCT AAA GGG TTT ATT GAG AAT ATG TTT TTC GTC TCA GCC AAT        545
Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe Val Ser Ala Asn
        305                 310                 315

CCC TGG GTG AGT TTC ACC AGT TTT GAT TTA AAC GTA GCC AAT ATG GAC        593
Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val Ala Asn Met Asp.
    320                 325                 330

AAC TTC TTC GCC CCC GTT TTC ACT ATG GGC AAA TAT TAT ACG CAA GGC        641
Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr Tyr Thr Gln Gly
335                 340                 345                 350

GAC AAG GTG CTG ATG CCG CTG GCG ATT CAG GTT CAT CAT GCC GTT TGT        689
Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His His Ala Val Cys
                355                 360                 365

GAT GGC TTC CAT GTC GGC AGA ATG CTT AAT GAA TTA CAA CAG TAC TGC        737
Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu Gln Gln Tyr Cys
            370                 375                 380

GAT GAG TGG CAG GGC GGG GCG TAATTTTTTT AAGGCAGTTA TTGGGTGCCC           788
Asp Glu Trp Gln Gly Gly Ala
        385

TTAAACGCCT GGTGCTAGAT CTTCC                                            813
```

(2) INFORMATION FOR SEQ ID NO: 273:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 273:

```
Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp
 1               5                  10                  15

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr
            20                  25                  30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val
        35                  40                  45

Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala

    50                  55                  60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly
65                  70                  75                  80

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His
                85                  90                  95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp
            100                 105                 110

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly
        115                 120                 125

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe
    130                 135                 140

145                 150                 155                 160

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr
                165                 170                 175

Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His
            180                 185                 190

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu
        195                 200                 205

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala
    210                 215
```

(2) INFORMATION FOR SEQ ID NO: 274:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2755 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic vector"

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:3..509
(D) OTHER INFORMATION:/product= "gIIIp ss, myc tag, amber codon"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:complement (1853..2509)
(D) OTHER INFORMATION:/product= "cat resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 274:

```
AA TTC GAG CAG AAG CTG ATC TCT GAG GAG GAT CTG TAG GGT GGT GGC          47
   Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  *  Gly Gly Gly
   220                 225                 230
```

```
TCT GGT TCC GGT GAT TTT GAT TAT GAA AAG ATG GCA AAC GCT AAT AAG        95
Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys
235                 240                 245                 250

GGG GCT ATG ACC GAA AAT GCC GAT GAA AAC GCG CTA CAG TCT GAC GCT       143
Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala
                255                 260                 265

AAA GGC AAA CTT GAT TCT GTC GCT ACT GAT TAC GGT GCT GCT ATC GAT       191
Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp
            270                 275                 280

GGT TTC ATT GGT GAC GTT TCC GGC CTT GCT AAT GGT AAT GGT GCT ACT       239
Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr
        285                 290                 295

GGT GAT TTT GCT GGC TCT AAT TCC CAA ATG GCT CAA GTC GGT GAC GGT       287
Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly
    300                 305                 310

GAT AAT TCA CCT TTA ATG AAT AAT TTC CGT CAA TAT TTA CCT TCC CTC       335
Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu
315                 320                 325                 330

CCT CAA TCG GTT GAA TGT CGC CCT TTT GTC TTT GGC GCT GGT AAA CCA       383
Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe Gly Ala Gly Lys Pro
                335                 340                 345

TAT GAA TTT TCT ATT GAT TGT GAC AAA ATA AAC TTA TTC CGT GGT GTC       431
Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn Leu Phe Arg Gly Val
            350                 355                 360

TTT GCG TTT CTT TTA TAT GTT GCC ACC TTT ATG TAT GTA TTT TCT ACG       479
Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr
        365                 370                 375

TTT GCT AAC ATA CTG CGT AAT AAG GAG TCT TGATAAGCTT GACCTGTGAA         529
Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
    380                 385

GTGAAAAATG GCGCAGATTG TGCGACATTT TTTTTGTCTG CCGTTTAATT AAAGGGGGGG     589

GGGGGCCGGC CTGGGGGGGG GTGTACATGA AATTGTAAAC GTTAATATTT TGTTAAAATT     649

CGCGTTAAAT TTTTGTTAAA TCAGCTCATT TTTTAACCAA TAGGCCGAAA TCGGCAAAAT     709

CCCTTATAAA TCAAAAGAAT AGACCGAGAT AGGGTTGAGT GTTGTTCCAG TTTGGAACAA     769

GAGTCCACTA TTAAAGAACG TGGACTCCAA CGTCAAAGGG CGAAAAACCG TCTATCAGGG     829

CGATGGCCCA CTACGAGAAC CATCACCCTA ATCAAGTTTT TTGGGGTCGA GGTGCCGTAA     889

AGCACTAAAT CGGAACCCTA AAGGGAGCCC CCGATTTAGA GCTTGACGGG GAAAGCCGGC     949

GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGCGCTAGGG CGCTGGCAAG    1009

TGTAGCGGTC ACGCTGCGCG TAACCACCAC ACCCGCCGCG CTTAATGCGC CGCTACAGGG    1069

CGCGTGCTAG CCATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA AAAGGCCGCG    1129
```

```
TTGCTGGCGT TTTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA    1189

AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCGTTTCC CCCTGGAAGC    1249

TCCCTCGTGC GCTCTCCTGT TCCGACCCTG CCGCTTACCG GATACCTGTC CGCCTTTCTC    1309

CCTTCGGGAA GCGTGGCGCT TTCTCATAGC TCACGCTGTA GGTATCTCAG TTCGGTGTAG    1369

GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC GAACCCCCCG TTCAGCCCGA CCGCTGCGCC    1429

TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGTAAGAC ACGACTTATC GCCACTGGCA    1489

GCAGCCACTG GTAACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC AGAGTTCTTG    1549

AAGTGGTGGC CTAACTACGG CTACACTAGA AGAACAGTAT TTGGTATCTG CGCTCTGCTG    1609

TAGCCAGTTA CCTTCGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAACA AACCACCGCT    1669

GGTAGCGGTG GTTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA    1729

GAAGATCCTT TGATCTTTTC TACGGGGTCT GACGCTCAGT GGAACGAAAA CTCACGTTAA    1789

GGGATTTTGG TCAGATCTAG CACCAGGCGT TTAAGGGCAC CAATAACTGC CTTAAAAAAA    1849

TTACGCCCCG CCCTGCCACT CATCGCAGTA CTGTTGTAAT TCATTAAGCA TTCTGCCGAC    1909

ATGGAAGCCA TCACAAACGG CATGATGAAC CTGAATCGCC AGCGGCATCA GCACCTTGTC    1969

GCCTTGCGTA TAATATTTGC CCATAGTGAA AACGGGGGCG AAGAAGTTGT CCATATTGGC    2029

TACGTTTAAA TCAAAACTGG TGAAACTCAC CCAGGGATTG GCTGAGACGA AAAACATATT    2089

CTCAATAAAC CCTTTAGGGA AATAGGCCAG GTTTTCACCG TAACACGCCA CATCTTGCGA    2149

ATATATGTGT AGAAACTGCC GGAAATCGTC GTGGTATTCA CTCCAGAGCG ATGAAAACGT    2209

TTCAGTTTGC TCATGGAAAA CGGTGTAACA AGGGTGAACA CTATCCCATA TCACCAGCTC    2269

ACCGTCTTTC ATTGCCATAC GGAACTCCGG GTGAGCATTC ATCAGGCGGG CAAGAATGTG    2329

AATAAAGGCC GGATAAAACT TGTGCTTATT TTTCTTTACG GTCTTTAAAA AGGCCGTAAT    2389

ATCCAGCTGA ACGGTCTGGT TATAGGTACA TTGAGCAACT GACTGAAATG CCTCAAAATG    2449

TTCTTTACGA TGCCATTGGG ATATATCAAC GGTGGTATAT CCAGTGATTT TTTTCTCCAT    2509

TTTAGCTTCC TTAGCTCCTG AAAATCTCGA TAACTCAAAA AATACGCCCG GTAGTGATCT    2569

TATTTCATTA TGGTGAAAGT TGGAACCTCA CCCGACGTCT AATGTGAGTT AGCTCACTCA    2629

TTAGGCACCC CAGGCTTTAC ACTTTATGCT TCCGGCTCGT ATGTTGTGTG GAATTGTGAG    2689

CGGATAACAA TTTCACACAG GAAACAGCTA TGACCATGAT TACGAATTTC TAGAGCATGC    2749

GGGGGG                                                               2755
```

(2) INFORMATION FOR SEQ ID NO: 275:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 169 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 275:

```
Phe Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu  *  Gly Gly Gly Ser

Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly
            20                  25                  30

Ala Met Thr Glu Asn Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys
        35                  40                  45

Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly
    50                  55                  60

Phe Ile Gly Asp Val Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly
65                  70                  75                  80

Asp Phe Ala Gly Ser Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp
                85                  90                  95

Asn Ser Pro Leu Met Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro
            100                 105                 110

Gln Ser Val Glu Cys Arg Pro Phe Val Phe Gly Ala Gly Lys Pro Tyr
        115                 120                 125

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

Ala Phe Leu Leu Tyr Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe
145                 150                 155                 160

Ala Asn Ile Leu Arg Asn Lys Glu Ser
                165
```

(2) INFORMATION FOR SEQ ID NO: 276:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 276:

```
Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp
1               5                  10                  15

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr
            20                  25                  30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val
        35                  40                  45
```

```
Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala
 50                  55                  60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly
 65                  70                  75                  80

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His
                 85                  90                  95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp
            100                 105                 110

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly
        115                 120                 125

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe
    130                 135                 140

Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val
145                 150                 155                 160

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr
                165                 170                 175

Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His
            180                 185                 190

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu
        195                 200                 205

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala
    210                 215
```

(2) INFORMATION FOR SEQ ID NO: 277:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 173 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 277:

```
GACGTCTTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC      60

CGGCTCGTAT GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG     120

ACCATGTCTA GAATAACTTC GTATAATGTA CGCTATACGA AGTTATCGCA TGC             173
```

(2) INFORMATION FOR SEQ ID NO: 278:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 278:
AGATCTCATA ACTTCGTATA ATGTATGCTA TACGAAGTTA TGACGTC 47

(2) INFORMATION FOR SEQ ID NO: 279:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1255 base pairs
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:1..1245
(D) OTHER INFORMATION:/product- "gIIIp, GGGGS linker"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 279:

```
GAA TTC GGT GGT GGT GGA TCT GCG TGC GCT GAA ACG GTT GAA AGT TGT      48
Glu Phe Gly Gly Gly Gly Ser Ala Cys Ala Glu Thr Val Glu Ser Cys
220                 225                 230                 235

TTA GCA AAA TCC CAT ACA GAA AAT TCA TTT ACT AAC GTC TGG AAA GAC      96
Leu Ala Lys Ser His Thr Glu Asn Ser Phe Thr Asn Val Trp Lys Asp
                240                 245                 250

GAC AAA ACT TTA GAT CGT TAC GCT AAC TAT GAG GGC TGT CTG TGG AAT     144
Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn
            255                 260                 265

GCT ACA GGC GTT GTA GTT TGT ACT GGT GAC GAA ACT CAG TGT TAC GGT     192
Ala Thr Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly
        270                 275                 280

ACA TGG GTT CCT ATT GGG CTT GCT ATC CCT GAA AAT GAG GGT GGT GGC     240
Thr Trp Val Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly
    285                 290                 295

TCT GAG GGT GGC GGT TCT GAG GGT GGC GGT TCT GAG GGT GGC GGT ACT     288
Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr
300                 305                 310                 315

AAA CCT CCT GAG TAC GGT GAT ACA CCT ATT CCG GGC TAT ACT TAT ATC     336
Lys Pro Pro Glu Tyr Gly Asp Thr Pro Ile Pro Gly Tyr Thr Tyr Ile
                320                 325                 330
```

```
AAC CCT CTC GAC GGC ACT TAT CCG CCT GGT ACT GAG CAA AAC CCC GCT        384
Asn Pro Leu Asp Gly Thr Tyr Pro Pro Gly Thr Glu Gln Asn Pro Ala
            335                 340                 345

AAT CCT AAT CCT TCT CTT GAG GAG TCT CAG CCT CTT AAT ACT TTC ATG        432
Asn Pro Asn Pro Ser Leu Glu Glu Ser Gln Pro Leu Asn Thr Phe Met
        350                 355                 360

TTT CAG AAT AAT AGG TTC CGA AAT AGG CAG GGG GCA TTA ACT GTT TAT        480
Phe Gln Asn Asn Arg Phe Arg Asn Arg Gln Gly Ala Leu Thr Val Tyr
    365                 370                 375

ACG GGC ACT GTT ACT CAA GGC ACT GAC CCC GTT AAA ACT TAT TAC CAG        528
Thr Gly Thr Val Thr Gln Gly Thr Asp Pro Val Lys Thr Tyr Tyr Gln
380                 385                 390                 395

TAC ACT CCT GTA TCA TCA AAA GCC ATG TAT GAC GCT TAC TGG AAC GGT        576
Tyr Thr Pro Val Ser Ser Lys Ala Met Tyr Asp Ala Tyr Trp Asn Gly
                400                 405                 410

AAA TTC AGA GAC TGC GCT TTC CAT TCT GGC TTT AAT GAG GAT TTA TTT        624
Lys Phe Arg Asp Cys Ala Phe His Ser Gly Phe Asn Glu Asp Leu Phe
            415                 420                 425

GTT TGT GAA TAT CAA GGC CAA TCG TCT GAC CTG CCT CAA CCT CCT GTC        672
Val Cys Glu Tyr Gln Gly Gln Ser Ser Asp Leu Pro Gln Pro Pro Val
        430                 435                 440

AAT GCT GGC GGC GGC TCT GGT GGT GGT TCT GGT GGC GGC TCT GAG GGT        720
Asn Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Glu Gly
    445                 450                 455

GGT GGC TCT GAG GGT GGC GGT TCT GAG GGT GGC GGC TCT GAG GGA GGC        768
Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly
460                 465                 470                 475

GGT TCC GGT GGT GGC TCT GGT TCC GGT GAT TTT GAT TAT GAA AAG ATG        816
Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met
                480                 485                 490

GCA AAC GCT AAT AAG GGG GCT ATG ACC GAA AAT GCC GAT GAA AAC GCG        864
Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala
            495                 500                 505

CTA CAG TCT GAC GCT AAA GGC AAA CTT GAT TCT GTC GCT ACT GAT TAC        912
Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr
        510                 515                 520

GGT GCT GCT ATC GAT GGT TTC ATT GGT GAC GTT TCC GGC CTT GCT AAT        960
Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn
    525                 530                 535

GGT AAT GGT GCT ACT GGT GAT TTT GCT GGC TCT AAT TCC CAA ATG GCT       1008
Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala
540                 545                 550                 555

CAA GTC GGT GAA GGT GAT AAT TCA CCT TTA ATG AAT AAT TTC CGT CAA       1056
Gln Val Gly Glu Gly Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln
                560                 565                 570
```

```
TAT TTA CCT TCC ATC CCT CAA TCG GTT GAA TGT CGC CCT TTT GTC TTT      1104
Tyr Leu Pro Ser Ile Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe
            575                 580                 585

GGC GCT GGT AAA CCC TAT GAA TTT TCT ATT GAT TGT GAC AAA ATA AAC      1152
Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn
        590                 595                 600

TTA TTC CGT GGT GTC TTT GCG TTT CTT TTA TAT GTT GCC ACC TTT ATG      1200
Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met
    605                 610                 615

TAT GTA TTT TCT ACG TTT GCT AAC ATA CTG CGT AAT AAG GAG TCT          1245
Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
620                 625                 630

TGATAAGCTT                                                          1255
```

(2) INFORMATION FOR SEQ ID NO: 280:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 415 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 280:

```
Glu Phe Gly Gly Gly Gly Ser Ala Cys Ala Glu Thr Val Glu Ser Cys
 1               5                  10                  15

            20                  25                  30

Asp Lys Thr Leu Asp Arg Tyr Ala Asn Tyr Glu Gly Cys Leu Trp Asn
        35                  40                  45

Ala Thr Gly Val Val Val Cys Thr Gly Asp Glu Thr Gln Cys Tyr Gly
    50                  55                  60

Thr Trp Val Pro Ile Gly Leu Ala Ile Pro Glu Asn Glu Gly Gly Gly
65                  70                  75                  80

Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Thr
                85                  90                  95

Lys Pro Pro Glu Tyr Gly Asp Thr Pro Ile Pro Gly Tyr Thr Tyr Ile
            100                 105                 110

Asn Pro Leu Asp Gly Thr Tyr Pro Pro Gly Thr Glu Gln Asn Pro Ala
        115                 120                 125

Asn Pro Asn Pro Ser Leu Glu Glu Ser Gln Pro Leu Asn Thr Phe Met
    130                 135                 140

Phe Gln Asn Asn Arg Phe Arg Asn Arg Gln Gly Ala Leu Thr Val Tyr
145                 150                 155                 160

Thr Gly Thr Val Thr Gln Gly Thr Asp Pro Val Lys Thr Tyr Tyr Gln
                165                 170                 175
```

```
Tyr Thr Pro Val Ser Ser Lys Ala Met Tyr Asp Ala Tyr Trp Asn Gly
            180                 185                 190

Lys Phe Arg Asp Cys Ala Phe His Ser Gly Phe Asn Glu Asp Leu Phe
        195                 200                 205

Val Cys Glu Tyr Gln Gly Gln Ser Ser Asp Leu Pro Gln Pro Pro Val
    210                 215                 220

Asn Ala Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly Gly Ser Glu Gly
225                 230                 235                 240

Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly Gly Ser Glu Gly Gly
                245                 250                 255

Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp Tyr Glu Lys Met
            260                 265                 270

Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala Asp Glu Asn Ala
        275                 280                 285

Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val Ala Thr Asp Tyr
    290                 295                 300

Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser Gly Leu Ala Asn
305                 310                 315                 320

Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn Ser Gln Met Ala
                325                 330                 335

Gln Val Gly Glu Gly Asp Asn Ser Pro Leu Met Asn Asn Phe Arg Gln
            340                 345                 350

Tyr Leu Pro Ser Ile Pro Gln Ser Val Glu Cys Arg Pro Phe Val Phe
        355                 360                 365

Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys Asp Lys Ile Asn
    370                 375                 380

Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val Ala Thr Phe Met
385                 390                 395                 400

Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn Lys Glu Ser
                405                 410                 415
```

(2) INFORMATION FOR SEQ ID NO: 281:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 502 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION:4..492
(D) OTHER INFORMATION:/product- "gIIIp ss"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 281:

```
CGG GAA TTC GGA GGC GGT TCC GGT GGT GGC TCT GGT TCC GGT GAT TTT          48
    Glu Phe Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe
                    420                 425                 430

GAT TAT GAA AAG ATG GCA AAC GCT AAT AAG GGG GCT ATG ACC GAA AAT          96
Asp Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn
                435                 440                 445

GCC GAT GAA AAC GCG CTA CAG TCT GAC GCT AAA GGC AAA CTT GAT TCT         144
Ala Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser
            450                 455                 460

GTC GCT ACT GAT TAC GGT GCT GCT ATC GAT GGT TTC ATT GGT GAC GTT         192
Val Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val
        465                 470                 475

TCC GGC CTT GCT AAT GGT AAT GGT GCT ACT GGT GAT TTT GCT GGC TCT         240
Ser Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser
    480                 485                 490

AAT TCC CAA ATG GCT CAA GTC GGT GAC GGT GAT AAT TCA CCT TTA ATG         288
Asn Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met
495                 500                 505                 510

AAT AAT TTC CGT CAA TAT TTA CCT TCC CTC CCT CAA TCG GTT GAA TGT         336
Asn Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys
                515                 520                 525

CGC CCT TTT GTC TTT GGC GCT GGT AAA CCA TAT GAA TTT TCT ATT GAT         384
Arg Pro Phe Val Phe Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp
            530                 535                 540

TGT GAC AAA ATA AAC TTA TTC CGT GGT GTC TTT GCG TTT CTT TTA TAT         432
Cys Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr
        545                 550                 555

GTT GCC ACC TTT ATG TAT GTA TTT TCT ACG TTT GCT AAC ATA CTG CGT         480
Val Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg
    560                 565                 570

AAT AAG GAG TCT TGATAAGCTT                                             502
Asn Lys Glu Ser
575
```

(2) INFORMATION FOR SEQ ID NO: 282:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 163 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 282:

```
Glu Phe Gly Gly Gly Ser Gly Gly Gly Ser Gly Ser Gly Asp Phe Asp
 1               5                  10                  15

Tyr Glu Lys Met Ala Asn Ala Asn Lys Gly Ala Met Thr Glu Asn Ala
            20                  25                  30

Asp Glu Asn Ala Leu Gln Ser Asp Ala Lys Gly Lys Leu Asp Ser Val
        35                  40                  45

Ala Thr Asp Tyr Gly Ala Ala Ile Asp Gly Phe Ile Gly Asp Val Ser
    50                  55                  60

Gly Leu Ala Asn Gly Asn Gly Ala Thr Gly Asp Phe Ala Gly Ser Asn
65                  70                  75                  80

Ser Gln Met Ala Gln Val Gly Asp Gly Asp Asn Ser Pro Leu Met Asn
                85                  90                  95

Asn Phe Arg Gln Tyr Leu Pro Ser Leu Pro Gln Ser Val Glu Cys Arg
            100                 105                 110

Pro Phe Val Phe Gly Ala Gly Lys Pro Tyr Glu Phe Ser Ile Asp Cys
        115                 120                 125

Asp Lys Ile Asn Leu Phe Arg Gly Val Phe Ala Phe Leu Leu Tyr Val
    130                 135                 140

Ala Thr Phe Met Tyr Val Phe Ser Thr Phe Ala Asn Ile Leu Arg Asn
145                 150                 155                 160

Lys Glu Ser
```

(2) INFORMATION FOR SEQ ID NO: 283:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 47 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 283:
GCATGCCATA ACTTCGTATA ATGTACGCTA TACGAAGTTA TAAGCTT 47

(2) INFORMATION FOR SEQ ID NO: 284:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1163 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:82..978
(D) OTHER INFORMATION:/product= "bla resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 284:

```
GGGGGTGTAC ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA      60

TAATATTGAA AAAGGAAGAG T ATG AGT ATT CAA CAT TTC CGT GTC GCC CTT       111
                        Met Ser Ile Gln His Phe Arg Val Ala Leu
                            165                 170

ATT CCC TTT TTT GCG GCA TTT TGC CTT CCT GTT TTT GCT CAC CCA GAA       159
Ile Pro Phe Phe Ala Ala Phe Cys Leu Pro Val Phe Ala His Pro Glu
    175                 180                 185

ACG CTG GTG AAA GTA AAA GAT GCT GAG GAT CAG TTG GGT GCG CGA GTG       207
Thr Leu Val Lys Val Lys Asp Ala Glu Asp Gln Leu Gly Ala Arg Val
190                 195                 200                 205

GGT TAC ATC GAA CTG GAT CTC AAC AGC GGT AAG ATC CTT GAG AGT TTT       255
Gly Tyr Ile Glu Leu Asp Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe
                210                 215                 220

CGC CCC GAA GAA CGT TTT CCA ATG ATG AGC ACT TTT AAA GTT CTG CTA       303
Arg Pro Glu Glu Arg Phe Pro Met Met Ser Thr Phe Lys Val Leu Leu
            225                 230                 235

TGT GGC GCG GTA TTA TCC CGT ATT GAC GCC GGG CAA GAG CAA CTC GGT       351
Cys Gly Ala Val Leu Ser Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly
        240                 245                 250

CGC CGC ATA CAC TAT TCT CAG AAT GAC TTG GTT GAG TAC TCA CCA GTC       399
Arg Arg Ile His Tyr Ser Gln Asn Asp Leu Val Glu Tyr Ser Pro Val
    255                 260                 265

ACA GAA AAG CAT CTT ACG GAT GGC ATG ACA GTA AGA GAA TTA TGC AGT       447
Thr Glu Lys His Leu Thr Asp Gly Met Thr Val Arg Glu Leu Cys Ser
270                 275                 280                 285

GCT GCC ATA ACC ATG AGT GAT AAC ACT GCG GCC AAC TTA CTT CTG ACA       495
Ala Ala Ile Thr Met Ser Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr
                290                 295                 300

ACG ATC GGA GGA CCG AAG GAG CTA ACC GCT TTT TTG CAC AAC ATG GGG       543
Thr Ile Gly Gly Pro Lys Glu Leu Thr Ala Phe Leu His Asn Met Gly
            305                 310                 315

GAT CAT GTA ACT CGC CTT GAT CGT TGG GAA CCG GAG CTG AAT GAA GCC       591
Asp His Val Thr Arg Leu Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala
        320                 325                 330

ATA CCA AAC GAC GAG CGT GAC ACC ACG ATG CCT GTA GCA ATG GCA ACA       639
Ile Pro Asn Asp Glu Arg Asp Thr Thr Met Pro Val Ala Met Ala Thr
```

```
    335                 340                 345

ACG TTG CGC AAA CTA TTA ACT GGC GAA CTA CTT ACT CTA GCT TCC CGG      687
Thr Leu Arg Lys Leu Leu Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg
350                 355                 360                 365

CAA CAG TTA ATA GAC TGG ATG GAG GCG GAT AAA GTT GCA GGA CCA CTT      735
Gln Gln Leu Ile Asp Trp Met Glu Ala Asp Lys Val Ala Gly Pro Leu
                370                 375                 380

CTG CGC TCG GCC CTT CCG GCT GGC TGG TTT ATT GCT GAT AAA TCT GGA      783
Leu Arg Ser Ala Leu Pro Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly
            385                 390                 395

GCC GGT GAG CGT GGG TCT CGC GGT ATC ATT GCA GCA CTG GGG CCA GAT      831
Ala Gly Glu Arg Gly Ser Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp
        400                 405                 410

GGT AAG CCC TCC CGT ATC GTA GTT ATC TAC ACG ACG GGG AGT CAG GCA      879
Gly Lys Pro Ser Arg Ile Val Val Ile Tyr Thr Thr Gly Ser Gln Ala
    415                 420                 425

ACT ATG GAT GAA CGA AAT AGA CAG ATC GCT GAG ATA GGT GCC TCA CTG      927
Thr Met Asp Glu Arg Asn Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu
430                 435                 440                 445

ATT AAG CAT TGG GTA ACT GTC AGA CCA AGT TTA CTC ATA TAT ACT TTA      975
Ile Lys His Trp Val Thr Val Arg Pro Ser Leu Leu Ile Tyr Thr Leu
                450                 455                 460

GAT TGATTTAAAA CTTCATTTTT AATTTAAAAG GATCTAGGTG AAGATCCTTT          1028
Asp

TTGATAATCT CATGACCAAA ATCCCTTAAC GTGAGTTTTC GTTCCACTGA GCGTCAGACC   1088

CCGTAGAAAA GATCAAAGGA TCTTCTTGAG ATCCTTTTTG ATAATGGCCG GCCCCCCCCC   1148

TTAATTAAGG GGGGG                                                   1163
```

(2) INFORMATION FOR SEQ ID NO: 285:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 299 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 285:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp Val Thr
        275                 280                 285

Val Arg Pro Ser Leu Leu Ile Tyr Thr Leu Asp
    290                 295
```

(2) INFORMATION FOR SEQ ID NO: 286:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 470 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 286:

```
GCTAGCACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG      60
TGACCGCTAC ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC     120
TCGCCACGTT CGCCGGCTTT CCCCGTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC     180
GATTTAGTGC TTTACGGCAC CTCGACCCCA AAAAACTTGA TTAGGGTGAT GGTTCTCGTA     240
GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC GTTGGAGTCC ACGTTCTTTA     300
ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC TATTCTTTTG     360
ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA AAATGAGCTG ATTTAACAAA     420
AATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT TTCATGTACA                 470
```

(2) INFORMATION FOR SEQ ID NO: 287:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 832 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 287:

AGATCTAATA AGATGATCTT CTTGAGATCG TTTTGGTCTG CGCGTAATCT CTTGCTCTGA 60

AAACGAAAAA ACCGCCTTGC AGGGCGGTTT TTCGTAGGTT CTCTGAGCTA CCAACTCTTT 120

GAACCGAGGT AACTGGCTTG GAGGAGCGCA GTCACTAAAA CTTGTCCTTT CAGTTTAGCC 180

TTAACCGGCG CATGACTTCA AGACTAACTC CTCTAAATCA ATTACCAGTG GCTGCTGCCA 240

GTGGTGCTTT TGCATGTCTT TCCGGGTTGG ACTCAAGACG ATAGTTACCG GATAAGGCGC 300

AGCGGTCGGA CTGAACGGGG GGTTCGTGCA TACAGTCCAG CTTGGAGCGA ACTGCCTACC 360

CGGAACTGAG TGTCAGGCGT GGAATGAGAC AAACGCGGCC ATAACAGCGG AATGACACCG 420

GTAAACCGAA AGGCAGGAAC AGGAGAGCGC AGGAGGGAGC CGCCAGGGGG AAACGCCTGG 480

TATCTTTATA GTCCTGTCGG GTTTCGCCAC CACTGATTTG AGCGTCAGAT TTCGTGATGC 540

TTGTCAGGGG GGCGGAGCCT ATGGAAAAAC GGCTTTGCCG CGGCCCTCTC ACTTCCCTGT 600

TAAGTATCTT CCTGGCATCT TCCAGGAAAT CTCCGCCCCG TTCGTAAGCC ATTTCCGCTC 660

GCCGCAGTCG AACGACCGAG CGTAGCGAGT CAGTGAGCGA GGAAGCGGAA TATATCCTGT 720

ATCACATATT CTGCTGACGC ACCGGTGCAG CCTTTTTTCT CCTGCCACAT GAAGCACTTC 780

ACTGACACCC TCATCAGTGC CAACATAGTA AGCCAGTATA CACTCCGCTA GC 832

(2) INFORMATION FOR SEQ ID NO: 288:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 49 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 288:
AGATCTCATA ACTTCGTATA ATGTATGCTA TACGAAGTTA TTCAGATCT 49

(2) INFORMATION FOR SEQ ID NO: 289:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 289:

```
TCTAGAGCAT GCGTAGGAGA AAATAAAATG AAACAAAGCA CTATTGCACT GGCACTCTTA      60

CCGTTGCTCT TCACCCCTGT TACCAAAGCC GAATTC                               96
```

(2) INFORMATION FOR SEQ ID NO: 290:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 120 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 290:

```
TCTAGAGCAT GCGTAGGAGA AAATAAAATG AAACAAAGCA CTATTGCACT GGCACTCTTA      60

CCGTTGCTCT TCACCCCTGT TACCAAAGCC GACTACAAAG ATGAAGTGCA ATTGGAATTC     120
```

(2) INFORMATION FOR SEQ ID NO: 291:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic DNA cassette"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 291:

```
TCTAGAGGTT GAGGTGATTT TATGAAAAAG AATATCGCAT TTCTTCTTGC ATCTATGTTC      60

GTTTTTTCTA TTGCTACAAA TGCATACGCT GAATTC                               96
```

(2) INFORMATION FOR SEQ ID NO: 292:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1221 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:79..1158
(D) OTHER INFORMATION:/product= "lacI"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 292:

```
GCTAGCATCG AATGGCGCAA AACCTTTCGC GGTATGGCAT GATAGCGCCC GGAAGAGAGT      60

CAATTCAGGG TGGTGAAT GTG AAA CCA GTA ACG TTA TAC GAT GTC GCA GAG       111
                    Val Lys Pro Val Thr Leu Tyr Asp Val Ala Glu
                    300                 305                 310

TAT GCC GGT GTC TCT TAT CAG ACC GTT TCC CGC GTG GTG AAC CAG GCC       159
Tyr Ala Gly Val Ser Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala
                315                 320                 325

AGC CAC GTT TCT GCG AAA ACG CGG GAA AAA GTG GAA GCG GCG ATG GCG       207
Ser His Val Ser Ala Lys Thr Arg Glu Lys Val Glu Ala Ala Met Ala
            330                 335                 340

GAG CTG AAT TAC ATT CCT AAC CGC GTG GCA CAA CAA CTG GCG GGC AAA       255
Glu Leu Asn Tyr Ile Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys
        345                 350                 355
```

```
CAG TCG TTG CTG ATT GGC GTT GCC ACC TCC AGT CTG GCC CTG CAC GCG       303
Gln Ser Leu Leu Ile Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala
    360                 365                 370

CCG TCG CAA ATT GTC GCG GCG ATT AAA TCT CGC GCC GAT CAA CTG GGT       351
Pro Ser Gln Ile Val Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly
375                 380                 385                 390

GCC AGC GTG GTC GTG TCG ATG GTA GAA CGA AGC GGC GTC GAA GCC TGT       399
Ala Ser Val Val Val Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys
                395                 400                 405

AAA GCG GCG GTG CAC AAT CTT CTC GCG CAA CGT GTC AGT GGG CTG ATT       447
Lys Ala Ala Val His Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile
            410                 415                 420

ATT AAC TAT CCG CTG GAT GAC CAG GAT GCT ATT GCT GTG GAA GCT GCC       495
Ile Asn Tyr Pro Leu Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala
        425                 430                 435

TGC ACT AAT GTT CCG GCG TTA TTT CTT GAT GTC TCT GAC CAG ACA CCC       543
Cys Thr Asn Val Pro Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro
    440                 445                 450

ATC AAC AGT ATT ATT TTC TCC CAT GAG GAC GGT ACG CGA CTG GGC GTG       591
Ile Asn Ser Ile Ile Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val
455                 460                 465                 470

GAG CAT CTG GTC GCA TTG GGC CAC CAG CAA ATC GCG CTG TTA GCT GGC       639
Glu His Leu Val Ala Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly
                475                 480                 485

CCA TTA AGT TCT GTC TCG GCG CGT CTG CGT CTG GCT GGC TGG CAT AAA       687
Pro Leu Ser Ser Val Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys
            490                 495                 500

TAT CTC ACT CGC AAT CAA ATT CAG CCG ATA GCG GAA CGG GAA GGC GAC       735
Tyr Leu Thr Arg Asn Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp
        505                 510                 515

TGG AGT GCC ATG TCC GGT TTT CAA CAA ACC ATG CAA ATG CTG AAT GAG       783
Trp Ser Ala Met Ser Gly Phe Gln Gln Thr Met Gln Met Leu Asn Glu
    520                 525                 530

GGC ATC GTT CCC ACT GCG ATG CTG GTT GCC AAC GAT CAG ATG GCG CTG       831
Gly Ile Val Pro Thr Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu
535                 540                 545                 550

GGC GCA ATG CGT GCC ATT ACC GAG TCC GGG CTG CGC GTT GGT GCG GAC       879
                555                 560                 565

ATC TCG GTA GTG GGA TAC GAC GAT ACC GAG GAC AGC TCA TGT TAT ATC       927
Ile Ser Val Val Gly Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile
            570                 575                 580

CCG CCG CTG ACC ACC ATC AAA CAG GAT TTT CGC CTG CTG GGG CAA ACC       975
Pro Pro Leu Thr Thr Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr
        585                 590                 595
```

```
AGC GTG GAC CGC TTG CTG CAA CTC TCT CAG GGC CAG GCG GTG AAG GGC      1023
Ser Val Asp Arg Leu Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly
    600                 605                 610

AAT CAG CTG TTG CCC GTC TCA CTG GTG AAA AGA AAA ACC ACC CTG GCT      1071
Asn Gln Leu Leu Pro Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala
615                 620                 625                 630

CCC AAT ACG CAA ACC GCC TCT CCC CGC GCG TTG GCC GAT TCA CTG ATG      1119
Pro Asn Thr Gln Thr Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met
                635                 640                 645

CAG CTG GCA CGA CAG GTT TCC CGA CTG GAA AGC GGG CAG TGAGGCTACC       1168
Gln Leu Ala Arg Gln Val Ser Arg Leu Glu Ser Gly Gln
            650                 655

CGATAAAAGC GGCTTCCTGA CAGGAGGCCG TTTTGTTTTG CAGCCCACTT AAG          1221
```

(2) INFORMATION FOR SEQ ID NO: 293:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 360 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 293:

```
Val Lys Pro Val Thr Leu Tyr Asp Val Ala Glu Tyr Ala Gly Val Ser
 1               5                  10                  15

Tyr Gln Thr Val Ser Arg Val Val Asn Gln Ala Ser His Val Ser Ala
            20                  25                  30

Lys Thr Arg Glu Lys Val Glu Ala Ala Met Ala Glu Leu Asn Tyr Ile
        35                  40                  45

Pro Asn Arg Val Ala Gln Gln Leu Ala Gly Lys Gln Ser Leu Leu Ile
    50                  55                  60

Gly Val Ala Thr Ser Ser Leu Ala Leu His Ala Pro Ser Gln Ile Val
65                  70                  75                  80

Ala Ala Ile Lys Ser Arg Ala Asp Gln Leu Gly Ala Ser Val Val Val
                85                  90                  95

Ser Met Val Glu Arg Ser Gly Val Glu Ala Cys Lys Ala Ala Val His
            100                 105                 110

Asn Leu Leu Ala Gln Arg Val Ser Gly Leu Ile Ile Asn Tyr Pro Leu
        115                 120                 125

Asp Asp Gln Asp Ala Ile Ala Val Glu Ala Ala Cys Thr Asn Val Pro
    130                 135                 140

Ala Leu Phe Leu Asp Val Ser Asp Gln Thr Pro Ile Asn Ser Ile Ile
145                 150                 155                 160

Phe Ser His Glu Asp Gly Thr Arg Leu Gly Val Glu His Leu Val Ala
```

```
                    165                 170                 175

Leu Gly His Gln Gln Ile Ala Leu Leu Ala Gly Pro Leu Ser Ser Val
            180                 185                 190

Ser Ala Arg Leu Arg Leu Ala Gly Trp His Lys Tyr Leu Thr Arg Asn
        195                 200                 205

Gln Ile Gln Pro Ile Ala Glu Arg Glu Gly Asp Trp Ser Ala Met Ser
    210                 215                 220

Gly Phe Gln Gln Thr Met Gln Met Leu Asn Glu Gly Ile Val Pro Thr
225                 230                 235                 240

Ala Met Leu Val Ala Asn Asp Gln Met Ala Leu Gly Ala Met Arg Ala
                245                 250                 255

Ile Thr Glu Ser Gly Leu Arg Val Gly Ala Asp Ile Ser Val Val Gly
            260                 265                 270

Tyr Asp Asp Thr Glu Asp Ser Ser Cys Tyr Ile Pro Pro Leu Thr Thr
        275                 280                 285

Ile Lys Gln Asp Phe Arg Leu Leu Gly Gln Thr Ser Val Asp Arg Leu
    290                 295                 300

Leu Gln Leu Ser Gln Gly Gln Ala Val Lys Gly Asn Gln Leu Leu Pro
305                 310                 315                 320

Val Ser Leu Val Lys Arg Lys Thr Thr Leu Ala Pro Asn Thr Gln Thr
                325                 330                 335

Ala Ser Pro Arg Ala Leu Ala Asp Ser Leu Met Gln Leu Ala Arg Gln
            340                 345                 350

Val Ser Arg Leu Glu Ser Gly Gln
        355                 360
```

(2) INFORMATION FOR SEQ ID NO: 294:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2380 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic vector"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:complement (51..707)
(D) OTHER INFORMATION:/product= "cat resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 294:

GATCTAGCAC CAGGCGTTTA AGGGCACCAA TAACTGCCTT AAAAAAATTA CGCCCCGCCC 60

TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTC TGCCGACATG GAAGCCATCA 120

CAAACGGCAT GATGAACCTG AATCGCCAGC GGCATCAGCA CCTTGTCGCC TTGCGTATAA 180

TATTTGCCCA TAGTGAAAAC GGGGGCGAAG AAGTTGTCCA TATTGGCTAC GTTTAAATCA 240

AAACTGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCCT 300

TTAGGGAAAT AGGCCAGGTT TTCACCGTAA CACGCCACAT CTTGCGAATA TATGTGTAGA 360

AACTGCCGGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA 420

TGGAAAACGG TGTAACAAGG GTGAACACTA TCCCATATCA CCAGCTCACC GTCTTTCATT 480

GCCATACGGA ACTCCGGGTG AGCATTCATC AGGCGGGCAA GAATGTGAAT AAAGGCCGGA 540

TAAAACTTGT GCTTATTTTT CTTTACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG 600

GTCTGGTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAAATGTTC TTTACGATGC 660

CATTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCCTTA 720

GCTCCTGAAA ATCTCGATAA CTCAAAAAAT ACGCCCGGTA GTGATCTTAT TTCATTATGG 780

TGAAAGTTGG AACCTCACCC GACGTCTAAT GTGAGTTAGC TCACTCATTA GGCACCCCAG 840

GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT 900

CACACAGGAA ACAGCTATGA CCATGATTAC GAATTTCTAG ACCCCCCCCC CGCATGCCAT 960

AACTTCGTAT AATGTACGCT ATACGAAGTT ATAAGCTTGA CCTGTGAAGT GAAAAATGGC 1020

GCAGATTGTG CGACATTTTT TTTGTCTGCC GTTTAATTAA AGGGGGGGGG GGGCCGGCCT 1080

GGGGGGGGGT GTACATGAAA TTGTAAACGT TAATATTTTG TTAAAATTCG CGTTAAATTT 1140

TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC GGCAAAATCC CTTATAAATC 1200

AAAAGAATAG ACCGAGATAG GGTTGAGTGT TGTTCCAGTT TGGAACAAGA GTCCACTATT 1260

AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC TATCAGGGCG ATGGCCCACT 1320

ACGAGAACCA TCACCCTAAT CAAGTTTTTT GGGGTCGAGG TGCCGTAAAG CACTAAATCG 1380

GAACCCTAAA GGGAGCCCCC GATTTAGAGC TTGACGGGGA AAGCCGGCGA ACGTGGCGAG 1440

AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG CTGGCAAGTG TAGCGGTCAC 1500

GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG CTACAGGGCG CGTGCTAGCG 1560

GAGTGTATAC TGGCTTACTA TGTTGGCACT GATGAGGGTG TCAGTGAAGT GCTTCATGTG 1620

GCAGGAGAAA AAAGGCTGCA CCGGTGCGTC AGCAGAATAT GTGATACAGG ATATATTCCG 1680

CTTCCTCGCT CACTGACTCG CTACGCTCGG TCGTTCGACT GCGGCGAGCG GAAATGGCTT 1740

ACGAACGGGG CGGAGATTTC CTGGAAGATG CCAGGAAGAT ACTTAACAGG GAAGTGAGAG 1800

GGCCGCGGCA AAGCCGTTTT TCCATAGGCT CCGCCCCCCT GACAAGCATC ACGAAATCTG 1860

```
ACGCTCAAAT CAGTGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC   1920

TGGCGGCTCC CTCCTGCGCT CTCCTGTTCC TGCCTTTCGG TTTACCGGTG TCATTCCGCT   1980

GTTATGGCCG CGTTTGTCTC ATTCCACGCC TGACACTCAG TTCCGGGTAG GCAGTTCGCT   2040

CCAAGCTGGA CTGTATGCAC GAACCCCCCG TTCAGTCCGA CCGCTGCGCC TTATCCGGTA   2100

ACTATCGTCT TGAGTCCAAC CCGGAAAGAC ATGCAAAAGC ACCACTGGCA GCAGCCACTG   2160

GTAATTGATT TAGAGGAGTT AGTCTTGAAG TCATGCGCCG GTTAAGGCTA AACTGAAAGG   2220

ACAAGTTTTA GTGACTGCGC TCCTCCAAGC CAGTTACCTC GGTTCAAAGA GTTGGTAGCT   2280

CAGAGAACCT ACGAAAAACC GCCCTGCAAG GCGGTTTTTT CGTTTTCAGA GCAAGAGATT   2340

ACGCGCAGAC CAAAACGATC TCAAGAAGAT CATCTTATTA                          2380
```

(2) INFORMATION FOR SEQ ID NO: 295:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 295:

```
Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr
            20                  25                  30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val
        35                  40                  45

Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala
    50                  55                  60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly
65                  70                  75                  80

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His
                85                  90                  95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp
            100                 105                 110

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly
        115                 120                 125

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe
    130                 135                 140

Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val
145                 150                 155                 160

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr
                165                 170                 175


Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His
            180                 185                 190

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu
        195                 200                 205

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala
    210                 215
```

(2) INFORMATION FOR SEQ ID NO: 296:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3488 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc - "synthetic vector"

(A) NAME/KEY: CDS
(B) LOCATION:complement (1341..1997)
(D) OTHER INFORMATION:/product= "cat resistance"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:complement (2521..3417)
(D) OTHER INFORMATION:/product= "bla resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 296:

GTACATGAAA TTGTAAACGT TAATATTTTG TTAAAATTCG CGTTAAATTT TTGTTAAATC 60

AGCTCATTTT TTAACCAATA GGCCGAAATC GGCAAAATCC CTTATAAATC AAAAGAATAG 120

ACCGAGATAG GGTTGAGTGT TGTTCCAGTT TGGAACAAGA GTCCACTATT AAAGAACGTG 180

GACTCCAACG TCAAAGGGCG AAAAACCGTC TATCAGGGCG ATGGCCCACT ACGAGAACCA 240

TCACCCTAAT CAAGTTTTTT GGGGTCGAGG TGCCGTAAAG CACTAAATCG GAACCCTAAA 300

GGGAGCCCCC GATTTAGAGC TTGACGGGGA AAGCCGGCGA ACGTGGCGAG AAAGGAAGGG 360

AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA 420

ACCACCACAC CCGCCGCGCT TAATGCGCCG CTACAGGGCG CGTGCTAGCG GAGTGTATAC 480

TGGCTTACTA TGTTGGCACT GATGAGGGTG TCAGTGAAGT GCTTCATGTG GCAGGAGAAA 540

AAAGGCTGCA CCGGTGCGTC AGCAGAATAT GTGATACAGG ATATATTCCG CTTCCTCGCT 600

CACTGACTCG CTACGCTCGG TCGTTCGACT GCGGCGAGCG GAAATGGCTT ACGAACGGGG 660

CGGAGATTTC CTGGAAGATG CCAGGAAGAT ACTTAACAGG GAAGTGAGAG GGCCGCGGCA 720

AAGCCGTTTT TCCATAGGCT CCGCCCCCCT GACAAGCATC ACGAAATCTG ACGCTCAAAT 780

CAGTGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGCGGCTCC 840

CTCCTGCGCT CTCCTGTTCC TGCCTTTCGG TTTACCGGTG TCATTCCGCT GTTATGGCCG 900

CGTTTGTCTC ATTCCACGCC TGACACTCAG TTCCGGGTAG GCAGTTCGCT CCAAGCTGGA 960

CTGTATGCAC GAACCCCCCG TTCAGTCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT 1020

TGAGTCCAAC CCGGAAAGAC ATGCAAAAGC ACCACTGGCA GCAGCCACTG GTAATTGATT 1080

TAGAGGAGTT AGTCTTGAAG TCATGCGCCG GTTAAGGCTA AACTGAAAGG ACAAGTTTTA 1140

GTGACTGCGC TCCTCCAAGC CAGTTACCTC GGTTCAAAGA GTTGGTAGCT CAGAGAACCT 1200

ACGAAAAACC GCCCTGCAAG GCGGTTTTTT CGTTTTCAGA GCAAGAGATT ACGCGCAGAC 1260

CAAAACGATC TCAAGAAGAT CATCTTATTA GATCTAGCAC CAGGCGTTTA AGGGCACCAA 1320

TAACTGCCTT AAAAAAATTA CGCCCCGCCC TGCCACTCAT CGCAGTACTG TTGTAATTCA 1380

TTAAGCATTC TGCCGACATG GAAGCCATCA CAAACGGCAT GATGAACCTG AATCGCCAGC 1440

GGCATCAGCA CCTTGTCGCC TTGCGTATAA TATTTGCCCA TAGTGAAAAC GGGGGCGAAG 1500

AAGTTGTCCA TATTGGCTAC GTTTAAATCA AAACTGGTGA AACTCACCCA GGGATTGGCT 1560

GAGACGAAAA ACATATTCTC AATAAACCCT TTAGGGAAAT AGGCCAGGTT TTCACCGTAA 1620

CACGCCACAT CTTGCGAATA TATGTGTAGA AACTGCCGGA AATCGTCGTG GTATTCACTC 1680

CAGAGCGATG AAAACGTTTC AGTTTGCTCA TGGAAAACGG TGTAACAAGG GTGAACACTA 1740

TCCCATATCA CCAGCTCACC GTCTTTCATT GCCATACGGA ACTCCGGGTG AGCATTCATC 1800

AGGCGGGCAA GAATGTGAAT AAAGGCCGGA TAAAACTTGT GCTTATTTTT CTTTACGGTC 1860

TTTAAAAAGG CCGTAATATC CAGCTGAACG GTCTGGTTAT AGGTACATTG AGCAACTGAC 1920

TGAAATGCCT CAAAATGTTC TTTACGATGC CATTGGGATA TATCAACGGT GGTATATCCA 1980

GTGATTTTTT TCTCCATTTT AGCTTCCTTA GCTCCTGAAA ATCTCGATAA CTCAAAAAAT 2040

ACGCCCGGTA GTGATCTTAT TTCATTATGG TGAAAGTTGG AACCTCACCC GACGTCTAAT 2100

GTGAGTTAGC TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG 2160

TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC 2220

GAATTTCTAG ACCCCCCCCC CGCATGCCAT AACTTCGTAT AATGTACGCT ATACGAAGTT 2280

ATAAGCTTGA CCTGTGAAGT GAAAAATGGC GCAGATTGTG CGACATTTTT TTTGTCTGCC 2340

GTTTAATTAA GGGGGGGGGC CGGCCATTAT CAAAAAGGAT CTCAAGAAGA TCCTTTGATC 2400

TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG 2460

AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA 2520

ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC CAATGCTTAA TCAGTGAGGC 2580

```
ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA    2640
GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA    2700
CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG    2760
CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAACTGTT GCCGGGAAGC    2820
TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT    2880
CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG    2940
GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT    3000
CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA    3060
TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA    3120
GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA    3180
TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG    3240
GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGCGC    3300
ACCCAACTGA TCCTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG    3360
AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT    3420
CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT    3480
ATTTGAAT                                                            3488
```

(2) INFORMATION FOR SEQ ID NO: 297:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 219 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 297:

```
Met Glu Lys Lys Ile Thr Gly Tyr Thr Thr Val Asp Ile Ser Gln Trp
 1               5                  10                  15

His Arg Lys Glu His Phe Glu Ala Phe Gln Ser Val Ala Gln Cys Thr
            20                  25                  30

Tyr Asn Gln Thr Val Gln Leu Asp Ile Thr Ala Phe Leu Lys Thr Val
        35                  40                  45

Lys Lys Asn Lys His Lys Phe Tyr Pro Ala Phe Ile His Ile Leu Ala
    50                  55                  60

Arg Leu Met Asn Ala His Pro Glu Phe Arg Met Ala Met Lys Asp Gly
65                  70                  75                  80

Glu Leu Val Ile Trp Asp Ser Val His Pro Cys Tyr Thr Val Phe His
                85                  90                  95

Glu Gln Thr Glu Thr Phe Ser Ser Leu Trp Ser Glu Tyr His Asp Asp
            100                 105                 110

Phe Arg Gln Phe Leu His Ile Tyr Ser Gln Asp Val Ala Cys Tyr Gly
        115                 120                 125

Glu Asn Leu Ala Tyr Phe Pro Lys Gly Phe Ile Glu Asn Met Phe Phe
    130                 135                 140

Val Ser Ala Asn Pro Trp Val Ser Phe Thr Ser Phe Asp Leu Asn Val
145                 150                 155                 160

Ala Asn Met Asp Asn Phe Phe Ala Pro Val Phe Thr Met Gly Lys Tyr
                165                 170                 175

Tyr Thr Gln Gly Asp Lys Val Leu Met Pro Leu Ala Ile Gln Val His
            180                 185                 190

His Ala Val Cys Asp Gly Phe His Val Gly Arg Met Leu Asn Glu Leu
        195                 200                 205

Gln Gln Tyr Cys Asp Glu Trp Gln Gly Gly Ala
    210                 215
```

(2) INFORMATION FOR SEQ ID NO: 298:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 299 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 298:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp Val Thr
        275                 280                 285

Val Arg Pro Ser Leu Leu Ile Tyr Thr Leu Asp
    290                 295
```

(2) INFORMATION FOR SEQ ID NO: 299:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 2728 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: circular

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic vector"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:complement (471..1367)
(D) OTHER INFORMATION:/product= "bla resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 299:

```
GATCTCATAA CTTCGTATAA TGTATGCTAT ACGAAGTTAT GACGTCTAAT GTGAGTTAGC      60

TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA     120

TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GAATTTCTAG     180

ACCCCCCCCC CGCATGCCAT AACTTCGTAT AATGTACGCT ATACGAAGTT ATAAGCTTGA     240
```

CCTGTGAAGT GAAAAATGGC GCAGATTGTG CGACATTTTT TTGTCTGCC GTTTAATTAA 300

GGGGGGGGGC CGGCCATTAT CAAAAAGGAT CTCAAGAAGA TCCTTTGATC TTTTCTACGG 360

GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG AGATTATCAA 420

AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA 480

TATATGAGTA AACTTGGTCT GACAGTTACC CAATGCTTAA TCAGTGAGGC ACCTATCTCA 540

GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG 600

ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA 660

CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT 720

CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAACTGTT GCCGGGAAGC TAGAGTAAGT 780

AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA 840

CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA 900

TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA 960

AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT 1020

GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA 1080

GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG 1140

CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACTC 1200

TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGCGC ACCCAACTGA 1260

TCCTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT 1320

GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT 1380

CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT 1440

ACATGAAATT GTAAACGTTA ATATTTTGTT AAAATTCGCG TTAAATTTTT GTTAAATCAG 1500

CTCATTTTTT AACCAATAGG CCGAAATCGG CAAAATCCCT TATAAATCAA AAGAATAGAC 1560

CGAGATAGGG TTGAGTGTTG TTCCAGTTTG GAACAAGAGT CCACTATTAA AGAACGTGGA 1620

CTCCAACGTC AAAGGGCGAA AAACCGTCTA TCAGGGCGAT GGCCCACTAC GAGAACCATC 1680

ACCCTAATCA AGTTTTTTGG GGTCGAGGTG CCGTAAAGCA CTAAATCGGA ACCCTAAAGG 1740

GAGCCCCCGA TTTAGAGCTT GACGGGGAAA GCCGGCGAAC GTGGCGAGAA AGGAAGGGAA 1800

GAAAGCGAAA GGAGCGGGCG CTAGGGCGCT GGCAAGTGTA GCGGTCACGC TGCGCGTAAC 1860

CACCACACCC GCCGCGCTTA ATGCGCCGCT ACAGGGCGCG TGCTAGCGGA GTGTATACTG 1920

GCTTACTATG TTGGCACTGA TGAGGGTGTC AGTGAAGTGC TTCATGTGGC AGGAGAAAAA 1980

AGGCTGCACC GGTGCGTCAG CAGAATATGT GATACAGGAT ATATTCCGCT TCCTCGCTCA 2040

```
CTGACTCGCT ACGCTCGGTC GTTCGACTGC GGCGAGCGGA AATGGCTTAC GAACGGGGCG    2100

GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG CCGCGGCAAA    2160

GCCGTTTTTC CATAGGCTCC GCCCCCCTGA CAAGCATCAC GAAATCTGAC GCTCAAATCA    2220

GTGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTTCCCCCTG GCGGCTCCCT    2280

CCTGCGCTCT CCTGTTCCTG CCTTTCGGTT TACCGGTGTC ATTCCGCTGT TATGGCCGCG    2340

TTTGTCTCAT TCCACGCCTG ACACTCAGTT CCGGGTAGGC AGTTCGCTCC AAGCTGGACT    2400

GTATGCACGA ACCCCCCGTT CAGTCCGACC GCTGCGCCTT ATCCGGTAAC TATCGTCTTG    2460

AGTCCAACCC GGAAAGACAT GCAAAAGCAC CACTGGCAGC AGCCACTGGT AATTGATTTA    2520

GAGGAGTTAG TCTTGAAGTC ATGCGCCGGT TAAGGCTAAA CTGAAAGGAC AAGTTTTAGT    2580

GACTGCGCTC CTCCAAGCCA GTTACCTCGG TTCAAAGAGT TGGTAGCTCA GAGAACCTAC    2640

GAAAAACCGC CCTGCAAGGC GGTTTTTTCG TTTTCAGAGC AAGAGATTAC GCGCAGACCA    2700

AAACGATCTC AAGAAGATCA TCTTATTA                                       2728
```

(2) INFORMATION FOR SEQ ID NO: 300:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 299 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 300:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp Val Thr
        275                 280                 285

Val Arg Pro Ser Leu Leu Ile Tyr Thr Leu Asp
    290                 295
```

(2) INFORMATION FOR SEQ ID NO: 301:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 301:
TATGAGATCT CATAACTTCG TATAATGTAC GCTATACGAA GTTAT 45

(2) INFORMATION FOR SEQ ID NO: 302:

(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 302:
TAATAACTTC GTATAGCATA CATTATACGA AGTTATGAGA TCTCA 45

(2) INFORMATION FOR SEQ ID NO: 303:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 91 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 303:

CATTTTTTGC CCTCGTTATC TACGCATGCG ATAACTTCGT ATAGCGTACA TTATACGAAG 60

TTATTCTAGA CATGGTCATA GCTGTTTCCT G 91

(2) INFORMATION FOR SEQ ID NO: 304:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 52 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 304:
GGGGGGAATT CGGTGGTGGT GGATCTGCGT GCGCTGAAAC GGTTGAAAGT TG 52

(2) INFORMATION FOR SEQ ID NO: 305:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 305:
CCCCCCCAAG CTTATCAAGA CTCCTTATTA CG 32

(2) INFORMATION FOR SEQ ID NO: 306:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 306:
GGGGGGGGAA TTCGGAGGCG GTTCCGGTGG TGGC 34

(2) INFORMATION FOR SEQ ID NO: 307:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 74 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 307:

GGGGGGGGAA TTCGAGCAGA AGCTGATCTC TGAGGAGGAT CTGTAGGGTG GTGGCTCTGG 60

TTCCGGTGAT TTTG 74

(2) INFORMATION FOR SEQ ID NO: 308:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 308:
CCATAACTTC GTATAATGTA CGCTATACGA AGTTATA 37

(2) INFORMATION FOR SEQ ID NO: 309:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 309:
AGCTTATAAC TTCGTATAGC GTACATTATA CGAAGTTATG GCATG 45

(2) INFORMATION FOR SEQ ID NO: 310:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 310:

AGCTTGACCT GTGAAGTGAA AAATGGCGCA GATTGTGCGA CATTTTTTTT GTCTGCCGTT 60

TAATTAAAGG GGGGGT 76

(2) INFORMATION FOR SEQ ID NO: 311:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 75 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 311:

GTACACCCCC CCCCAGGCCG GCCCCCCCCC CCCTTTAATT AAACGGCAGA CAAAAAAAAT 60

GTCGCACAAT CTGCG 75

(2) INFORMATION FOR SEQ ID NO: 312:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 312:
GGGGGGGTGT ACATTCAAAT ATGTATCCGC TCATG 35

(2) INFORMATION FOR SEQ ID NO: 313:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 313:
GGGTTACATC GAACTGGATC TC 22

(2) INFORMATION FOR SEQ ID NO: 314:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 59 base pairs

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 314:
CCAGTTCGAT GTAACCCACT CGCGCACCCA ACTGATCCTC AGCATCTTTT ACTTTCACC 59

(2) INFORMATION FOR SEQ ID NO: 315:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 315:
ACTCTAGCTT CCCGGCAACA GTTAATAGAC TGGATGGAGG CGG 43

(2) INFORMATION FOR SEQ ID NO: 316:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 316:
CTGTTGCCGG GAAGCTAGAG TAAG 24

(2) INFORMATION FOR SEQ ID NO: 317:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 58 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 317:
CCCCCCCTTA ATTAAGGGGG GGGGCCGGCC ATTATCAAAA AGGATCTCAA GAAGATCC 58

(2) INFORMATION FOR SEQ ID NO: 318:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 318:
GGGGGGGGCT AGCACGCGCC CTGTAGCGGC GCATTAA 37

(2) INFORMATION FOR SEQ ID NO: 319:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 319:
CCCCCCCTGT ACATGAAATT GTAAACGTTA ATATTTTG 38

(2) INFORMATION FOR SEQ ID NO: 320:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 320:
GGGCGATGGC CCACTACGAG AACCATCACC CTAATC 36

(2) INFORMATION FOR SEQ ID NO: 321:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 321:
GGGGGGAGAT CTAATAAGAT GATCTTCTTG AG 32

(2) INFORMATION FOR SEQ ID NO: 322:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 322:
GAGTTGGTAG CTCAGAGAAC CTACGAAAAA CCGCCCTGCA AGGCG 45

(2) INFORMATION FOR SEQ ID NO: 323:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 323:
GTAGGTTCTC TGAGCTACCA ACTC 24

(2) INFORMATION FOR SEQ ID NO: 324:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 324:
GTTTCCCCCT GGCGGCTCCC TCCTGCGCTC TCCTGTTCCT GCC 43

(2) INFORMATION FOR SEQ ID NO: 325:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 325:
AGGAGGGAGC CGCCAGGGGG AAAC 24

(2) INFORMATION FOR SEQ ID NO: 326:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 326:
GACATCAGCG CTAGCGGAGT GTATAC 26

(2) INFORMATION FOR SEQ ID NO: 327:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 43 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 327:
GATCTCATAA CTTCGTATAA TGTATGCTAT ACGAAGTTAT TCA 43

(2) INFORMATION FOR SEQ ID NO: 328:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 45 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 328:
GATCTGAATA ACTTCGTATA GCATACATTA TACGAAGTTA TGAGA 45

(2) INFORMATION FOR SEQ ID NO: 329:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 329:
GGGGGGGAGA TCTGACCAAA ATCCCTTAAC GTGAG 35

(2) INFORMATION FOR SEQ ID NO: 330:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 330:
GGTATCTGCG CTCTGCTGTA GCCAGTTACC TTCGG 35

(2) INFORMATION FOR SEQ ID NO: 331:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 331:
CCCCCCCGCT AGCCATGTGA GCAAAAGGCC AGCAA 35

(2) INFORMATION FOR SEQ ID NO: 332:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 332:
GGGACGTCGG GTGAGGTTCC AAC 23

(2) INFORMATION FOR SEQ ID NO: 333:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 333:
CCATACGGAA CTCCGGGTGA GCATTCATC 29

(2) INFORMATION FOR SEQ ID NO: 334:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 334:
CCGGAGTTCC GTATGG 16

(2) INFORMATION FOR SEQ ID NO: 335:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 335:
ACGTTTAAAT CAAAACTGG 19

(2) INFORMATION FOR SEQ ID NO: 336:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 69 base pairs

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 336:

CCAGTTTTGA TTTAAACGTA GCCAATATGG ACAACTTCTT CGCCCCCGTT TTCACTATGG 60

GCAAATATT 69

(2) INFORMATION FOR SEQ ID NO: 337:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 337:
GGAAGATCTA GCACCAGGCG TTTAAG 26

(2) INFORMATION FOR SEQ ID NO: 338:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 338:
GAGGCCGGCC ATCGAATGGC GCAAAAC 27

(2) INFORMATION FOR SEQ ID NO: 339:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 339:
CGCGTACCGT CCTCATGGGA GAAAATAATA C 31

(2) INFORMATION FOR SEQ ID NO: 340:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 83 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 340:

CCATGAGGAC GGTACGCGAC TGGGCGTGGA GCATCTGGTC GCATTGGGTC ACCAGCAAAT 60

CCGCTGTTAG CTGGCCCATT AAG 83

(2) INFORMATION FOR SEQ ID NO: 341:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 341:
GTCAGCGGCG GGATATAACA TGAGCTGTCC TCGGTATCGT CG 42

(2) INFORMATION FOR SEQ ID NO: 342:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 342:
GTTATATCCC GCCGCTGACC ACCATCAAAC 30

(2) INFORMATION FOR SEQ ID NO: 343:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 65 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(ix) FEATURE:

(A) NAME/KEY: conflict
(B) LOCATION:replace(42..44, "")
(D) OTHER INFORMATION:/note= "in Fig.35b, M41, LAC6: T4T; but see Fig.35a, M41: LAC6 pos. 1055-1119 on complementary strand, 1076 to 1078: TAT"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 343:

CATCAGTGAA TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGAG CCAGGGTGGT 60

TTTTC 65

(2) INFORMATION FOR SEQ ID NO: 344:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 73 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 344:
GGTTAATTAA CCTCACTGCC CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATCAG 60
TGAATCGGCC AAC 73

(2) INFORMATION FOR SEQ ID NO: 345:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 345:
CTAGACTAGT GTTTAAACCG GACCGGGGGG GGGCTTAAGG GGGGGGGGGG 50

(2) INFORMATION FOR SEQ ID NO: 346:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 346:
CTAGCCCCCC CCCCCCTTAA GCCCCCCCCC GGTCCGGTTT AAACACTAGT 50

(2) INFORMATION FOR SEQ ID NO: 347:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 50 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 347:
CTAGACTAGT GTTTAAACCG GACCGGGGGG GGGCTTAAGG GGGGGGGGGG 50

(2) INFORMATION FOR SEQ ID NO: 348:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 82 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 348:

CCCCCCCTTA AGTGGGCTGC AAAACAAAAC GGCCTCCTGT CAGGAAGCCG CTTTTATCGG 60

GTAGCCTCAC TGCCCGCTTT CC 82

(2) INFORMATION FOR SEQ ID NO: 349:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 40 base pairs
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 349:
GTTGTTGTGC CACGCGGTTA GGAATGTAAT TCAGCTCCGC 40

(2) INFORMATION FOR SEQ ID NO: 350:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 350:
AACCGCGTGG CACAACAAC 19

(2) INFORMATION FOR SEQ ID NO: 351:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 41 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 351:
CTTCGTTCTA CCATCGACAC GACCACGCTG GCACCCAGTT G 41

(2) INFORMATION FOR SEQ ID NO: 352:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 352:
GTGTCGATGG TAGAACGAAG 20

(2) INFORMATION FOR SEQ ID NO: 353:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 353:

CCACAGCAAT AGCATCCTGG TCATCCAGCG GATAGTTAAT AATCAGCCCA CTGACACGTT 60

GCGCGAG 67

(2) INFORMATION FOR SEQ ID NO: 354:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 354:
GACCAGGATG CTATTGCTGT GG 22

(2) INFORMATION FOR SEQ ID NO: 355:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 355:
CAGCGCGATT TGCTGGTGGC CCAATGCGAC CAGATGC 37

(2) INFORMATION FOR SEQ ID NO: 356:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 356:
CACCAGCAAA TCGCGCTG 18

(2) INFORMATION FOR SEQ ID NO: 357:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 357:
CCCGGACTCG GTAATGGCAC GCATTGCGCC CAGCGCC 37

(2) INFORMATION FOR SEQ ID NO: 358:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 358:
GCCATTACCG AGTCCGGG 18

(2) INFORMATION FOR SEQ ID NO: 359:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 359:
AATTCCACCA TCATCACCAT TGACGTCTA 29

(2) INFORMATION FOR SEQ ID NO: 360:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 360:
AGCTTAGACG TCAATGGTGA TGATGGTGG 29

(2) INFORMATION FOR SEQ ID NO: 361:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1289 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic gene cassette"

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION:complement (280..1137)
(D) OTHER INFORMATION:/product= "bla resistance"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 361:

```
CGCGTTAACC TCAGGTGACC AAGCCCCTGG CCAAGGTCCC GTACGTTCGA AGATTACCAT      60
CACGTGGATC CGGTACCAGG CCGGCCATTA TCAAAAAGGA TCTCAAGAAG ATCCTTTGAT     120
CTTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA CGTTAAGGGA TTTTGGTCAT     180
GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTTAAAT TAAAAATGAA GTTTTAAATC     240
AATCTAAAGT ATATATGAGT AAACTTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC     300
ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA     360
GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA     420
CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA GGGCCGAGCG     480
CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAACTGTT GCCGGGAAGC     540
TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT     600
CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG     660
GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT     720
CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA     780
TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA     840
GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA     900
TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG     960
GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC    1020
ACCCAACTGA TCTTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG    1080
AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT    1140
CTTCCTTTTT CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT    1200
ATTTGAATGT ACTCGGCCGC ACGAGCTGCA GGCGCCATTA ATGGCTCGAG CGCGCTTCAG    1260
CGCTTTGTCT TCCGGATGTA CATGAAATT                                      1289
```

(2) INFORMATION FOR SEQ ID NO: 362:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 286 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 362:

```
Met Ser Ile Gln His Phe Arg Val Ala Leu Ile Pro Phe Phe Ala Ala
 1               5                  10                  15

Phe Cys Leu Pro Val Phe Ala His Pro Glu Thr Leu Val Lys Val Lys
            20                  25                  30

Asp Ala Glu Asp Gln Leu Gly Ala Arg Val Gly Tyr Ile Glu Leu Asp
        35                  40                  45

Leu Asn Ser Gly Lys Ile Leu Glu Ser Phe Arg Pro Glu Glu Arg Phe
    50                  55                  60

Pro Met Met Ser Thr Phe Lys Val Leu Leu Cys Gly Ala Val Leu Ser
65                  70                  75                  80

Arg Ile Asp Ala Gly Gln Glu Gln Leu Gly Arg Arg Ile His Tyr Ser
                85                  90                  95

Gln Asn Asp Leu Val Glu Tyr Ser Pro Val Thr Glu Lys His Leu Thr
            100                 105                 110

Asp Gly Met Thr Val Arg Glu Leu Cys Ser Ala Ala Ile Thr Met Ser
        115                 120                 125

Asp Asn Thr Ala Ala Asn Leu Leu Leu Thr Thr Ile Gly Gly Pro Lys
    130                 135                 140

Glu Leu Thr Ala Phe Leu His Asn Met Gly Asp His Val Thr Arg Leu
145                 150                 155                 160

Asp Arg Trp Glu Pro Glu Leu Asn Glu Ala Ile Pro Asn Asp Glu Arg
                165                 170                 175

Asp Thr Thr Met Pro Val Ala Met Ala Thr Thr Leu Arg Lys Leu Leu
            180                 185                 190

Thr Gly Glu Leu Leu Thr Leu Ala Ser Arg Gln Gln Leu Ile Asp Trp
        195                 200                 205

Met Glu Ala Asp Lys Val Ala Gly Pro Leu Leu Arg Ser Ala Leu Pro
    210                 215                 220

Ala Gly Trp Phe Ile Ala Asp Lys Ser Gly Ala Gly Glu Arg Gly Ser
225                 230                 235                 240

Arg Gly Ile Ile Ala Ala Leu Gly Pro Asp Gly Lys Pro Ser Arg Ile
                245                 250                 255

Val Val Ile Tyr Thr Thr Gly Ser Gln Ala Thr Met Asp Glu Arg Asn
            260                 265                 270

Arg Gln Ile Ala Glu Ile Gly Ala Ser Leu Ile Lys His Trp
            275                 280                 285
```

(2) INFORMATION FOR SEQ ID NO: 363:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 363:
GCCCTGCAAG CGGAAGAC 18

(2) INFORMATION FOR SEQ ID NO: 364:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 364:
GGCTTTCGAA TGGCCAAAGG 20

(2) INFORMATION FOR SEQ ID NO: 365:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:25..27
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (ACT/GTT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:37..39
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (TTT,CAT,CTT,ATG,CAG)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:43..45
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (18 codons, no Pro, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:46..48
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:49..51
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:52..54
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:55..57
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (CCT/TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:58..60
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19 aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 365:

GCCCTGCAAG CGGAAGACTT TGCGRYTTAT TATTGCHWKC AGNNKDVTDV TNNKYCTNNK 60

ACCTTTGGCC ATTCGAAAGC C 81

(2) INFORMATION FOR SEQ ID NO: 366:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:37..39
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (TTT,CAT,CTT,ATG,CAG)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature

(B) LOCATION:43..45
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (18 codons, no Pro, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:46..48
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:49..51
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:52..54
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:55..57
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (CCT/TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:58..60
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19 aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 366:

```
GCCCTGCAAG CGGAAGACGT GGGCGTGTAT TATTGCHWKC AGNNKDVTDV TNNKYCTNNK      60

ACCTTTGGCC ATTCGAAAGC C                                                81
```

(2) INFORMATION FOR SEQ ID NO: 367:

(i) SEQUENCE CHARACTERISTICS:

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

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:37..39

(D) OTHER INFORMATION:/product= "random codon by trinucleotides (TTT,CAT,CTT,ATG,CAG)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:43..45
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (18 codons, no Pro, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:46..48
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:49..51
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (GAT, GGT, AAT, TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:52..54
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:55..57
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (CCT/TCT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:58..60
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19 aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 367:

```
GCCCTGCAAG CGGAAGACGT GGCGGTGTAT TATTGCHWKC AGNNKDVTDV TNNKYCTNNK      60

ACCTTTGGCC ATTCGAAAGC C                                                81
```

(2) INFORMATION FOR SEQ ID NO: 368:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 108 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:41..43
(D) OTHER INFORMATION:/Product= "random codon by trinucleotides (CGT, TGG, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:47..61
(D) OTHER INFORMATION:/product= "random codons by trinucleotides (18 aa, no Trp, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:62..64
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 368:

CCTGCAAGCG GAAGACGAAG CGGATTATTA TTGCCAGAGC YRKGACNNKN NKNNKNNKNN 60

KNNKGGCGGC GGCACGAAGT TAACCGTTCT TGGCCAGGAA TTCGAGCC 108

(2) INFORMATION FOR SEQ ID NO: 369:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 105 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:41..43
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (CGT, TGG, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:47..58
(D) OTHER INFORMATION:/product- "random codons by trinucleotides (18 aa, no Trp, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:59..61
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 369:

CCTGCAAGCG GAAGACGAAG CGGATTATTA TTGCCAGAGC YRKGACNNKN NKNNKNNKNN 60

KGGCGGCGGC ACGAAGTTAA CCGTTCTTGG CCAGGAATTC GAGCC 105

(2) INFORMATION FOR SEQ ID NO: 370:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 102 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(A) DESCRIPTION: /desc = "synthetic oligonucleotide library"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:41..43
(D) OTHER INFORMATION:/product= "random codon by trinucleotides (CGT, TGG, TAT)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:47..55
(D) OTHER INFORMATION:/product= "random codons by trinucleotides (18 aa, no Trp, no Cys)"

(ix) FEATURE:

(A) NAME/KEY: misc_feature
(B) LOCATION:56..58
(D) OTHER INFORMATION:/product= "random codon by trinucleotide mutagenesis (19aa, no Cys)"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 370:

CCTGCAAGCG GAAGACGAAG CGGATTATTA TTGCCAGAGC YRKGACNNKN NKNNKNNKGG 60

CGGCGGCACG AAGTTAACCG TTCTTGGCCA GGAATTCGAG CC 102

(2) INFORMATION FOR SEQ ID NO: 371:

(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 = "synthetic oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 371:

GGCTCGAATT CCTGGCC 17

Claims

1. A modular vector, comprising (i) a nucleotide sequence encoding an immunoglobulin variable region, comprising a modular sequence of four consensus framework regions and complementarity determining regions CDR1, CDR2, and CDR3, wherein said nucleotide sequence comprises DNA cleavage sites at the boundary of each consensus framework region and each complementarity determining region, and (ii) (a) vector module(s), wherein each vector module is flanked by DNA cleavage sites, wherein each of said cleavage sites of (i) and (ii) is unique within said vector, and wherein said immunoglobulin variable region is a heavy chain or a light chain.

2. The modular vector according to claim 1, wherein said modules are taken from the list comprising: origins of single-stranded replication, origins of double-stranded replication for high- and low copy number plasmids, promoter/operator, repressor or terminator elements, resistance genes, potential recombination sites, gene III for display on filamentous phages, truncated gene III for display on filamentous phages, signal sequences, purification and detection tags, and sequences of additional moieties.

3. The modular vector according to claim 2, wherein said additional moieties are selected from the list of: a toxin, a cytokine, a reporter enzyme, a moiety being capable of binding a metal ion, a peptide, a tag suitable for detection and/or purification, or a homo-or hetero-association domain.

4. The modular vector of any one of claims 1 to 3, which is a cloning vector.

5. The modular vector of any one of claims 1 to 3, which is an expression vector.

6. The modular vector of claim 5, which is a vector suitable for expression and screening of libraries.

Patentansprüche

1. Modularer Vektor, umfassend (i) eine Nucleotidsequenz, die eine variable Immunglobulinregion codiert, die eine modulare Sequenz von vier Konsensus-Gerüstregionen und die komplementaritäts-bestimmenden Regionen CDR1, CDR2 und CDR3 umfasst, wobei die Nucleotidsequenz an der Grenze zu jeder der Konsensus-Gerüstregionen und jeder der komplementaritäts-bestimmenden Regionen DNA-Schnittstellen umfasst, und (ii) (ein) Vektormodul(e), wobei jedes Vektormodul von DNA-Schnittstellen flankiert wird, wobei jede der Schnittstellen von (i) und (ii) einmalig innerhalb des Vektors ist und wobei die variable Immunoglobulinregion eine schwere oder eine leichte Kette ist.

2. Modularer Vektor nach Anspruch 1, wobei die Module aus der Liste stammen, umfassend: Startpunkte der Einzelstrangreplikation, Startpunkte der Doppelstrangreplikation für Plasmide mit hoher und niedriger Kopiezahl, Promotor-/Operator-, Repressor- oder Terminatorelemente, Resistenzgene, potenzielle Rekombinationsstellen, Gen III zur Darstellung auf filamentösen Phagen, verkürztes Gen III zur Darstellung auf filamentösen Phagen, Signalsequenzen, Reinigungs- und Nachweistags und Sequenzen zusätzlicher Einheiten.

3. Modularer Vektor nach Anspruch 2, wobei die zusätzlichen Einheiten ausgewählt sind aus der Liste aus: einem Toxin, einem Cytokin, einem Reporterenzym, einer Einheit, die in der Lage ist, ein Metallion zu binden, einem Peptid, einem Tag, der für den Nachweis und/oder die Reinigung geeignet ist, oder einer Homo- oder Hetero-Assoziationsdomäne.

4. Modularer Vektor nach einem der Ansprüche 1 bis 3, der ein Clonierungsvektor ist.

5. Modularer Vektor nach einem der Ansprüche 1 bis 3, der ein Expressionsvektor ist.

6. Modularer Vektor nach Anspruch 5, der ein für die Expression und das Durchmustern von Bibliotheken geeigneter Vektor ist.

Revendications

1. Vecteur modulaire, comprenant (i) une séquence de nucléotides codant pour une région variable d'immunoglobuline, comprenant une séquence modulaire de quatre régions charpente consensus et régions déterminant la complémentarité CDR1, CDR2, et CDR3, dans lequel ladite séquence de nucléotides comprend des sites de coupure d'ADN à la limite de chaque région charpente consensus et de chaque région déterminant la complémentarité, et (ii) (un) ou des module(s) de vecteur, dans lequel chaque module de vecteur est bordé par des sites de coupure d'ADN, dans lequel chacun desdits sites de coupure de (i) et (ii) est unique dans ledit vecteur, et dans lequel ladite région variable de l'immunoglobuline est une chaîne lourde ou une chaîne légère.

2. Vecteur modulaire selon la revendication 1, dans lequel lesdits modules sont choisis parmi la liste comprenant : les origines de réplication monocaténaire, les origines de réplication bicaténaire pour les plasmides à haut et à faible nombre de copies, les promoteurs/opérateurs, les éléments de répression ou de terminaison, les gènes de résistance, les sites de recombinaison potentiels, le gène III pour une présentation sur les phages filamenteux, le gène III tronqué pour une présentation sur les phages filamenteux, les séquences signal, les marqueurs de purification et de détection, et les séquences de fragments supplémentaires.

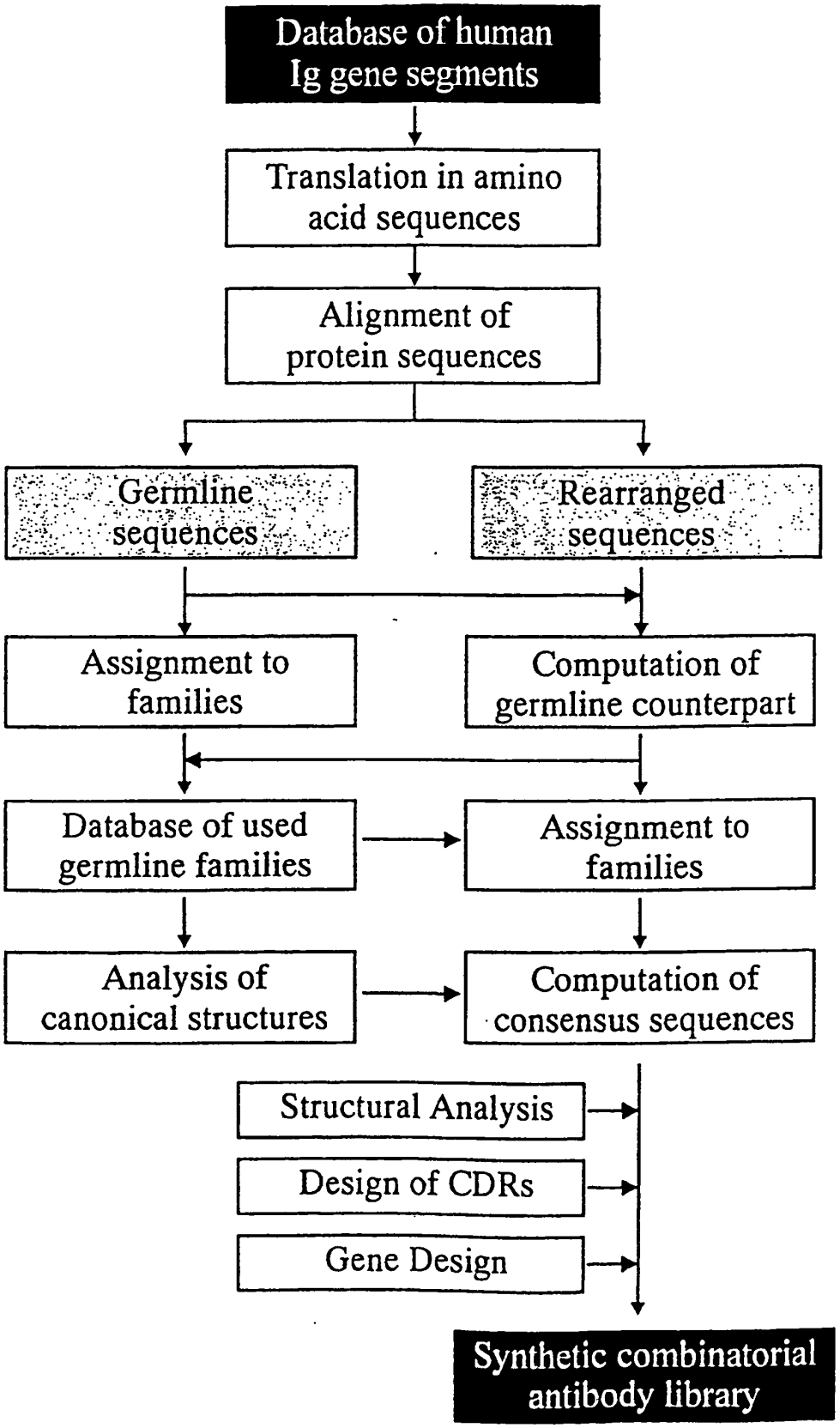
3. Vecteur modulaire selon la revendication 2, dans lequel lesdits fragments supplémentaires sont choisis parmi la liste constituée par : une toxine, une cytokine, une enzyme rapporteuse, fragment capable de fixer un ion métallique, un peptide, un marqueur approprié à la détection et/ou à la purification, ou un domaine d'homo- ou hétéro-association.

4. Vecteur modulaire selon l'une quelconque des revendications 1 à 3, qui est un vecteur de clonage.

5. Vecteur modulaire selon l'une quelconque des revendications 1 à 3, qui est un vecteur d'expression.

6. Vecteur modulaire selon la revendication 5, qui est un vecteur adapté pour l'expression et le criblage de banques.

Figure 1: construction of a synthetic human antibody library based on consensus sequences



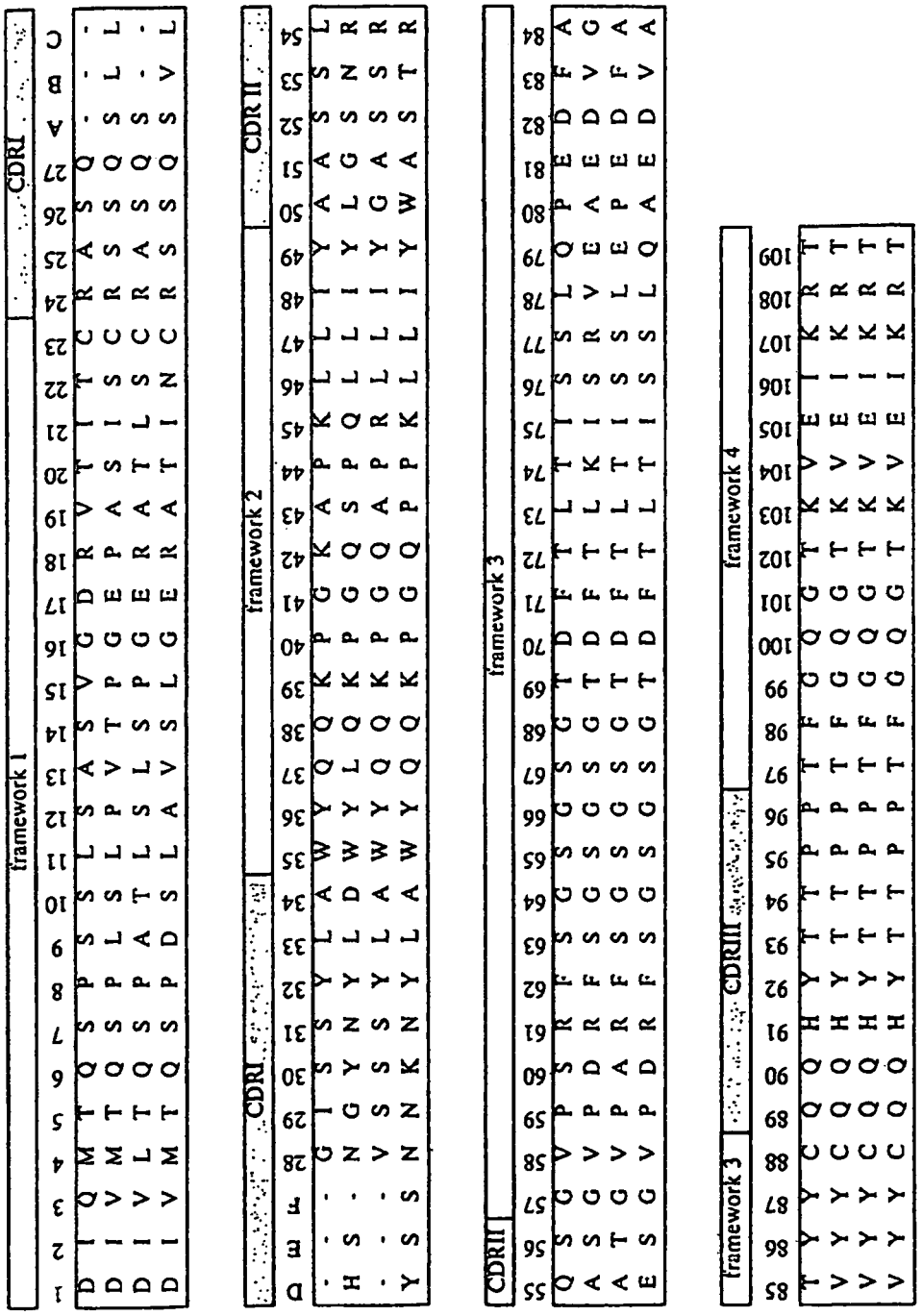


Figure 2A: VL kappa consensus sequences

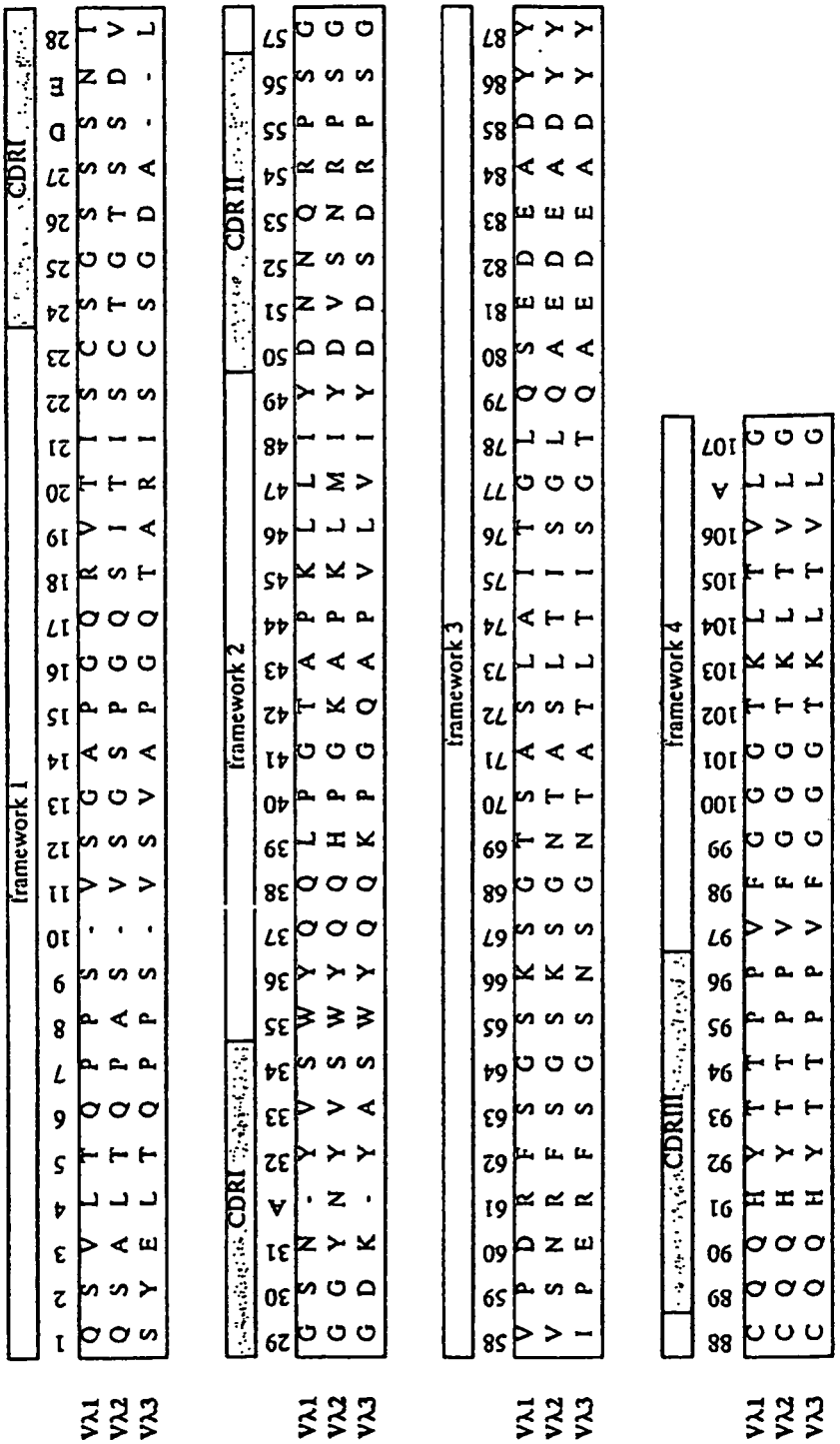


Figure 2B: VL lambda consensus sequences

Figure 2C: V heavy chain consensus sequences

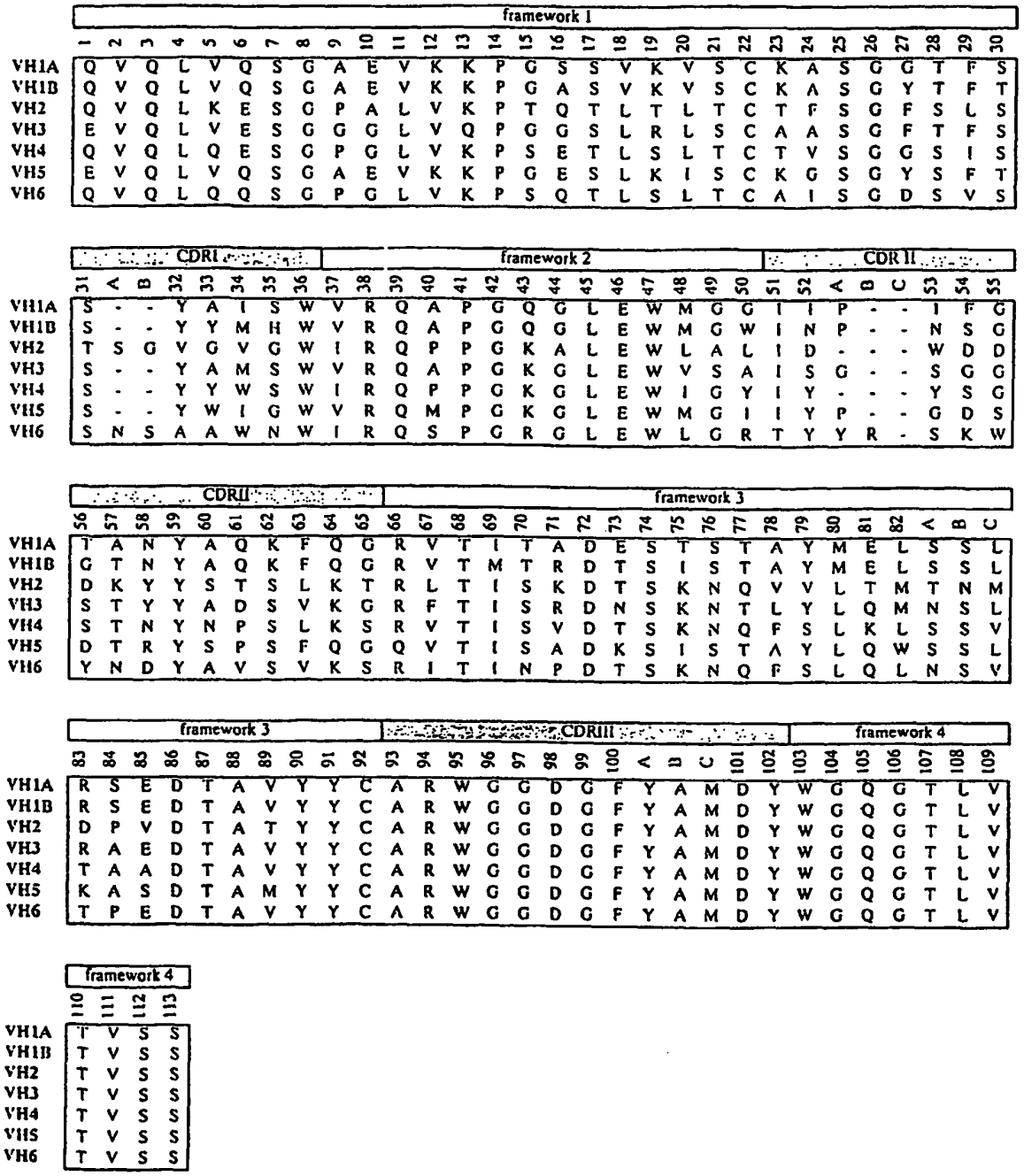


Figure 3A: V kappa 1 (Vκ1) gene sequence

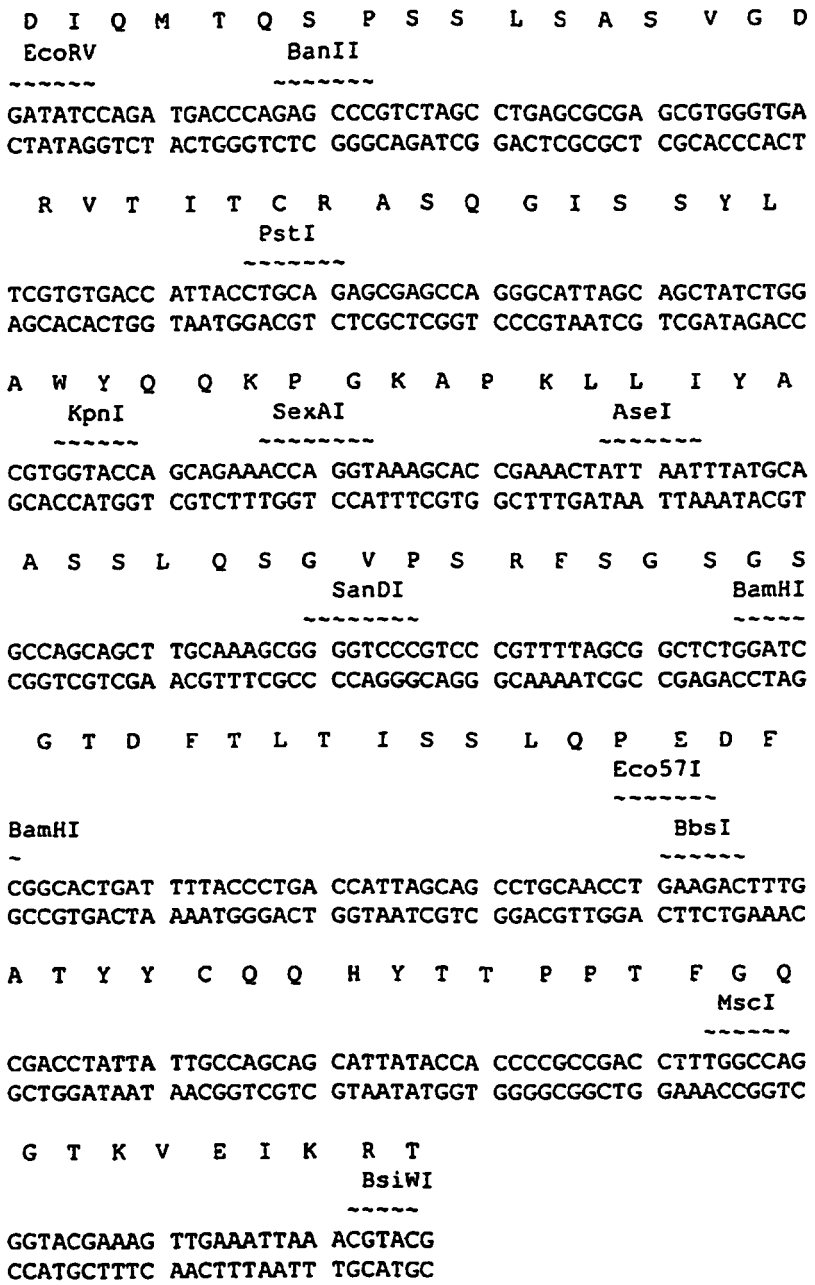


Figure 3B: V kappa 2 (Vκ2) gene sequence

```
 D  I  V  M   T  Q  S   P  L  S   L  P  V  T   P  G  E
 EcoRV          BanII
~~~~~~         ~~~~~~~
GATATCGTGA TGACCCAGAG CCCACTGAGC CTGCCAGTGA CTCCGGGCGA
CTATAGCACT ACTGGGTCTC GGGTGACTCG GACGGTCACT GAGGCCCGCT

  P  A  S   I  S  C  R   S  S  Q   S  L  L   H  S  N
                PstI
               ~~~~~~~
GCCTGCGAGC ATTAGCTGCA GAAGCAGCCA AAGCCTGCTG CATAGCAACG
CGGACGCTCG TAATCGACGT CTTCGTCGGT TTCGGACGAC GTATCGTTGC

G  Y  N  Y   L  D  W   Y  L  Q  K   P  G  Q   S  P  Q
                  KpnI          SexAI
                 ~~~~~~~        ~~~~~~~
GCTATAACTA TCTGGATTGG TACCTTCAAA AACCAGGTCA AAGCCCGCAG
CGATATTGAT AGACCTAACC ATGGAAGTTT TTGGTCCAGT TTCGGGCGTC

 L  L  I  Y   L  G  S   N  R  A   S  G  V  P   D  R  F
   AseI                              SanDI
  ~~~~~~                            ~~~~~~~~
CTATTAATTT ATCTGGGCAG CAACCGTGCC AGTGGGGTCC CGGATCGTTT
GATAATTAAA TAGACCCGTC GTTGGCACGG TCACCCCAGG GCCTAGCAAA

  S  G  S   G  S  G  T   D  F  T   L  K  I   S  R  V
            BamHI
           ~~~~~~
TAGCGGCTCT GGATCCGGCA CCGATTTTAC CCTGAAAATT AGCCGTGTGG
ATCGCCGAGA CCTAGGCCGT GGCTAAAATG GGACTTTTAA TCGGCACACC

E  A  E  D   V  G  V   Y  Y  C  Q   Q  H  Y   T  T  P
   Eco57I
   ~~~~~~
      BbsI
     ~~~~~~~
AAGCTGAAGA CGTGGGCGTG TATTATTGCC AGCAGCATTA TACCACCCCG
TTCGACTTCT GCACCCGCAC ATAATAACGG TCGTCGTAAT ATGGTGGGGC

 P  T  F  G   Q  G  T   K  V  E   I  K  R  T
        MscI                          BsiWI
       ~~~~~~~                        ~~~~~~
CCGACCTTTG GCCAGGGTAC GAAAGTTGAA ATTAAACGTA CG
GGCTGGAAAC CGGTCCCATG CTTTCAACTT TAATTTGCAT GC
```

Figure 3C: V kappa 3 (Vκ3) gene sequence

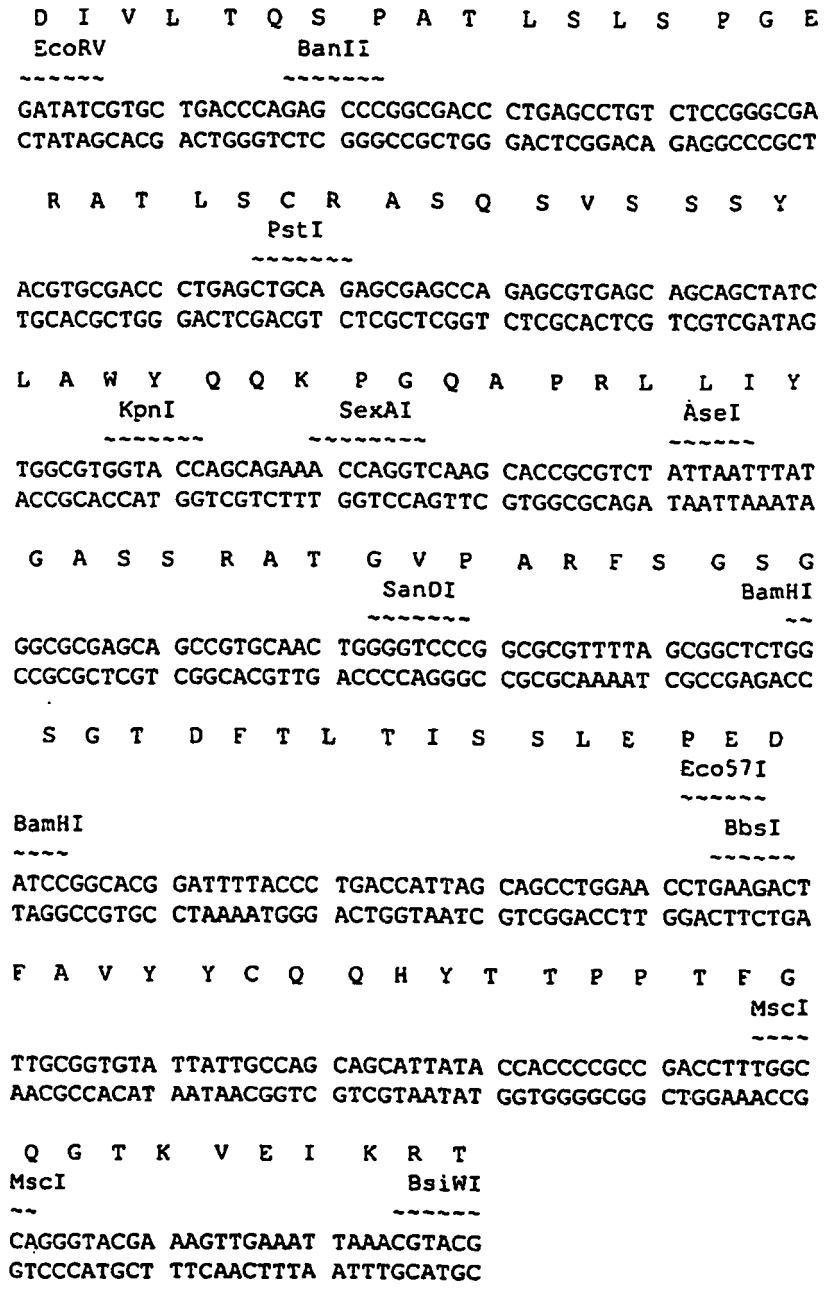


Figure 3D: V kappa 4 (Vκ4) gene sequence

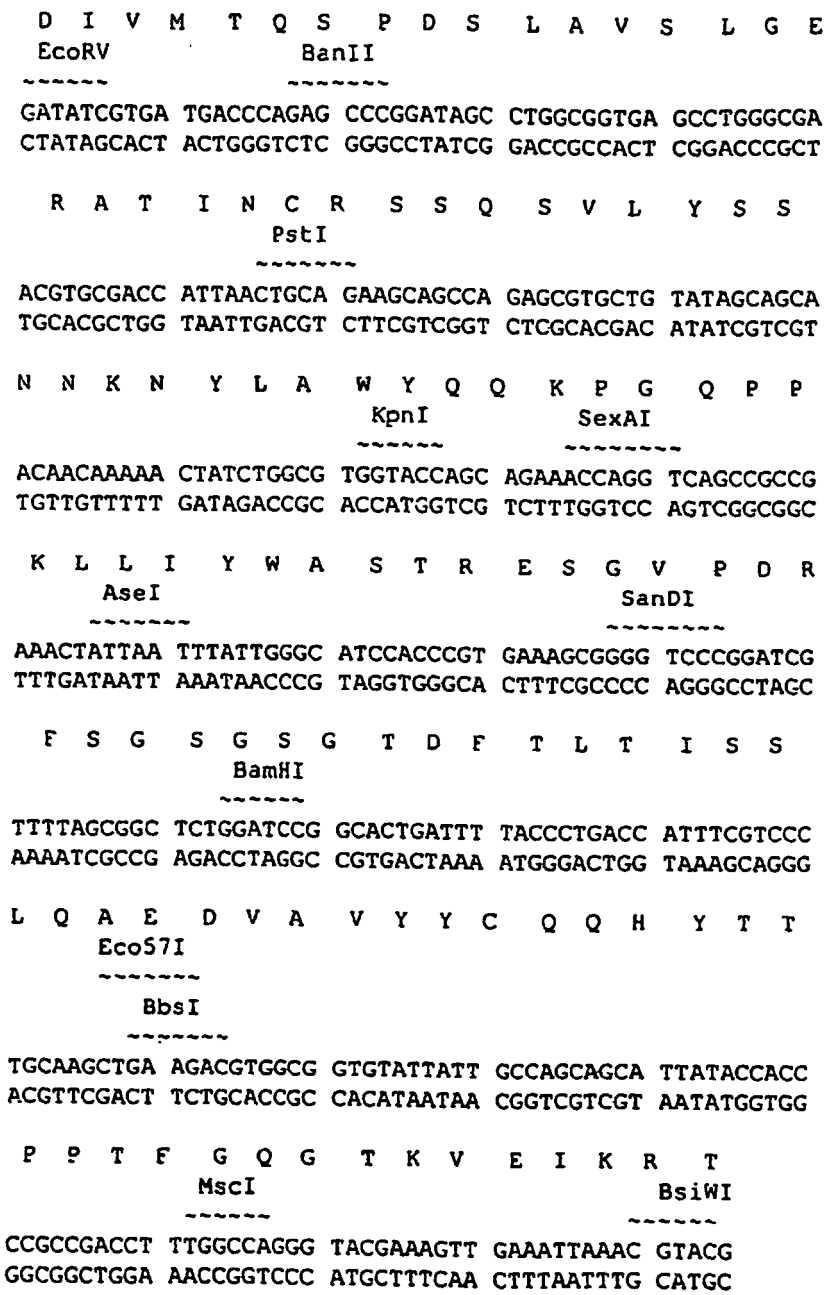


Figure 4A: V lambda 1 (Vλ1) gene sequence

```
 Q  S  V  L   T  Q  P   P  S  V   S  G  A  P   G  Q  R
                                              SexAI
                                             ~~~~~~~~
CAGAGCGTGC TGACCCAGCC GCCTTCAGTG AGTGGCGCAC CAGGTCAGCG
GTCTCGCACG ACTGGGTCGG CGGAAGTCAC TCACCGCGTG GTCCAGTCGC
                        Eco57I
                       ~~~~~~

  V  T  I   S  C  S  G   S  S  S   N  I  G   S  N  Y
         BssSI
       ~~~~~~~
TGTGACCATC TCGTGTAGCG GCAGCAGCAG CAACATTGGC AGCAACTATG
ACACTGGTAG AGCACATCGC CGTCGTCGTC GTTGTAACCG TCGTTGATAC

V  S  W  Y   Q  Q  L   P  G  T  A   P  K  L   L  I  Y
      KpnI            XmaI    BbeI
     ~~~~~~~         ~~~~~~  ~~~~~~~
TGAGCTGGTA CCAGCAGTTG CCCGGGACGG CGCCGAAACT GCTGATTTAT
ACTCGACCAT GGTCGTCAAC GGGCCCTGCC GCGGCTTTGA CGACTAAATA

 D  N  N  Q   R  P  S   G  V  P   D  R  F  S   G  S  K
               Bsu36I                          BamHI
              ~~~~~~~~                        ~~~~~~
GATAACAACC AGCGTCCCTC AGGCGTGCCG GATCGTTTTA GCGGATCCAA
CTATTGTTGG TCGCAGGGAG TCCGCACGGC CTAGCAAAAT CGCCTAGGTT

  S  G  T   S  A  S  L   A  I  T   G  L  Q   S  E  D
                                               BbsI
                                              ~~~~~~
AAGCGGCACC AGCGCGAGCC TTGCGATTAC GGGCCTGCAA AGCGAAGACG
TTCGCCGTGG TCGCGCTCGG AACGCTAATG CCCGGACGTT TCGCTTCTGC

E  A  D  Y   Y  C  Q   Q  H  Y  T   T  P  P   V  F  G
AAGCGGATTA TTATTGCCAG CAGCATTATA CCACCCCGCC TGTGTTTGGC
TTCGCCTAAT AATAACGGTC GTCGTAATAT GGTGGGGCGG ACACAAACCG

 G  G  T  K   L  T  V   L  G
             HpaI       MscI
            ~~~~~~     ~~~
GGCGGCACGA AGTTAACCGT TCTTGGC
CCGCCGTGCT TCAATTGGCA AGAACCG
```

Figure 4B: V lambda 2 (Vλ2) gene sequence

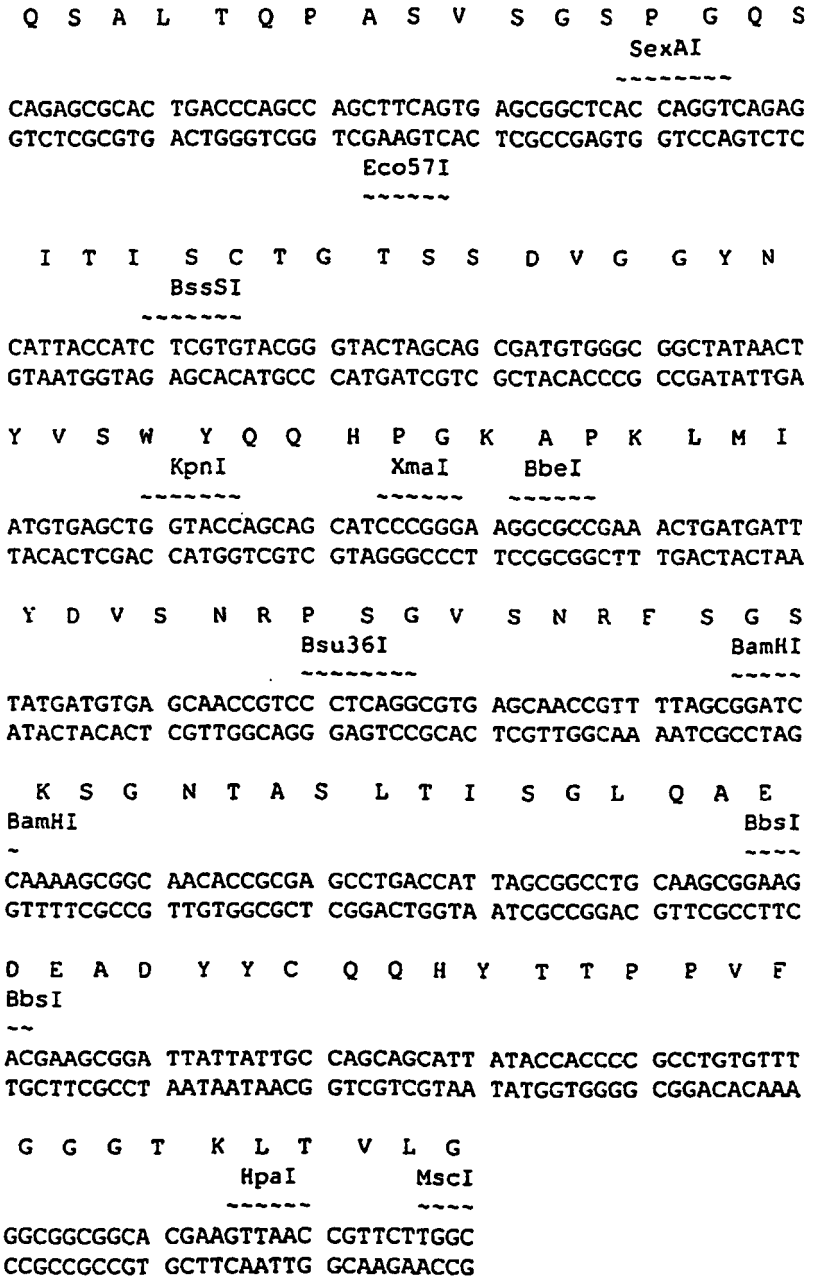


Figure 4C: V lambda 3 (Vλ3) gene sequence

```
 S  Y  E  L  T  Q  P   P  S  V   S  V  A  P   G  Q  T
                                          SexAI
                                         ~~~~~~~~
AGCTATGAAC TGACCCAGCC GCCTTCAGTG AGCGTTGCAC CAGGTCAGAC
TCGATACTTG ACTGGGTCGG CGGAAGTCAC TCGCAACGTG GTCCAGTCTG
                      Eco57I
                      ~~~~~~

  A  R  I   S  C  S  G   D  A  L   G  D  K   Y  A  S
          BssSI
        ~~~~~~~
CGCGCGTATC TCGTGTAGCG GCGATGCGCT GGGCGATAAA TACGCGAGCT
GCGCGCATAG AGCACATCGC CGCTACGCGA CCCGCTATTT ATGCGCTCGA

W  Y  Q  Q   K  P  G   Q  A  P  V   L  V  I   Y  D  D
 KpnI          XmaI     BbeI
~~~~~~        ~~~~~~   ~~~~~~
GGTACCAGCA GAAACCCGGG CAGGCGCCAG TTCTGGTGAT TTATGATGAT
CCATGGTCGT CTTTGGGCCC GTCCGCGGTC AAGACCACTA AATACTACTA

 S  D  R  P   S  G  I   P  E  R   F  S  G  S   N  S  G
          Bsu36I                      BamHI
          ~~~~~~~                    ~~~~~~~
TCTGACCGTC CCTCAGGCAT CCCGGAACGC TTTAGCGGAT CCAACAGCGG
AGACTGGCAG GGAGTCCGTA GGGCCTTGCG AAATCGCCTA GGTTGTCGCC

  N  T  A   T  L  T  I   S  G  T   Q  A  E   D  E  A
                                        BbsI
                                       ~~~~~~~
CAACACCGCG ACCCTGACCA TTAGCGGCAC TCAGGCGGAA GACGAAGCGG
GTTGTGGCGC TGGGACTGGT AATCGCCGTG AGTCCGCCTT CTGCTTCGCC

D  Y  Y  C   Q  Q  H   Y  T  T  P   P  V  F   G  G  G
ATTATTATTG CCAGCAGCAT TATACCACCC CGCCTGTGTT TGGCGGCGGC
TAATAATAAC GGTCGTCGTA ATATGGTGGG GCGGACACAA ACCGCCGCCG

 T  K  L  T   V  L  G
     HpaI       MscI
    ~~~~~~~     ~~~~
ACGAAGTTAA CCGTTCTTGG C
TGCTTCAATT GGCAAGAACC G
```

Figure 5A: V heavy chain 1A (VH1A) gene sequence

```
 Q  V  Q  L   V  Q  S   G  A  E   V  K  K  P   G  S  S
       MfeI
      -------
CAGGTGCAAT TGGTTCAGTC TGGCGCGGAA GTGAAAAAAC CGGGCAGCAG
GTCCACGTTA ACCAAGTCAG ACCGCGCCTT CACTTTTTTG GCCCGTCGTC

  V  K  V   S  C  K  A   S  G  G   T  F  S   S  Y  A
                        BspEI
                       ------
CGTGAAAGTG AGCTGCAAAG CCTCCGGAGG CACTTTTAGC AGCTATGCGA
GCACTTTCAC TCGACGTTTC GGAGGCCTCC GTGAAAATCG TCGATACGCT

I  S  W  V   R  Q  A   P  G  Q  G   L  E  W   M  G  G
               BstXI              XhoI
             ~~~~~~~~~~~~~       ------
TTAGCTGGGT GCGCCAAGCC CCTGGGCAGG GTCTCGAGTG GATGGGCGGC
AATCGACCCA CGCGGTTCGG GGACCCGTCC CAGAGCTCAC CTACCCGCCG

 I  I  P  I   F  G  T   A  N  Y   A  Q  K  F   Q  G  R
ATTATTCCGA TTTTTGGCAC GGCGAACTAC GCGCAGAAGT TTCAGGGCCG
TAATAAGGCT AAAAACCGTG CCGCTTGATG CGCGTCTTCA AAGTCCCGGC

  V  T  I   T  A  D  E   S  T  S   T  A  Y   M  E  L
BstEII
-------
GGTGACCATT ACCGCGGATG AAAGCACCAG CACCGCGTAT ATGGAACTGA
CCACTGGTAA TGGCGCCTAC TTTCGTGGTC GTGGCGCATA TACCTTGACT

S  S  L  R   S  E  D   T  A  V  Y   Y  C  A   R  W  G
                      EagI             BssHII
                     ------           -------
GCAGCCTGCG TAGCGAAGAT ACGGCCGTGT ATTATTGCGC GCGTTGGGGC
CGTCGGACGC ATCGCTTCTA TGCCGGCACA TAATAACGCG CGCAACCCCG

 G  D  G  F   Y  A  M   D  Y  W   G  Q  G  T   L  V  T
                                   StyI
                                  ------
GGCGATGGCT TTTATGCGAT GGATTATTGG GGCCAAGGCA CCCTGGTGAC
CCGCTACCGA AAATACGCTA CCTAATAACC CCGGTTCCGT GGGACCACTG

  V  S  S
     BlpI
    ------
GGTTAGCTCA G
CCAATCGAGT C
```

Figure 5B: V heavy chain 1B (VH1B) gene sequence

```
 Q  V  Q  L   V  Q  S   G  A  E   V  K  K  P   G  A  S
       MfeI
      ~~~~~~~
CAGGTGCAAT TGGTTCAGAG CGGCGCGGAA GTGAAAAAAC CGGGCGCGAG
GTCCACGTTA ACCAAGTCTC GCCGCGCCTT CACTTTTTTG GCCCGCGCTC

  V  K  V   S  C  K  A   S  G  Y   T  F  T   S  Y  Y
                        BspEI
                       ~~~~~~
CGTGAAAGTG AGCTGCAAAG CCTCCGGATA TACCTTTACC AGCTATTATA
GCACTTTCAC TCGACGTTTC GGAGGCCTAT ATGGAAATGG TCGATAATAT

M  H  W  V   R  Q  A   P  G  Q  G   L  E  W   M  G  W
                BstXI              XhoI
              ~~~~~~~~~~~~~       ~~~~~~
TGCACTGGGT CCGCCAAGCC CCTGGGCAGG GTCTCGAGTG GATGGGCTGG
ACGTGACCCA GGCGGTTCGG GGACCCGTCC CAGAGCTCAC CTACCCGACC

 I  N  P  N   S  G  G   T  N  Y   A  Q  K  F   Q  G  R
ATTAACCCGA ATAGCGGCGG CACGAACTAC GCGCAGAAGT TTCAGGGCCG
TAATTGGGCT TATCGCCGCC GTGCTTGATG CGCGTCTTCA AAGTCCCGGC

  V  T  M   T  R  D  T   S  I  S   T  A  Y   M  E  L
BstEII
~~~~~~~
GGTGACCATG ACCCGTGATA CCAGCATTAG CACCGCGTAT ATGGAACTGA
CCACTGGTAC TGGGCACTAT GGTCGTAATC GTGGCGCATA TACCTTGACT

S  S  L  R   S  E  D   T  A  V  Y   Y  C  A   R  W  G
                      EagI            BssHII
                     ~~~~~~          ~~~~~~~
GCAGCCTGCG TAGCGAAGAT ACGGCCGTGT ATTATTGCGC GCGTTGGGGC
CGTCGGACGC ATCGCTTCTA TGCCGGCACA TAATAACGCG CGCAACCCCG

 G  D  G  F   Y  A  M   D  Y  W   G  Q  G  T   L  V  T
                                   StyI
                                  ~~~~~~
GGCGATGGCT TTTATGCGAT GGATTATTGG GGCCAAGGCA CCCTGGTGAC
CCGCTACCGA AAATACGCTA CCTAATAACC CCGGTTCCGT GGGACCACTG

  V  S  S
     BlpI
    ~~~~~~
GGTTAGCTCA G
CCAATCGAGT C
```

Figure 5C: V heavy chain 2 (VH2) gene sequence

```
 Q  V  Q  L   K  E  S   G  P  A   L  V  K  P   T  Q  T
       MfeI
      ~~~~~~~
CAGGTGCAAT TGAAAGAAAG CGGCCCGGCC CTGGTGAAAC CGACCCAAAC
GTCCACGTTA ACTTTCTTTC GCCGGGCCGG GACCACTTTG GCTGGGTTTG

  L  T  L   T  C  T  F   S  G  F   S  L  S   T  S  G
                        BspEI
                       ~~~~~~
CCTGACCCTG ACCTGTACCT TTTCCGGATT TAGCCTGTCC ACGTCTGGCG
GGACTGGGAC TGGACATGGA AAAGGCCTAA ATCGGACAGG TGCAGACCGC

V  G  V  G   W  I  R   Q  P  P  G   K  A  L   E  W  L
                        BstXI           XhoI
                    ~~~~~~~~~~~~~~      ~~~~~~~
TTGGCGTGGG CTGGATTCGC CAGCCGCCTG GGAAAGCCCT CGAGTGGCTG
AACCGCACCC GACCTAAGCG GTCGGCGGAC CCTTTCGGGA GCTCACCGAC

 A  L  I  D   W  D  D   D  K  Y   Y  S  T  S   L  K  T
                                                 MluI
                                                  ~~
GCTCTGATTG ATTGGGATGA TGATAAGTAT TATAGCACCA GCCTGAAAAC
CGAGACTAAC TAACCCTACT ACTATTCATA ATATCGTGGT CGGACTTTTG

  R  L  T   I  S  K  D   T  S  K   N  Q  V   V  L  T
MluI                     NspV
~~~~                     ~~~~~~
GCGTCTGACC ATTAGCAAAG ATACTTCGAA AAATCAGGTG GTGCTGACTA
CGCAGACTGG TAATCGTTTC TATGAAGCTT TTTAGTCCAC CACGACTGAT

M  T  N  M   D  P  V   D  T  A  T   Y  Y  C   A  R  W
                                        BssHII
                                       ~~~~~~~
TGACCAACAT GGACCCGGTG GATACGGCCA CCTATTATTG CGCGCGTTGG
ACTGGTTGTA CCTGGGCCAC CTATGCCGGT GGATAATAAC GCGCGCAACC

 G  G  D  G   F  Y  A   M  D  Y   W  G  Q  G   T  L  V
                                     StyI
                                    ~~~~~~~
GGCGGCGATG GCTTTTATGC GATGGATTAT TGGGGCCAAG GCACCCTGGT
CCGCCGCTAC CGAAAATACG CTACCTAATA ACCCCGGTTC CGTGGGACCA

  T  V  S   S
        BlpI
       ~~~~~~
GACGGTTAGC TCAG
CTGCCAATCG AGTC
```

Figure 5D: V heavy chain 3 (VH3) gene sequence

```
 E  V  Q  L   V  E  S   G  G  G   L  V  Q  P   G  G  S
       MfeI
      ~~~~~~~
GAAGTGCAAT TGGTGGAAAG CGGCGGCGGC CTGGTGCAAC CGGGCGGCAG
CTTCACGTTA ACCACCTTTC GCCGCCGCCG GACCACGTTG GCCCGCCGTC

  L  R  L   S  C  A  A   S  G  F   T  F  S   S  Y  A
                        BspEI
                       ~~~~~~
CCTGCGTCTG AGCTGCGCGG CCTCCGGATT TACCTTTAGC AGCTATGCGA
GGACGCAGAC TCGACGCGCC GGAGGCCTAA ATGGAAATCG TCGATACGCT

M  S  W  V   R  Q  A   P  G  K  G   L  E  W   V  S  A
               BstXI                 XhoI
             ~~~~~~~~~~~~~          ~~~~~~
TGAGCTGGGT GCGCCAAGCC CCTGGGAAGG GTCTCGAGTG GGTGAGCGCG
ACTCGACCCA CGCGGTTCGG GGACCCTTCC CAGAGCTCAC CCACTCGCGC

 I  S  G  S   G  G  S   T  Y  Y   A  D  S  V   K  G  R
ATTAGCGGTA GCGGCGGCAG CACCTATTAT GCGGATAGCG TGAAAGGCCG
TAATCGCCAT CGCCGCCGTC GTGGATAATA CGCCTATCGC ACTTTCCGGC

  F  T  I   S  R  D  N   S  K  N   T  L  Y   L  Q  M
            PmlI       NspV
           ~~~~~~     ~~~~~~
TTTTACCATT TCACGTGATA ATTCGAAAAA CACCCTGTAT CTGCAAATGA
AAAATGGTAA AGTGCACTAT TAAGCTTTTT GTGGGACATA GACGTTTACT

N  S  L  R   A  E  D   T  A  V  Y   Y  C  A   R  W  G
                      EagI           BssHII
                     ~~~~~~         ~~~~~~~
ACAGCCTGCG TGCGGAAGAT ACGGCCGTGT ATTATTGCGC GCGTTGGGGC
TGTCGGACGC ACGCCTTCTA TGCCGGCACA TAATAACGCG CGCAACCCCG

 G  D  G  F   Y  A  M   D  Y  W   G  Q  G  T   L  V  T
                                   StyI
                                  ~~~~~~
GGCGATGGCT TTTATGCGAT GGATTATTGG GGCCAAGGCA CCCTGGTGAC
CCGCTACCGA AAATACGCTA CCTAATAACC CCGGTTCCGT GGGACCACTG

  V  S  S
     BlpI
    ~~~~~~
GGTTAGCTCA G
CCAATCGAGT C
```

Figure 5E: V heavy chain 4 (VH4) gene sequence

```
 Q  V  Q  L   Q  E  S   G  P  G   L  V  K  P   S  E  T
       MfeI
      ~~~~~~~
CAGGTGCAAT TGCAAGAAAG TGGTCCGGGC CTGGTGAAAC CGAGCGAAAC
GTCCACGTTA ACGTTCTTTC ACCAGGCCCG GACCACTTTG GCTCGCTTTG

  L  S  L   T  C  T  V   S  G  G   S  I  S   S  Y  Y
                         BspEI
                        ~~~~~~
CCTGAGCCTG ACCTGCACCG TTTCCGGAGG CAGCATTAGC AGCTATTATT
GGACTCGGAC TGGACGTGGC AAAGGCCTCC GTCGTAATCG TCGATAATAA

W  S  W  I   R  Q  P   P  G  K  G   L  E  W   I  G  Y
               BstXI                 XhoI
             ~~~~~~~~~~~~~         ~~~~~~
GGAGCTGGAT TCGCCAGCCG CCTGGGAAGG GTCTCGAGTG GATTGGCTAT
CCTCGACCTA AGCGGTCGGC GGACCCTTCC CAGAGCTCAC CTAACCGATA

 I  Y  Y  S   G  S  T   N  Y  N   P  S  L  K   S  R  V
                                                BstEII
                                                  ~~~
ATTTATTATA GCGGCAGCAC CAACTATAAT CCGAGCCTGA AAAGCCGGGT
TAAATAATAT CGCCGTCGTG GTTGATATTA GGCTCGGACT TTTCGGCCCA

  T  I  S   V  D  T  S   K  N  Q   F  S  L   K  L  S
BstEII          NspV
~~~~            ~~~~~~~
GACCATTAGC GTTGATACTT CGAAAAACCA GTTTAGCCTG AAACTGAGCA
CTGGTAATCG CAACTATGAA GCTTTTTGGT CAAATCGGAC TTTGACTCGT

S  V  T  A   A  D  T   A  V  Y  Y   C  A  R   W  G  G
                EagI                 BssHII
              ~~~~~~~               ~~~~~~
GCGTGACGGC GGCGGATACG GCCGTGTATT ATTGCGCGCG TTGGGGCGGC
CGCACTGCCG CCGCCTATGC CGGCACATAA TAACGCGCGC AACCCCGCCG

 D  G  F  Y   A  M  D   Y  W  G   Q  G  T  L   V  T  V
                                  StyI
                                ~~~~~~~
GATGGCTTTT ATGCGATGGA TTATTGGGGC CAAGGCACCC TGGTGACGGT
CTACCGAAAA TACGCTACCT AATAACCCCG GTTCCGTGGG ACCACTGCCA

  S  S
   BlpI
  ~~~~~
TAGCTCAG
ATCGAGTC
```

Figure 5F: V heavy chain 5 (VH5) gene sequence

```
 E  V  Q  L   V  Q  S   G  A  E   V  K  K  P   G  E  S
       MfeI
      ~~~~~~~
GAAGTGCAAT TGGTTCAGAG CGGCGCGGAA GTGAAAAAAC CGGGCGAAAG
CTTCACGTTA ACCAAGTCTC GCCGCGCCTT CACTTTTTTG GCCCGCTTTC

  L  K  I   S  C  K  G   S  G  Y   S  F  T   S  Y  W
                        BspEI
                       ~~~~~~
CCTGAAAATT AGCTGCAAAG GTTCCGGATA TTCCTTTACG AGCTATTGGA
GGACTTTTAA TCGACGTTTC CAAGGCCTAT AAGGAAATGC TCGATAACCT

I  G  W  V   R  Q  M   P  G  K  G   L  E  W   M  G  I
               BstXI              XhoI
             ~~~~~~~~~~~~~       ~~~~~~
TTGGCTGGGT GCGCCAGATG CCTGGGAAGG GTCTCGAGTG GATGGGCATT
AACCGACCCA CGCGGTCTAC GGACCCTTCC CAGAGCTCAC CTACCCGTAA

 I  Y  P  G   D  S  D   T  R  Y   S  P  S  F   Q  G  Q
ATTTATCCGG GCGATAGCGA TACCCGTTAT TCTCCGAGCT TTCAGGGCCA
TAAATAGGCC CGCTATCGCT ATGGGCAATA AGAGGCTCGA AAGTCCCGGT

  V  T  I   S  A  D  K   S  I  S   T  A  Y   L  Q  W
BstEII
~~~~~~~
GGTGACCATT AGCGCGGATA AAAGCATTAG CACCGCGTAT CTTCAATGGA
CCACTGGTAA TCGCGCCTAT TTTCGTAATC GTGGCGCATA GAAGTTACCT

S  S  L  K   A  S  D   T  A  M  Y   Y  C  A   R  W  G
                                      BssHII
                                     ~~~~~~~
GCAGCCTGAA AGCGAGCGAT ACGGCCATGT ATTATTGCGC GCGTTGGGGC
CGTCGGACTT TCGCTCGCTA TGCCGGTACA TAATAACGCG CGCAACCCCG

 G  D  G  F   Y  A  M   D  Y  W   G  Q  G  T   L  V  T
                                    StyI
                                   ~~~~~~
GGCGATGGCT TTTATGCGAT GGATTATTGG GGCCAAGGCA CCCTGGTGAC
CCGCTACCGA AAATACGCTA CCTAATAACC CCGGTTCCGT GGGACCACTG

  V  S  S
     BlpI
    ~~~~~~
GGTTAGCTCA G
CCAATCGAGT C
```

Figure 5G: V heavy chain 6 (VH6) gene sequence

```
 Q  V  Q  L  Q  Q  S  G  P  G  L  V  K  P  S  Q  T
      MfeI
      ~~~~~~~
CAGGTGCAAT TGCAACAGTC TGGTCCGGGC CTGGTGAAAC CGAGCCAAAC
GTCCACGTTA ACGTTGTCAG ACCAGGCCCG GACCACTTTG GCTCGGTTTG

  L  S  L  T  C  A  I  S  G  D  S  V  S  S  N  S
                      BspEI
                     ~~~~~~
CCTGAGCCTG ACCTGTGCGA TTTCCGGAGA TAGCGTGAGC AGCAACAGCG
GGACTCGGAC TGGACACGCT AAAGGCCTCT ATCGCACTCG TCGTTGTCGC

A  A  W  N  W  I  R  Q  S  P  G  R  G  L  E  W  L
                      BstXI          XhoI
                  ~~~~~~~~~~~~~~    ~~~~~~~
CGGCGTGGAA CTGGATTCGC CAGTCTCCTG GGCGTGGCCT CGAGTGGCTG
GCCGCACCTT GACCTAAGCG GTCAGAGGAC CCGCACCGGA GCTCACCGAC

 G  R  T  Y  Y  R  S  K  W  Y  N  D  Y  A  V  S  V
GGCCGTACCT ATTATCGTAG CAAATGGTAT AACGATTATG CGGTGAGCGT
CCGGCATGGA TAATAGCATC GTTTACCATA TTGCTAATAC GCCACTCGCA

  K  S  R  I  T  I  N  P  D  T  S  K  N  Q  F  S
           BsaBI              NspV
        ~~~~~~~~~~~         ~~~~~~
GAAAAGCCGG ATTACCATCA ACCCGGATAC TTCGAAAAAC CAGTTTAGCC
CTTTTCGGCC TAATGGTAGT TGGGCCTATG AAGCTTTTTG GTCAAATCGG

L  Q  L  N  S  V  T  P  E  D  T  A  V  Y  Y  C  A
                                 EagI         BssHII
                                ~~~~~~        ~~~~~
TGCAACTGAA CAGCGTGACC CCGGAAGATA CGGCCGTGTA TTATTGCGCG
ACGTTGACTT GTCGCACTGG GGCCTTCTAT GCCGGCACAT AATAACGCGC

 R  W  G  G  D  G  F  Y  A  M  D  Y  W  G  Q  G  T
BssHII                                      StyI
~                                          ~~~~~~
CGTTGGGGCG GCGATGGCTT TTATGCGATG GATTATTGGG GCCAAGGCAC
GCAACCCCGC CGCTACCGAA AATACGCTAC CTAATAACCC CGGTTCCGTG

  L  V  T  V  S  S
              BlpI
             ~~~~~~
CCTGGTGACG GTTAGCTCAG
GGACCACTGC CAATCGAGTC
```

Figure 6: oligonucleotides for gene synthesis

O1K1	5'- GAATGCATACGCTGATATCCAGATGACCCAGAGCCCGTCTAGCCTGAGC -3'
O1K2	5'- CGCTCTGCAGGTAATGGTCACACGATCACCCACGCTCGCGCTCAGGCTAGACGGGC -3'
O1K3	5'- GACCATTACCTGCAGAGCGAGCCAGGGCATTAGCAGCTATCTGGCGTGGTACCAGCAG -3'
O1K4	5'- CTTTGCAAGCTGCTGGCTGCATAAATTAATAGTTTCGGTGCTTTACCTGGTTTCTGCTGGTACCACGCCAG -3'
O1K5	5'- CAGCCAGCAGCTTGCAAAGCGGGGTCCCGTCCCGTTTTAGCGGCTCTGGATCCGGCACTGATTTTAC -3'
O1K6	5'- GATAATAGGTCGCAAAGTCTTCAGGTTGCAGGCTGCTAATGGTCAGGGTAAAATCAGTGCCGGATCC -3'
O2K1	5'- CGATATCGTGATGACCCAGAGCCCACTGAGCCTGCCAGTGACTCCGGGCGAGCC -3'
O2K2	5'- GCCGTTGCTATGCAGCAGGCTTTGGCTGCTTCTGCAGCTAATGCTCGCAGGCTCGCCCGGAGTCAC -3'
O2K3	5'- CTGCTGCATAGCAACGGCTATAACTATCTGGATTGGTACCTTCAAAAACCAGGTCAAAGCCC -3'
O2K4	5'- CGATCCGGGACCCCACTGGCACGGTTGCTGCCCAGATAAATTAATAGCTGCGGGCTTTGACCTGGTTTTTG -3'
O2K5	5'- AGTGGGGTCCCGGATCGTTTTAGCGGCTCTGGATCCGGCACCGATTTTACCCTGAAAATTAGCCGTGTG -3'
O2K6	5'- CCATGCAATAATACACGCCCACGTCTTCAGCTTCCACACGGCTAATTTTCAGGG -3'
O3K1	5'- GAATGCATACGCTGATATCGTGCTGACCCAGAGCCCGG -3'
O3K2	5'- CGCTCTGCAGCTCAGGGTCGCACGTTCGCCCGGAGACAGGCTCAGGGTCGCCGGGCTCTGGGTCAGC -3'
O3K3	5'- CCCTGAGCTGCAGAGCGAGCCAGAGCGTGAGCAGCAGCTATCTGGCGTGGTACCAG -3'
O3K4	5'- GCACGGCTGCTCGCGCCATAAATTAATAGACGCGGTGCTTGACCTGGTTTCTGCTGGTACCACGCCAGATAG -3'
O3K5	5'- GCGCGAGCAGCCGTGCAACTGGGGTCCCGGCGCGTTTTAGCGGCTCTGGATCCGGCACGGATTTTAC -3'
O3K6	5'- GATAATACACCGCAAAGTCTTCAGGTTCCAGGCTGCTAATGGTCAGGGTAAAATCCGTGCCGGATC -3'
O4K1	5'- GAATGCATACGCTGATATCGTGATGACCCAGAGCCCGGATAGCCTGGCG -3'
O4K2	5'- GCTTCTGCAGTTAATGGTCGCACGTTCGCCCAGGCTCACCGCCAGGCTATCCGGGC -3'
O4K3	5'- CGACCATTAACTGCAGAAGCAGCCAGAGCGTGCTGTATAGCAGCAACAACAAAAACTATCTGGCGTGGTACCAG -3'
O4K4	5'- GATGCCCAATAAATTAATAGTTTCGGCGGCTGACCTGGTTTCTGCTGGTACCACGCCAGATAG -3'
O4K5	5'- AAACTATTAATTTATTGGGCATCCACCCGTGAAAGCGGGGTCCCGGATCGTTTTAGCGGCTCTGGATCCGGCAC -3'
O4K6	5'- GATAATACACCGCCACGTCTTCAGCTTGCAGGGACGAAATGGTCAGGGTAAAATCAGTGCCGGATCCAGAGCC -3'
O1L1	5'- GAATGCATACGCTCAGAGCGTGCTGACCCAGCCGCCTTCAGTGAGTGG -3'
O1L2	5'- CAATGTTGCTGCTGCTGCCGCTACACGAGATGGTCACACGCTGACCTGGTGCGCCACTCACTGAAGGCGGC -3'
O1L3	5'- GGCAGCAGCAGCAACATTGGCAGCAACTATGTGAGCTGGTACCAGCAGTTGCCCGGGAC -3'
O1L4	5'- CCGGCACGCCTGAGGGACGCTGGTTGTTATCATAAATCAGCAGTTTCGGCGCCGTCCCGGGCAACTGC -3'
O1L5	5'- CCCTCAGGCGTGCCGGATCGTTTTAGCGGATCCAAAAGCGGCACCAGCGCGAGCCTTGCG -3'
O1L6	5'- CCGCTTCGTCTTCGCTTTGCAGGCCCGTAATCGCAAGGCTCGCGCTGG -3'
O2L1	5'- GAATGCATACGCTCAGAGCGCACTGACCCAGCCAGCTTCAGTGAGCGGC -3'
O2L2	5'- CGCTGCTAGTACCCGTACACGAGATGGTAATGCTCTGACCTGGTGAGCCGCTCACTGAAGCTGG -3'
O2L3	5'- GTACGGGTACTAGCAGCGATGTGGGCGGCTATAACTATGTGAGCTGGTACCAGCAGCATCCCGG -3'
O2L4	5'- CGCCTGAGGGACGGTTGCTCACATCATAAATCATCAGTTTCGGCGCCTTCCCGGGATGCTGCTGGTAC -3'
O2L5	5'- CAACCGTCCCTCAGGCGTGAGCAACCGTTTTAGCGGATCCAAAAGCGGCAACACCGCGAGCC -3'
O2L6	5'- CCGCTTCGTCTTCCGCTTGCAGGCCGCTAATGGTCAGGCTCGCGGTGTTGCCG -3'
O3L1	5'- GAATGCATACGCTAGCTATGAACTGACCCAGCCGCCTTCAGTGAGCG -3'
O3L2	5'- CGCCCAGCGCATCGCCGCTACACGAGATACGCGCGGTCTGACCTGGTGCAACGCTCACTGAAGGCGGC -3'
O3L3	5'- GGCGATGCGCTGGGCGATAAATACGCGAGCTGGTACCAGCAGAAACCCGGGCAGGCGC -3'
O3L4	5'- GCGTTCCGGGATGCCTGAGGGACGGTCAGAATCATCATAAATCACCAGAACTGGCGCCTGCCCGGGTTTC -3'
O3L5	5'- CAGGCATCCCGGAACGCTTTAGCGGATCCAACAGCGGCAACACCGCGACCCTGACCATTAGCGG -3'
O3L6	5'- CCGCTTCGTCTTCCGCCTGAGTGCCGCTAATGGTCAGGGTC -3'
O1246H1	5'- GCTCTTCACCCCTGTTACCAAAGCCCAGGTGCAATTG -3'
O1AH2	5'- GGCTTTGCAGCTCACTTTCACGCTGCTGCCCGGTTTTTTCACTTCCGCGCCAGACTGAACCAATTGCACCTGGGCTTTG -3'

Figure 6: (continued)

```
O1AH3    5'- GAAAGTGAGCTGCAAAGCCTCCGGAGGCACTTTTAGCAGCTATGCGATTAGCTGGGTGCGCCAAGCCCCTGGGCAGGGTC -3'
O1AH4    5'- GCCCTGAAACTTCTGCGCGTAGTTCGCCGTGCCAAAAATCGGAATAATGCCGCCCATCCACTCGAGACCCTGCCCAGGGGC -3'
O1AH5    5'- GCGCAGAAGTTTCAGGGCCGGGTGACCATTACCGCGGATGAAAGCACCAGCACCGCGTATATGGAACTGAGCAGCCTGCG -3'
O1ABH6   5'- GCGCGCAATAATACACGGCCGTATCTTCGCTACGCAGGCTGCTCAGTTCC -3'
O1BH2    5'- GGCTTTGCAGCTCACTTTCACGCTCGCGCCCGGTTTTTTCACTTCCGCGCCGCTCTGAACCAATTGCACCTGGGCTTTG -3'
O1BH3    5'- GAAAGTGAGCTGCAAAGCCTCCGGATATACCTTTACCAGCTATTATATGCACTGGGTCCGCCAAGCCCCTGGGCAGGGTC -3'
O1BH4    5'- GCCCTGAAACTTCTGCGCGTAGTTCGTGCCGCCGCTATTCGGGTTAATCCAGCCCATCCACTCGAGACCCTGCCCAGGGGC -3'
O1BH5    5'- GCGCAGAAGTTTCAGGGCCGGGTGACCATGACCCGTGATACCAGCATTAGCACCGCGTATATGGAACTGAGCAGCCTGCG -3'
O2H2     5'- GGTACAGGTCAGGGTCAGGGTTTGGGTCGGTTTCACCAGGGCCGGGCCGCTTTCTTTCAATTGCACCTGGGCTTTG -3'
O2H3     5'- CTGACCCTGACCTGTACCTTTTCCGGATTTAGCCTGTCCACGTCTGGCGTTGGCGTGGGCTGGATTCGCCAGCCGCCTGGGAAAG -3'
O2H4     5'- GCGTTTTCAGGCTGGTGCTATAATACTTATCATCATCCCAATCAATCAGAGCCAGCCACTCGAGGGCTTTCCCAGGCGGCTGG -3'
O2H5     5'- GCACCAGCCTGAAAACGCGTCTGACCATTAGCAAAGATACTTCGAAAAATCAGGTGGTGCTGACTATGACCAACATGG -3'
O2H6     5'- GCGCGCAATAATAGGTGGCCGTATCCACCGGGTCCATGTTGGTCATAGTCAGC -3'
O3H1     5'- CGAAGTGCAATTGGTGGAAAGCGGCGGCGGCCTGGTGCAACCGGGCGGCAG -3'
O3H2     5'- CATAGCTGCTAAAGGTAAATCCGGAGGCCGCGCAGCTCAGACGCAGGCTGCCGCCCGGTTGCAC -3'
O3H3     5'- GATTTACCTTTAGCAGCTATGCGATGAGCTGGGTGCGCCAAGCCCCTGGGAAGGGTCTCGAGTGGGTGAG -3'
O3H4     5'- GGCCTTTCACGCTATCCGCATAATAGGTGCTGCCGCCGCTACCGCTAATCGCGCTCACCCACTCGAGACCC -3'
O3H5     5'- CGGATAGCGTGAAAGGCCGTTTTACCATTTCACGTGATAATTCGAAAAACACCCTGTATCTGCAAATGAACAG -3'
O3H6     5'- CACGCGCGCAATAATACACGGCCGTATCTTCCGCACGCAGGCTGTTCATTTGCAGATACAGG -3'
O4H2     5'- GGTCAGGCTCAGGGTTTCGCTCGGTTTCACCAGGCCCGGACCACTTTCTTGCAATTGCACCTGGGCTTTG -3'
O4H3     5'- GAAACCCTGAGCCTGACCTGCACCGTTTCCGGAGGCAGCATTAGCAGCTATTATTGGAGCTGGATTCGCCAGCCGC -3'
O4H4     5'- GATTATAGTTGGTGCTGCCGCTATAATAAATATAGCCAATCCACTCGAGACCCTTCCCAGGCGGCTGGCGAATCCAG -3'
O4H5     5'- CGGCAGCACCAACTATAATCCGAGCCTGAAAAGCCGGGTGACCATTAGCGTTGATACTTCGAAAAACCAGTTTAGCCTG -3'
O4H6     5'- GCGCGCAATAATACACGGCCGTATCCGCCGCCGTCACGCTGCTCAGTTTCAGGCTAAACTGGTTTTTCG -3'
O5H1     5'- GCTCTTCACCCCTGTTACCAAAGCCGAAGTGCAATTG -3'
O5H2     5'- CCTTTGCAGCTAATTTTCAGGCTTTCGCCCGGTTTTTTCACTTCCGCGCCGCTCTGAACCAATTGCACTTCGGCTTTGG -3'
O5H3     5'- CCTGAAAATTAGCTGCAAAGGTTCCGGATATTCCTTTACGAGCTATTGGATTGGCTGGGTGCGCCAGATGCCTGG -3'
O5H4     5'- CGGAGAATAACGGGTATCGCTATCGCCCGGATAAATAATGCCCATCCACTCGAGACCCTTCCCAGGCATCTGGCGCAC -3'
O5H5     5'- CGATACCCGTTATTCTCCGAGCTTTCAGGGCCAGGTGACCATTAGCGCGGATAAAAGCATTAGCACCGCGTATCTTC -3'
O5H6     5'- GCGCGCAATAATACATGGCCGTATCGCTCGCTTTCAGGCTGCTCCATTGAAGATACGCGGTGCTAATG -3'
O6H2     5'- GAAATCGCACAGGTCAGGCTCAGGGTTTGGCTCGGTTTCACCAGGCCCGGACCAGACTGTTGCAATTGCACCTGGGCTTTG -3'
O6H3     5'- GCCTGACCTGTGCGATTTCCGGAGATAGCGTGAGCAGCAACAGCGCGGCGTGGAACTGGATTCGCCAGTCTCCTGGGCG -3'
O6H4     5'- CACCGCATAATCGTTATACCATTTGCTACGATAATAGGTACGGCCCAGCCACTCGAGGCCACGCCCAGGAGACTGGCG -3'
O6H5     5'- GGTATAACGATTATGCGGTGAGCGTGAAAAGCCGGATTACCATCAACCCGGATACTTCGAAAAACCAGTTTAGCCTGC -3'
O6H6     5'- GCGCGCAATAATACACGGCCGTATCTTCCGGGGTCACGCTGTTCAGTTGCAGGCTAAACTGGTTTTTC -3'
OCLK1    5'- GGCTGAAGACGTGGGCGTGTATTATTGCCAGCAGCATTATACCACCCCGCCGACCTTTGGCCAGGGTAC -3'
OCLK2    5'- GCGGAAAAATAAACACGCTCGGAGCAGCCACCGTACGTTTAATTTCAACTTTCGTACCCTGGCCAAAGGTC -3'
OCLK3    5'- GAGCGTGTTTATTTTTCCGCCGAGCGATGAACAACTGAAAAGCGGCACGGCGAGCGTGGTGTGCCTGCTG -3'
OCLK4    5'- CAGCGCGTTGTCTACTTTCCACTGAACTTTCGCTTCACGCGGATAAAAGTTGTTCAGCAGGCACACCACGC -3'
OCLK5    5'- GAAAGTAGACAACGCGCTGCAAAGCGGCAACAGCCAGGAAAGCGTGACCGAACAGGATAGCAAAGATAG -3'
OCLK6    5'- GTTTTTCATAATCCGCTTTGCTCAGGGTCAGGGTGCTGCTCAGAGAATAGGTGCTATCTTTGCTATCCTGTTCG -3'
OCLK7    5'- GCAAAGCGGATTATGAAAAACATAAAGTGTATGCGTGCGAAGTGACCCATCAAGGTCTGAGCAGCCCGGTG -3'
OCLK8    5'- GGCATGCTTATCAGGCCTCGCCACGATTAAAAGATTTAGTCACCGGGCTGCTCAGAC -3'
OCH1     5'- GGCGTCTAGAGGCCAAGGCACCCTGGTGACGGTTAGCTCAGCGTCGAC -3'
```

Figure 6: (continued)

OCH2	5'- GTGCTTTTGCTGCTCGGAGCCAGCGGAAACACGCTTGGACCTTTGGTCGACGCTGAGCTAACC -3'
OCH3	5'- CTCCGAGCAGCAAAAGCACCAGCGGCGGCACGGCTGCCCTGGGCTGCCTGGTTAAAGATTATTTCC -3'
OCH4	5'- CTGGTCAGCGCCCCCGCTGTTCCAGCTCACGGTGACTGGTTCCGGGAAATAATCTTTAACCAGGCA -3'
OCH5	5'- AGCGGGGCGCTGACCAGCGGCGTGCATACCTTTCCGGCGGTGCTGCAAGGCAGCGGCCTG -3'
OCH6	5'- GTGCCTAAGCTGCTGCTCGGCACGGTCACAACGCTGCTCAGGCTATACAGGCCGCTGCTTTGCAG -3'
OCH7	5'- GAGCAGCAGCTTAGGCACTCAGACCTATATATTTGCAACGTGAACCATAAACCGAGCAACACC -3'
OCH8	5'- GCGCGAATTCGCTTTTTCGGTTCCACTTTTTTATCCACTTTGGTGTTGCTCGGTTTATGG -3'

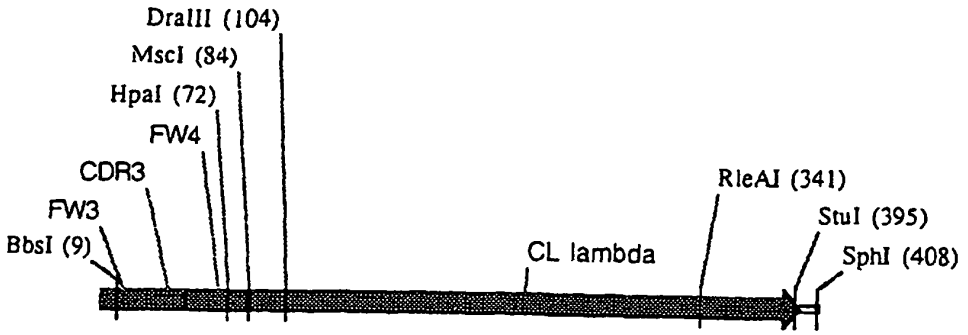
Figure 7A: sequence of the synthetic Cκ gene segment

```
      V  A   A  P  S   V  F  I   F  P  P  S   D  E  Q
 BsiWI
~~~~~~
CGTACGGTGG CTGCTCCGAG CGTGTTTATT TTTCCGCCGA GCGATGAACA
GCATGCCACC GACGAGGCTC GCACAAATAA AAAGGCGGCT CGCTACTTGT

  L  K  S   G  T  A  S   V  V  C   L  L  N   N  F  Y
ACTGAAAAGC GGCACGGCGA GCGTGGTGTG CCTGCTGAAC AACTTTTATC
TGACTTTTCG CCGTGCCGCT CGCACCACAC GGACGACTTG TTGAAAATAG

P  R  E  A   K  V  Q   W  K  V  D   N  A  L   Q  S  G
CGCGTGAAGC GAAAGTTCAG TGGAAAGTAG ACAACGCGCT GCAAAGCGGC
GCGCACTTCG CTTTCAAGTC ACCTTTCATC TGTTGCGCGA CGTTTCGCCG

 N  S  Q  E   S  V  T   E  Q  D   S  K  D  S   T  Y  S
AACAGCCAGG AAAGCGTGAC CGAACAGGAT AGCAAAGATA GCACCTATTC
TTGTCGGTCC TTTCGCACTG GCTTGTCCTA TCGTTTCTAT CGTGGATAAG

  L  S  S   T  L  T  L   S  K  A   D  Y  E   K  H  K
TCTGAGCAGC ACCCTGACCC TGAGCAAAGC GGATTATGAA AAACATAAAG
AGACTCGTCG TGGGACTGGG ACTCGTTTCG CCTAATACTT TTTGTATTTC

V  Y  A  C   E  V  T   H  Q  G  L   S  S  P   V  T  K
TGTATGCGTG CGAAGTGACC CATCAAGGTC TGAGCAGCCC GGTGACTAAA
ACATACGCAC GCTTCACTGG GTAGTTCCAG ACTCGTCGGG CCACTGATTT

 S  F  N  R   G  E  A   *
                StuI          SphI
               ~~~~~~~      ~~~~~~
TCTTTTAATC GTGGCGAGGC CTGATAAGCA TGC
AGAAAATTAG CACCGCTCCG GACTATTCGT ACG
```

Figure 7B: sequence of the synthetic CH1 gene segment

```
     A  S   T  K  G   P  S  V  F   P  L  A   P  S  S
 BlpI   SalI
~~~~~~~~~~~~~~~
GCTCAGCGTC GACCAAAGGT CCAAGCGTGT TTCCGCTGGC TCCGAGCAGC
CGAGTCGCAG CTGGTTTCCA GGTTCGCACA AAGGCGACCG AGGCTCGTCG

 K  S  T  S   G  G  T   A  A  L   G  C  L  V   K  D  Y
AAAAGCACCA GCGGCGGCAC GGCTGCCCTG GGCTGCCTGG TTAAAGATTA
TTTTCGTGGT CGCCGCCGTG CCGACGGGAC CCGACGGACC AATTTCTAAT

  F  P  E   P  V  T  V   S  W  N   S  G  A   L  T  S
TTTCCCGGAA CCAGTCACCG TGAGCTGGAA CAGCGGGGCG CTGACCAGCG
AAAGGGCCTT GGTCAGTGGC ACTCGACCTT GTCGCCCCGC GACTGGTCGC

G  V  H  T   F  P  A   V  L  Q  S   S  G  L   Y  S  L
GCGTGCATAC CTTTCCGGCG GTGCTGCAAA GCAGCGGCCT GTATAGCCTG
CGCACGTATG GAAAGGCCGC CACGACGTTT CGTCGCCGGA CATATCGGAC

 S  S  V  V   T  V  P   S  S  S   L  G  T  Q   T  Y  I
AGCAGCGTTG TGACCGTGCC GAGCAGCAGC TTAGGCACTC AGACCTATAT
TCGTCGCAAC ACTGGCACGG CTCGTCGTCG AATCCGTGAG TCTGGATATA

  C  N  V   N  H  K  P   S  N  T   K  V  D   K  K  V
TTGCAACGTG AACCATAAAC CGAGCAACAC CAAAGTGGAT AAAAAAGTGG
AACGTTGCAC TTGGTATTTG GCTCGTTGTG GTTTCACCTA TTTTTTCACC

E  P  K  S   E  F  *
             EcoRI     HindIII
            ~~~~~~     ~~~~~
AACCGAAAAG CGAATTCTGA TAAGCTT
TTGGCTTTTC GCTTAAGACT ATTCGAA
```

Figure 7C: functional map and sequence of module 24 comprising the synthetic Cλ gene segment (huCL lambda)



huCL lambda
408 bp

huCL lambda

```
      BbsI                                                                  HpaI
      ------                                                                --
  1   GAAGACGAAG CGGATTATTA TTGCCAGCAG CATTATACCA CCCCGCCTGT GTTTGGCGGC GGCACGAAGT
      CTTCTGCTTC GCCTAATAAT AACGGTCGTC GTAATATGGT GGGGCGGACA CAAACCGCCG CCGTGCTTCA
-----------------------------------------------------------------------------------
      HpaI        MscI              DraIII
      ----        ------            ----------
 71   TAACCGTTCT TGGCCAGCCG AAAGCCGCAC CGAGTGTGAC GCTGTTTCCG CCGAGCAGCG AAGAATTGCA
      ATTGGCAAGA ACCGGTCGGC TTTCGGCGTG GCTCACACTG CGACAAAGGC GGCTCGTCGC TTCTTAACGT
-----------------------------------------------------------------------------------
141   GGCGAACAAA GCGACCCTGG TGTGCCTGAT TAGCGACTTT TATCCGGGAG CCGTGACAGT GGCCTGGAAG
      CCGCTTGTTT CGCTGGGACC ACACGGACTA ATCGCTGAAA ATAGGCCCTC GGCACTGTCA CCGGACCTTC
-----------------------------------------------------------------------------------
211   GCAGATAGCA GCCCCGTCAA GGCGGGAGTG GAGACCACCA CACCCTCCAA ACAAAGCAAC AACAAGTACG
      CGTCTATCGT CGGGGCAGTT CCGCCCTCAC CTCTGGTGGT GTGGGAGGTT TGTTTCGTTG TTGTTCATGC
-----------------------------------------------------------------------------------
                                                  RleAI
                                                  ------
281   CGGCCAGCAG CTATCTGAGC CTGACGCCTG AGCAGTGGAA GTCCCACAGA AGCTACAGCT GCCAGGTCAC
      GCCGGTCGTC GATAGACTCG GACTGCGGAC TCGTCACCTT CAGGGTGTCT TCGATGTCGA CGGTCCAGTG
-----------------------------------------------------------------------------------
                                                  StuI        SphI
                                                  ------      ------
351   GCATGAGGGG AGCACCGTGG AAAAAACCGT TGCGCCGACT GAGGCCTGAT AAGCATGC
      CGTACTCCCC TCGTGGCACC TTTTTTGGCA ACGCGGCTGA CTCCGGACTA TTCGTACG
-----------------------------------------------------------------------
```

Figure 7D: oligonucleotides used for synthesis of module M24 containing Cλ gene segment

M24: assembly PCR

M24-A
GAAGACAAGCGGATTATTATTGCCAGCAGCATTATACCACCCCGCCTGTGTTTGGCGGCGGCACG
AAGTTAACCGTTC

M24-B
CAATTCTTCGCTGCTCGGCGGAAACAGCGTCACACTCGGTGCGGCTTTCGGCTGGCCAAGAACGGT
TAACTTCGTGCCGC

M24-C
CGCCGAGCAGCGAAGAATTGCAGGCGAACAAAGCGACCCTGGTGTGCCTGATTAGCGACTTTTAT
CCGGGAGCCGTGACA

M24-D
TGTTTGGAGGGTGTGGTGGTCTCCACTCCCGCCTTGACGGGGCTGCTATCTGCCTTCCAGGCCACTG
TCACGGCTCCCGG

M24-E
CCACACCCTCCAAACAAAGCAACAACAAGTACGCGGCCAGCAGCTATCTGAGCCTGACGCCTGA
GCAGTGGAAGTCCCACAGAAGCTACAGCTG

M24-F
GCATGCTTATCAGGCCTCAGTCGGCGCAACGGTTTTTTCCACGGTGCTCCCCTCATGCGTGACCTGG
CAGCTGTAGCTTC

Figure 8: sequence and restriction map of the synthetic gene encoding the consensus single-chain fragment VH3-Vκ2

```
 M  K  Q  S  T  I  A  L  A  L  L  P  L  L  F  T  P
                                        SapI
                                       ~~~~~~~~
ATGAAACAAA GCACTATTGC ACTGGCACTC TTACCGTTGC TCTTCACCCC
TACTTTGTTT CGTGATAACG TGACCGTGAG AATGGCAACG AGAAGTGGGG

  V  T  K  A  D  Y  K  D  E  V  Q  L  V  E  S  G
                                 MfeI
                                ~~~~~~
TGTTACCAAA GCCGACTACA AAGATGAAGT GCAATTGGTG GAAAGCGGCG
ACAATGGTTT CGGCTGATGT TTCTACTTCA CGTTAACCAC CTTTCGCCGC

G  G  L  V  Q  P  G  G  S  L  R  L  S  C  A  A  S
                                                BspEI
                                                 ~~~
GCGGCCTGGT GCAACCGGGC GGCAGCCTGC GTCTGAGCTG CGCGGCCTCC
CGCCGGACCA CGTTGGCCCG CCGTCGGACG CAGACTCGAC GCGCCGGAGG

 G  F  T  F  S  S  Y  A  M  S  W  V  R  Q  A  P  G
BspEI                                        BstXI
~~~                                         ~~~~~~~~~~~~
GGATTTACCT TTAGCAGCTA TGCGATGAGC TGGGTGCGCC AAGCCCCTGG
CCTAAATGGA AATCGTCGAT ACGCTACTCG ACCCACGCGG TTCGGGGACC

  K  G  L  E  W  V  S  A  I  S  G  S  G  G  S  T
        XhoI
       ~~~~~~~
GAAGGGTCTC GAGTGGGTGA GCGCGATTAG CGGTAGCGGC GGCAGCACCT
CTTCCCAGAG CTCACCCACT CGCGCTAATC GCCATCGCCG CCGTCGTGGA

Y  Y  A  D  S  V  K  G  R  F  T  I  S  R  D  N  S
                                    PmlI      NspV
                                   ~~~~~~~    ~~~~
ATTATGCGGA TAGCGTGAAA GGCCGTTTTA CCATTTCACG TGATAATTCG
TAATACGCCT ATCGCACTTT CCGGCAAAAT GGTAAAGTGC ACTATTAAGC

 K  N  T  L  Y  L  Q  M  N  S  L  R  A  E  D  T  A
NspV                                          EagI
~~                                            ~~~~
AAAAACACCC TGTATCTGCA AATGAACAGC CTGCGTGCGG AAGATACGGC
TTTTTGTGGG ACATAGACGT TTACTTGTCG GACGCACGCC TTCTATGCCG

  V  Y  Y  C  A  R  W  G  G  D  G  F  Y  A  M  D
EagI      BssHII
~~        ~~~~~~
CGTGTATTAT TGCGCGCGTT GGGGCGGCGA TGGCTTTTAT GCGATGGATT
GCACATAATA ACGCGCGCAA CCCCGCCGCT ACCGAAAATA CGCTACCTAA
```

Figure 8: (continued)

```
Y  W  G  Q   G  T  L   V  T  V  S   S  A  G   G  G  S
       StyI                        BlpI
      ~~~~~~~                     ~~~~~~~
ATTGGGGCCA AGGCACCCTG GTGACGGTTA GCTCAGCGGG TGGCGGTTCT
TAACCCCGGT TCCGTGGGAC CACTGCCAAT CGAGTCGCCC ACCGCCAAGA

 G  G  G  G   S  G  G   G  G  S   G  G  G  G   S  D  I
                                                    EcoRV
                                                    ~~~~~
GGCGGCGGTG GGAGCGGTGG CGGTGGTTCT GGCGGTGGTG GTTCCGATAT
CCGCCGCCAC CCTCGCCACC GCCACCAAGA CCGCCACCAC CAAGGCTATA

  V  M  T   Q  S  P  L   S  L  P   V  T  P   G  E  P
EcoRV        BanII
~           ~~~~~~
CGTGATGACC CAGAGCCCAC TGAGCCTGCC AGTGACTCCG GGCGAGCCTG
GCACTACTGG GTCTCGGGTG ACTCGGACGG TCACTGAGGC CCGCTCGGAC

A  S  I  S   C  R  S   S  Q  S  L   L  H  S   N  G  Y
            PstI
           ~~~~~~
CGAGCATTAG CTGCAGAAGC AGCCAAAGCC TGCTGCATAG CAACGGCTAT
GCTCGTAATC GACGTCTTCG TCGGTTTCGG ACGACGTATC GTTGCCGATA

 N  Y  L  D   W  Y  L   Q  K  P   G  Q  S   P  Q  L  L
              KpnI          SexAI               AseI
             ~~~~~~        ~~~~~~~~             ~~~
AACTATCTGG ATTGGTACCT TCAAAAACCA GGTCAAAGCC CGCAGCTATT
TTGATAGACC TAACCATGGA AGTTTTTGGT CCAGTTTCGG GCGTCGATAA

  I  Y  L   G  S  N  R   A  S  G   V  P  D   R  F  S
AseI                          Eco0109I
~~~                           ~~~~~~~~
AATTTATCTG GGCAGCAACC GTGCCAGTGG GGTCCCGGAT CGTTTTAGCG
TTAAATAGAC CCGTCGTTGG CACGGTCACC CCAGGGCCTA GCAAAATCGC

G  S  G  S   G  T  D   F  T  L  K   I  S  R   V  E  A
      BamHI
     ~~~~~~~
GCTCTGGATC CGGCACCGAT TTTACCCTGA AAATTAGCCG TGTGGAAGCT
CGAGACCTAG GCCGTGGCTA AAATGGGACT TTTAATCGGC ACACCTTCGA

 E  D  V  G   V  Y  Y   C  Q  Q   H  Y  T  T   P  P  T
 BbsI
~~~~~~
GAAGACGTGG GCGTGTATTA TTGCCAGCAG CATTATACCA CCCCGCCGAC
CTTCTGCACC CGCACATAAT AACGGTCGTC GTAATATGGT GGGGCGGCTG

  F  G  Q   G  T  K  V   E  I  K   R  T  E   F
    MscI                          BsiWI EcoRI
   ~~~~~~                        ~~~~~~~~~~~~
CTTTGGCCAG GGTACGAAAG TTGAAATTAA ACGTACGGAA TTC
GAAACCGGTC CCATGCTTTC AACTTTAATT TGCATGCCTT AAG
```

Figure 9: Phage display vector pIG10.3

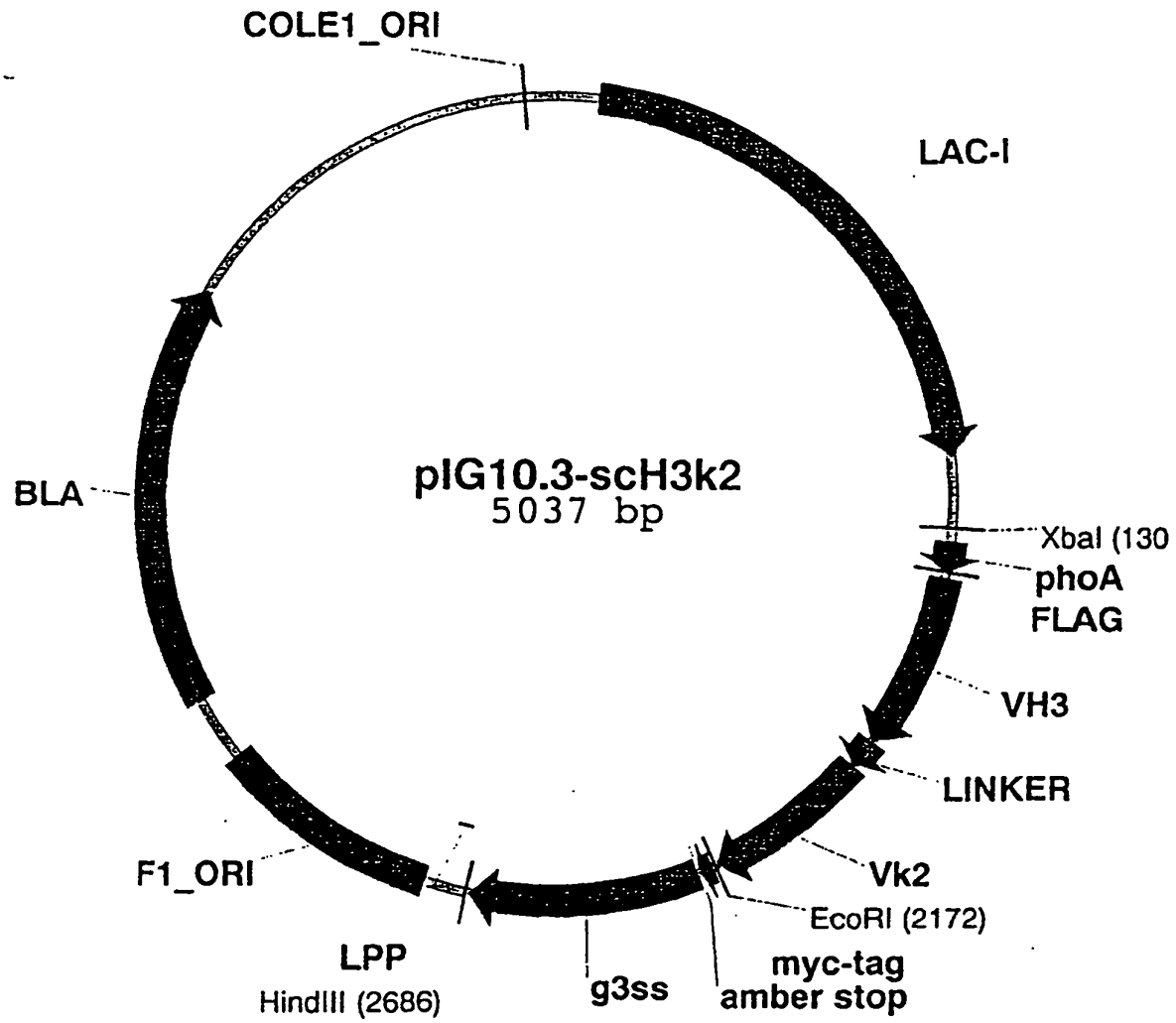


Figure 10: Sequence analysis of initial libraries

	92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103
A	C	A	R	W	G	G	D	G	.F	Y	A	-	-	M	D	Y	W
B	C	A	R	F	G	K	M	N	Y	-	-	-	-	-	D	Y	W
	C	A	R	H	R	T	E	W	H	-	-	-	-	-	D	Y	W
	C	A	R	V	R	E	L	Y	H	-	-	-	-	-	D	Y	W
	C	A	R	K	F	L	K	A	R	-	-	-	-	-	D	Y	W
	C	A	R	W	N	T	T	G	Y	-	-	-	-	-	D	Y	W
	C	A	R	I	N	E	A	Q	P	-	-	-	-	-	D	Y	W
	C	A	R	T	A	I	T	R	-	-	-	-	-	-	D	Y	W
	C	A	R	W	Y	N	R	N	S	-	-	-	-	-	D	Y	W
	C	A	R	S	V	G	D	S	K	-	-	-	-	-	D	Y	W
	C	A	R	S	K	T	F	A	A	-	-	-	-	-	D	Y	W
	C	A	R	V	A	P	Q	Y	D	-	-	-	-	-	D	Y	W
	C	A	R	M	Q	S	E	W	M	-	-	-	-	-	D	Y	W
C	C	A	R	Y	F	V	H	F	L	Y	T	M	V	M	D	V	W
	C	A	R	M	A	L	R	A	S	G	K	Y	I	M	D	V	W
	C	A	R	K	N	Q	M	V	F	H	A	R	K	F	D	V	W
	C	A	R	T	Q	S	F	W	E	Q	Q	K	V	M	D	Y	W
	C	A	R	Y	P	Y	R	S	N	F	F	M	P	M	D	V	W
	C	A	R	*	G	S	G	S	E	H	W	S	I	F	D	V	W
	C	A	R	R	N	P	W	N	V	N	Y	L	H	F	D	V	W
	C	A	R	M	K	P	M	L	N	R	D	G	T	M	D	V	W
	C	A	R	K	G	S	E	F	L	E	T	D	V	M	D	Y	W
	C	A	R	S	W	T	N	D	K	P	N	F	I	M	D	V	W
	C	A	R	Y	A	G	T	T	F	K	Q	G	P	M	D	Y	W

Figure 11: Expression analysis of initial library

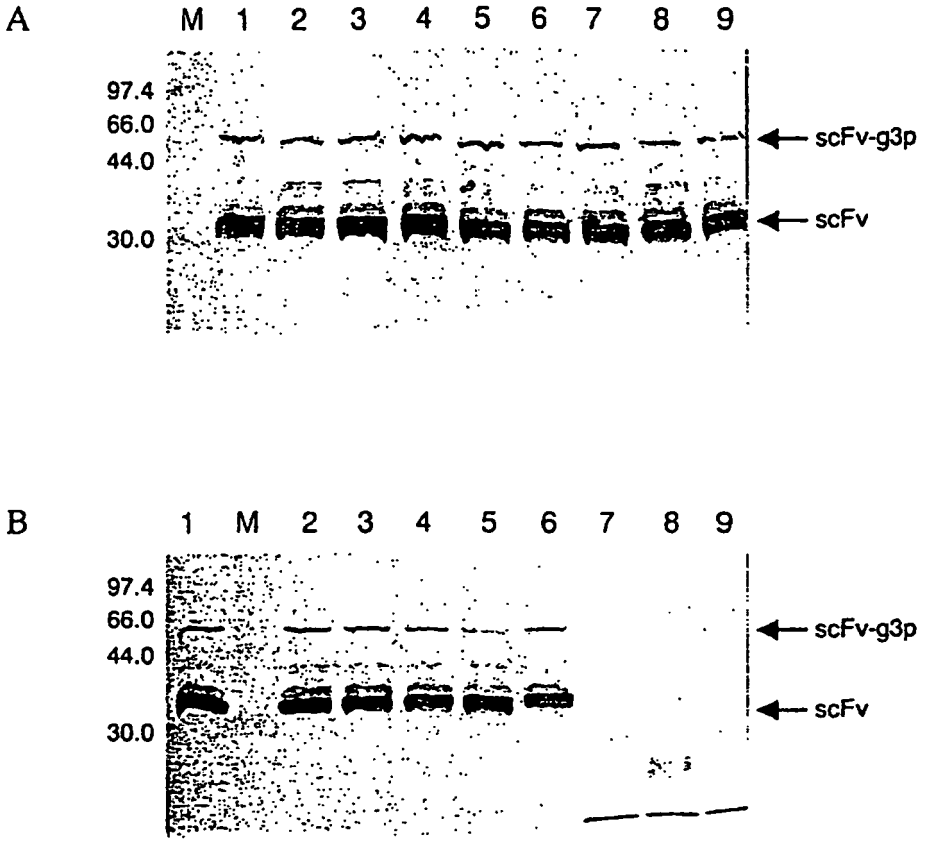


Figure 12: Increase of specificity during the panning rounds

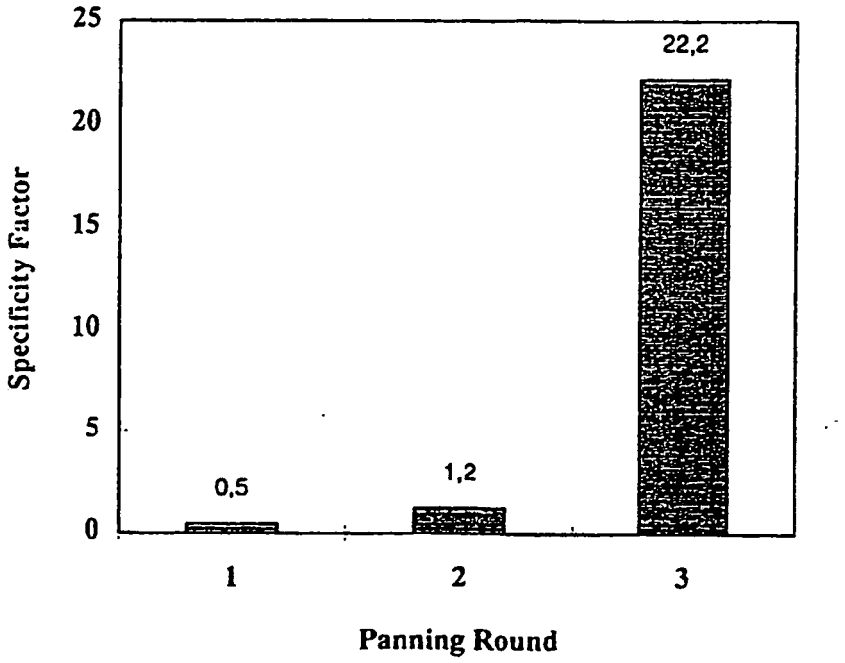


Figure 13: Phage ELISA of clones after the 3rd round of panning

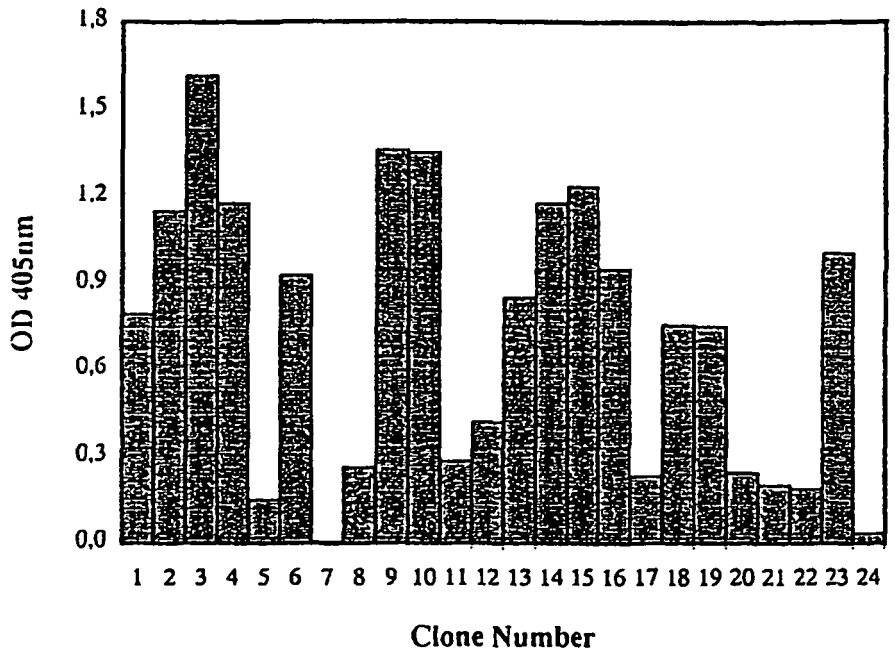


Figure 14: Competition ELISA

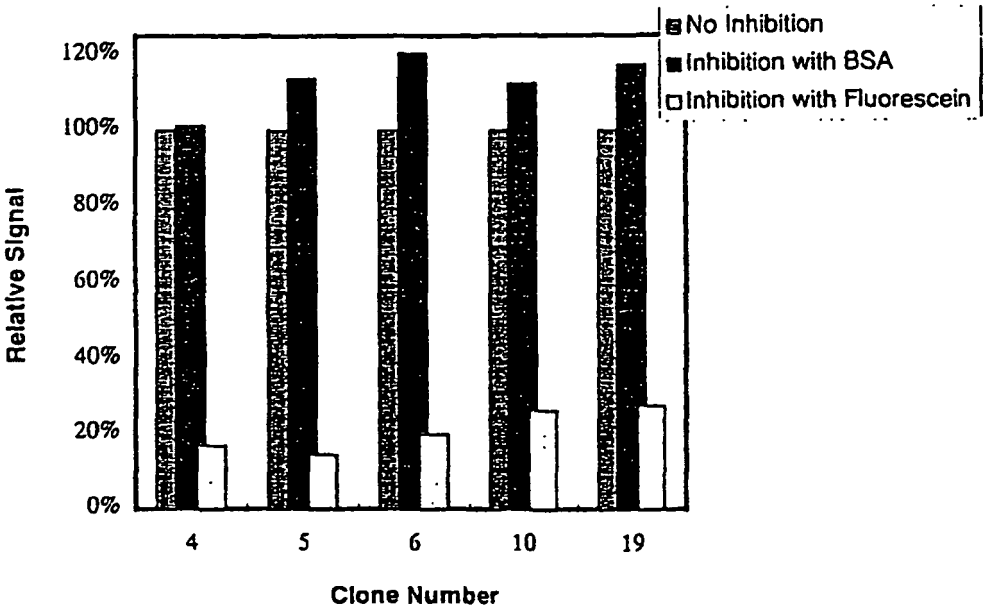


Figure 15: Sequence analysis of fluorescein binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103	Frequency
C	A	R	K	R	M	M	Q	N	P	R	F	R	F	D	V	W	1
C	A	R	R	S	K	Q	K	R	K	M	R	R	F	D	V	W	3
C	A	R	R	N	G	K	R	H	L	R	H	R	F	D	V	W	1
C	A	R	R	K	M	R	K	R	I	K	R	R	F	D	V	W	2
C	A	R	Y	R	K	I	M	K	W	K	N	S	F	D	V	W	1
C	A	R	L	I	E	V	H	P	S	F	D	Q	M	D	V	W	1
C	A	R	R	K	P	M	F	L	K	K	A	V	F	D	V	W	1
C	A	R	R	K	F	H	R	Y	S	T	V	K	F	D	Y	W	2
C	A	R	R	K	T	M	R	S	R	V	K	Y	F	D	Y	W	1
C	A	R	K	K	R	S	W	R	R	M	D	R	F	D	V	W	1
C	A	R	R	N	P	R	R	G	R	M	N	R	F	D	V	W	1
C	A	R	K	G	K	K	K	F	A	R	P	R	F	D	V	W	1
C	A	R	R	M	V	H	K	G	K	R	K	I	F	D	V	W	1
C	A	R	R	K	H	I	T	Y	P	R	K	Q	F	D	V	W	1
C	A	R	R	W	T	K	R	R	S	F	A	R	F	D	V	W	1
C	A	R	K	K	L	K	Q	Y	T	F	S	R	F	D	Y	W	1

Figure 16: Purification of fluorescein binding scFv fragments

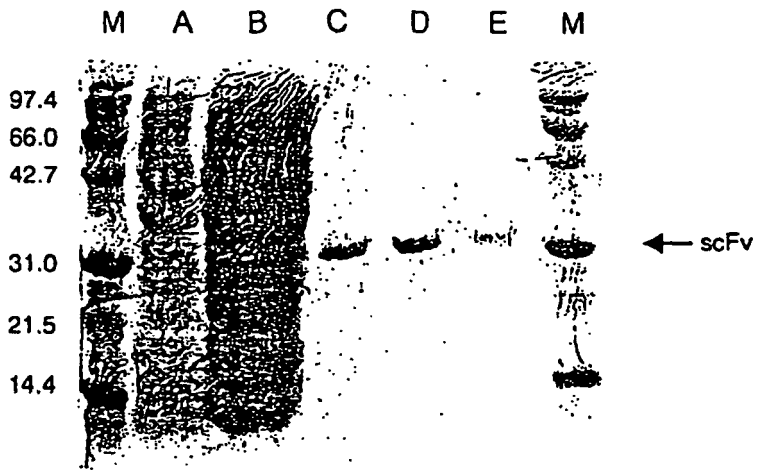
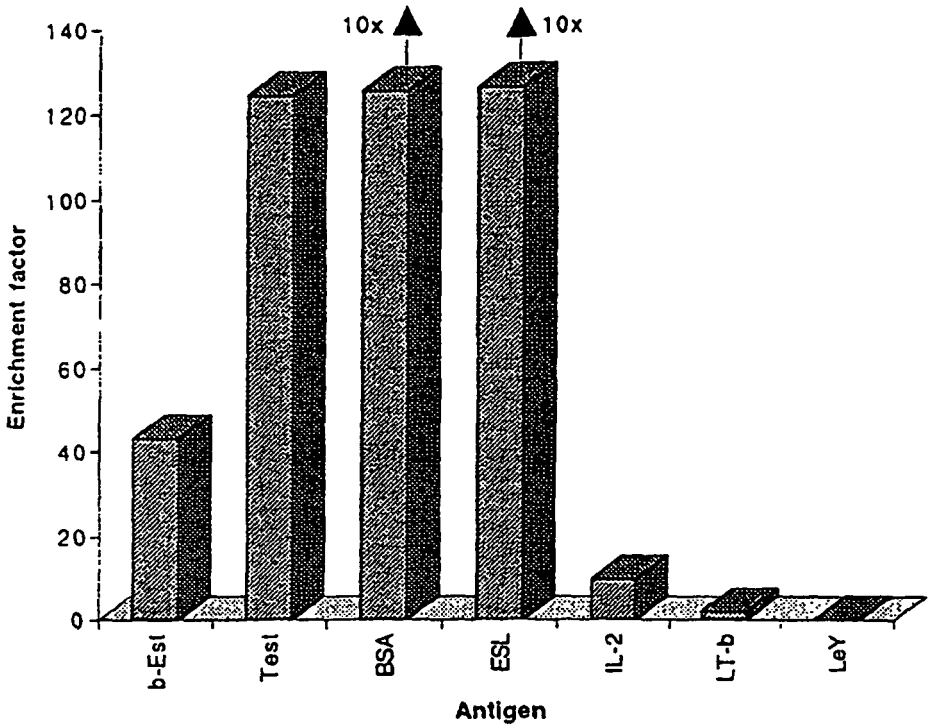


Figure 17: Enrichment factors after three rounds of panning



estradiol

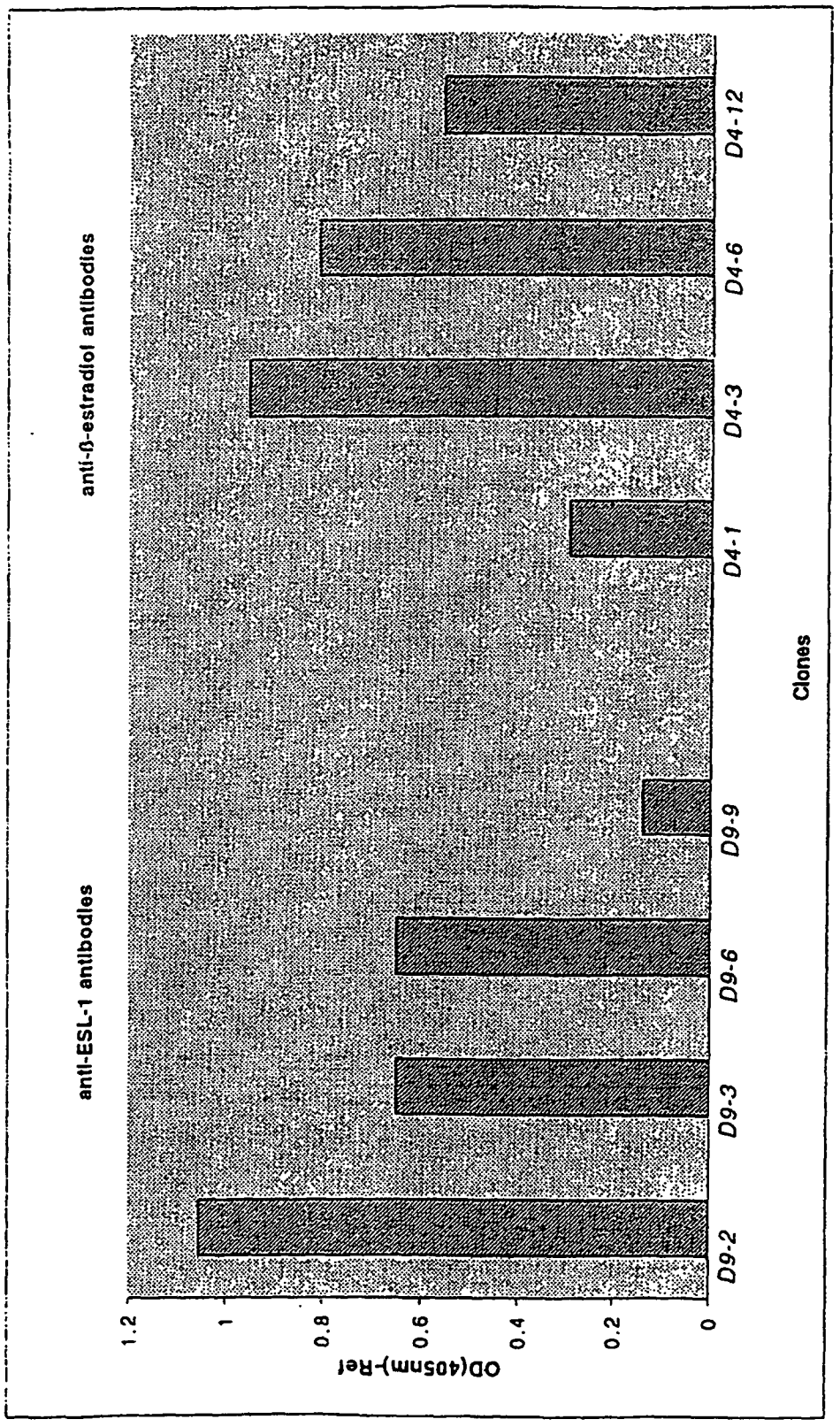


Figure 18: ELISA of anti-ESL-1 and anti-β-estradiol antibodies

Figure 19: Selectivity and cross-reactivity of HuCAL antibodies

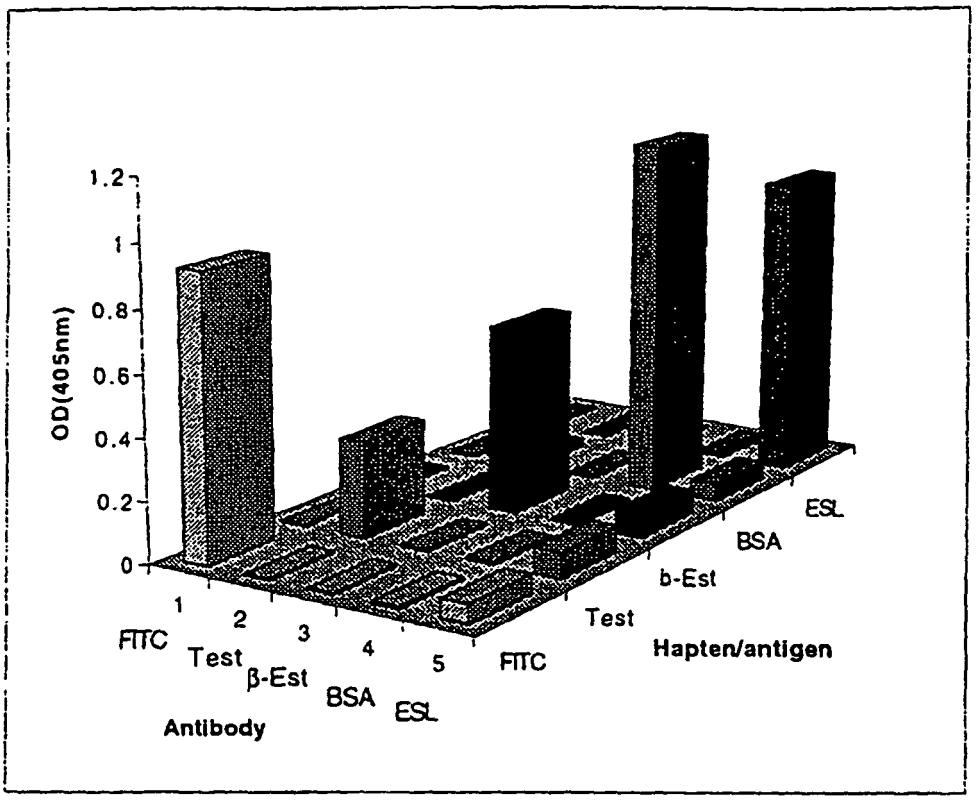


Figure 20: Sequence analysis of estradiol binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103	Frequency
C	A	R	T	R	P	W	Q	A	T	R	K	G	F	D	V	W	3
C	A	R	N	Q	W	E	F	K	N	R	R	K	M	D	Y	W	8
C	A	R	K	R	W	M	W	P	I	G	K	R	F	D	Y	W	7
C	A	R	Y	S	L	W	R	L	D	E	Y	F	F	D	Y	W	1
C	A	R	V	P	W	G	D	F	W	S	W	H	M	D	V	W	1
C	A	R	N	G	L	E	P	R	H	R	K	M	M	D	Y	W	1
C	A	R	I	M	K	A	P	P	-	-	-	-	-	D	Y	W	1
C	A	R	R	K	T	W	H	W	F	Y	K	R	M	D	Y	W	1
C	A	R	W	K	D	M	W	S	Q	V	Y	V	M	D	Y	W	1
C	A	R	N	K	Q	Q	M	R	F	R	R	F	M	D	Y	W	5
C	A	R	N	M	L	A	L	S	R	G	K	E	M	D	V	W	4
C	A	R	N	M	R	L	M	R	M	R	K	N	F	D	V	W	1

Figure 21: Sequence analysis of testosterone binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103	Frequency
C	A	R	Y	I	K	Q	A	K	R	K	L	A	F	D	Y	W	4
C	A	R	Y	N	R	H	A	W	Q	K	M	Q	F	D	Y	W	3
C	A	R	Y	V	K	Y	A	R	N	K	M	Q	F	D	Y	W	2
C	A	R	Y	K	R	G	A	W	M	K	T	M	F	D	V	W	1
C	A	R	R	K	P	L	R	R	I	M	K	W	F	D	Y	W	1
C	A	R	Y	R	K	R	A	S	R	Q	M	Q	F	D	Y	W	1

Figure 22: Sequence analysis of lymphotoxin-ß binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103	Frequency
C	A	R	Q	R	Y	R	S	K	I	K	G	H	F	D	V	W	16
C	A	R	_	W	R	D	F	N	S	Y	D	P	M	D	Y	W	1
C	A	R	M	A	D	L	D	N	Y	W	V	Q	F	D	Y	W	1
C	A	R	L	Q	A	Y	L	K	P	H	H	W	M	D	Y	W	1
C	A	R	R	L	I	E	Q	A	R	D	H	V	M	D	Y	W	1
C	A	R	S	W	H	N	S	Q	F	T	Q	S	F	D	V	W	1
C	A	R	V	D	H	F	Q	T	E	N	E	W	M	D	Y	W	1
C	A	R	D	W	P	T	L	I	F	W	Y	W	F	D	Y	W	1

Figure 23: Sequence analysis of ESL-1 binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100Ca	100D	100E	101	102	103	Frequency
C	A	R	G	F	G	F	T	E	-	-	-	-	-	-	D	Y	W	4
C	A	R	Q	F	D	E	D	S	F	V	R	-	R	F	D	V	W	4
C	A	R	I	L	K	E	S	S	K	S	R	-	Q	M	D	V	W	2
C	A	R	E	Q	D	E	Y	G	A	I	R	-	I	M	D	Y	W	1
C	A	R	N	H	F	E	A	S	W	P	R	R	Q	M	D	V	W	1
C	A	R	E	N	E	W	V	D	M	I	L	-	D	M	D	Y	W	2
C	A	R	Q	Y	S	E	T	R	W	V	R	-	K	F	D	Y	W	1
C	A	R	Q	F	K	E	S	K	T	R	R	-	K	F	D	V	W	13
C	A	R	K	K	T	Q	Y	V	H	D	W	-	R	M	D	V	W	3
C	A	R	R	W	R	E	T	K	S	K	R	-	F	F	D	V	W	1
C	A	R	D	Y	I	M	E	F	-	-	-	-	-	-	D	Y	W	1
C	A	R	Q	F	E	E	T	K	Q	R	R	-	L	M	D	Y	W	1

Figure 24: Sequence analysis of BSA binders

92	93	94	95	96	97	98	99	100	100A	100B	100C	100D	100E	101	102	103	Frequency
C	A	R	D	Q	G	F	Y	A	I	D	Y	V	M	D	Y	W	5
C	A	R	V	F	T	Y	M	Y	N	Y	F	R	F	D	V	W	1
C	A	R	V	F	F	E	Q	M	E	V	V	R	M	D	V	W	1
C	A	R	E	K	E	Y	R	L	S	W	S	Q	M	D	Y	W	1
C	A	R	Y	P	S	R	W	A	P	N	W	Y	M	D	Y	W	1
C	A	R	D	G	G	F	K	P	L	T	H	F	F	D	V	W	1

Figure 25: modular pCAL vector system

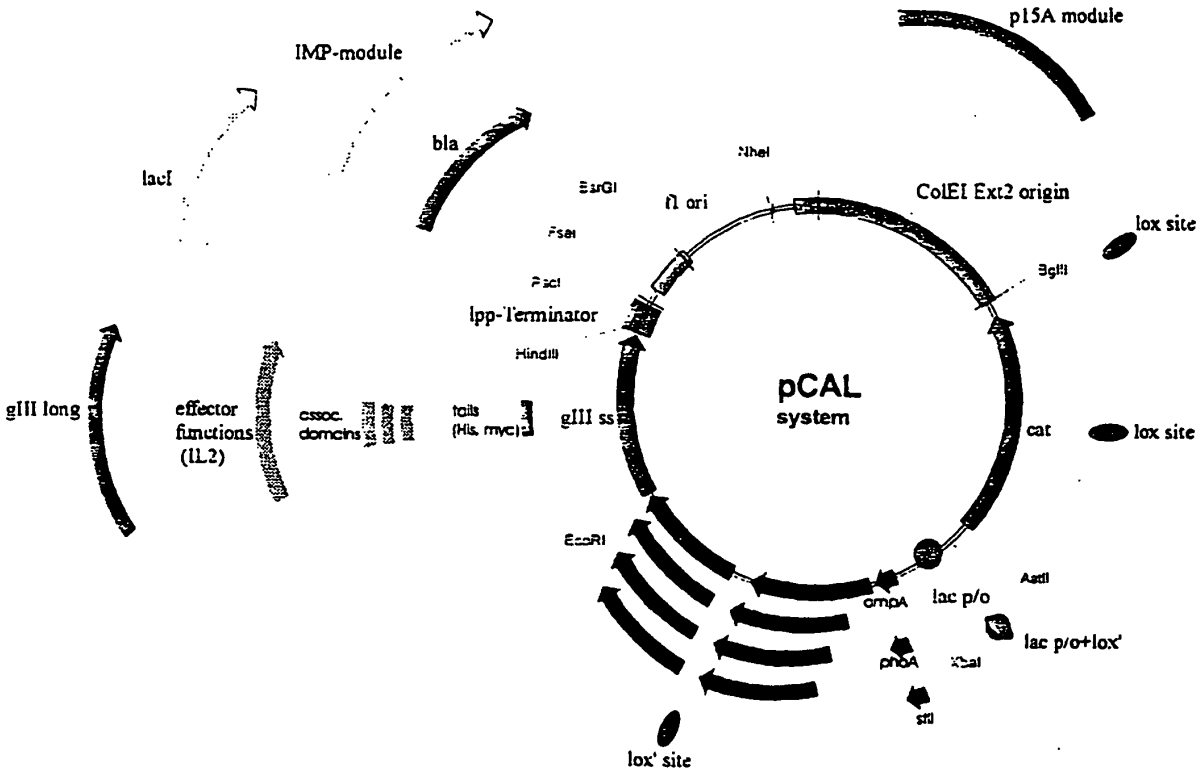


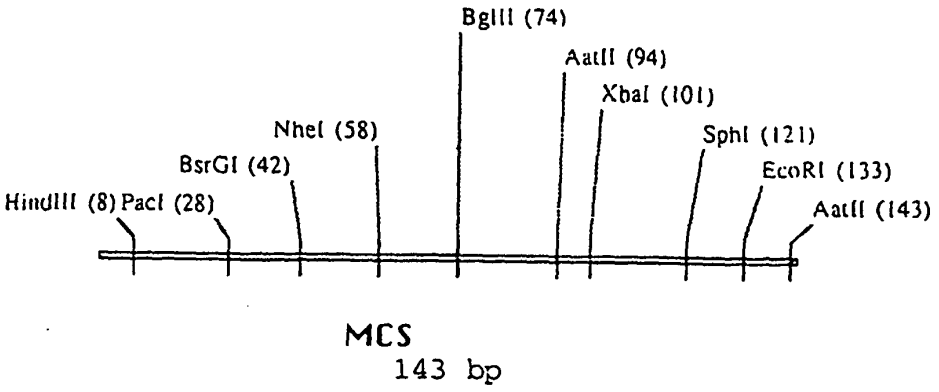
Figure 25a: List of unique restriction sites used in or suitable for HuCAL genes or pCAL vectors

unique restriction site	Isoschizomers
AatII	/
AflII	BfrI, BspTI, Bst98I
AscI	/
AseI	VspI, AsnI, PshBI
BamHI	BstI
BbeI	EheI, KasI, NarI
BbsI	BpuAI, BpiI
BglII	/
BlpI	Bpu1102I,CelII, BtpI
BsaBI	MamI, Bsh1365I, BsrBRI
BsiWI	Pfl23II, SpII, SunI
BspEI	AccIII, BseAI, BsiMI, Kpn2I, MroI
BsrGI	Bsp1407I, SspBI
BssHII	PauI
BstEII	BstPI, Eco91I, EcoO65I
BstXI	/
Bsu36I	AocI, CviI, Eco81I
DraIII	/
DsmAI	
EagI	BstZI, EclXI, Eco52I, XmaIII
Eco57I	/
EcoO109I	DraII
EcoRI	/
EcoRV	Eco32I
FseI	/
HindIII	/
HpaI	/
KpnI	Acc65I, Asp718I
MluI	/
MscI	BalI, MluNI
MunI	MfeI
NheI	/
NsiI	Ppu10I, EcoT22I, Mph1103I
NspV	Bsp119I, BstBI, Csp45I, LspI, SfuI
PacI	/
PmeI	/
PmlI	BbrPI, Eco72I, PmaCI
Psp5II	PpuMI
PstI	/
RsrII	(RsrII), CpoI, CspI
SanDI	/
SapI	/
SexAI	/
SpeI	/
SfiI	/
SphI	BbuI, PaeI,NspI
StuI	AatII, Eco147I
StyI	Eco130I, EcoT14I
XbaI	BspLU11I
XhoI	PaeR7I
XmaI	AvaI, SmaI, Cfr9I, PspAI

Figure26: list of pCAL vector modules

No	module/flanking restriction sites	functional element	sites to be removed	sites to be inserted	template	reference
M1	AatII-lacp/o-XbaI	lac promotor/operator	2x VspI (AseI)	AatII	vector pASK30	Skerra et al. (1991) Bio/Technology 9, 273-278
M2	BglII-lox-AatII	Cre/lox recombination site	2x VspI (AseI)	lox, BglII	(synthetic)	Hoess et al. (1986) Nucleic Acids Res. 2287-2300
M3	XbaI-lox'-SphI	Cre/lox' recombination site	none	lox', SphI	(synthetic)	see M2
M7-I	EcoRI-gIIIlong-HindIII	gIIIp of filamentous phage with N-terminal myctail/amber codon	SphI, BamHI	none	vector pIG10	Ge et al., (1994) Expressing antibodies in E. coli. In: Antibody engineering: A practical approach. IRL Press, New York, pp 229-266
M7-II	EcoRI-gIIIss-HindIII	truncated gIIIp of filamentous phage with N-terminal Gly-Ser linker	SphI		vector pIG10	see M7-I
M7-III	EcoRI-gIIIss-HindIII	truncated gIIIp of filamentous phage with N-terminal myctail/amber codon	SphI, BbsI		vector pIG10	see M7-I
M8	SphI-lox-HindIII	Cre/lox recombination site	none	lox	(synthetic)	see M3
M9-II	HindIII-tpp-PacI	tpp-terminator	none	PacI, FseI	(synthetic)	see M1
M10-II	PacI/FseI-bla-BsrGI	beta-lactamase/bla (ampR)	VspI, Eco57I, BssSI	PacI, FseI, BsrGI	pASK30	see M1
M11-II	BsrGI-f1 ori-NheI	origin of single-stranded replication	DraIII (BanII not removed)	BsrGI, NheI	pASK3O	see M1
M11-III	BsrGI-f1 ori-NheI	origin of single-stranded replication	DraIII, BanII	BsrGI, NheI	pASK30	see M1
M12	NheI-p15A-BglII	origin of double-stranded replication	BssSI, VspI, NspV	NheI, BglII	pACYC184	Rose, R.E. (1988) Nucleic Acids Res. 16, 355
M13	BglII-lox-BglII	Cre/lox recombination site	none	BglII, lox, XmnI	(synthetic)	see M3
M14-Ext2	BglII-ColEI-NheI	origin of double-stranded replication	Eco57I (BssSI not removed)	BglII, NheI	pUC19	Yanisch-Peron, C. (1985) Gene 33.103-119
M17	AatII-cat-BglII	chloramphenicol-acetyltransferase/ cat (camR)	BspEI, MscI, StyI/NcoI		pACYC184	Cardoso, M. & Schwarz,S. (1992) J. Appl. Bacteriol.72, 289-293
M19	XbaI-phoA-EcoRI	signal sequence of phosphatase A	(synthetic)		(synthetic)	see M1
M20	XbaI-phoA-FLAG-EcoRI	signal sequence of phosphatase A + FLAG detection tag	(synthetic)		(synthetic)	Knappik, A & Plückthun, A. (1994) BioTechniques 17, 754-761
M21	XbaI-stII-SapI	heat-stable enterotoxin II signal sequence	(synthetic)		(synthetic)	Lee et al. (1983) Infect. Immunol. 264-268
M41	AflII-lacI-NheI	lac-repressor	BstXI, MluI,BbsI, BanII, BstEII, HpaI, BbeI, VspI		pASK30	see M1
M42	EcoRI-Histail-HindIII	poly-histidine tail	(synthetic)		(synthetic)	Lindner et al., (1992) Methods: a companion to methods in enzymology 4, 41-56

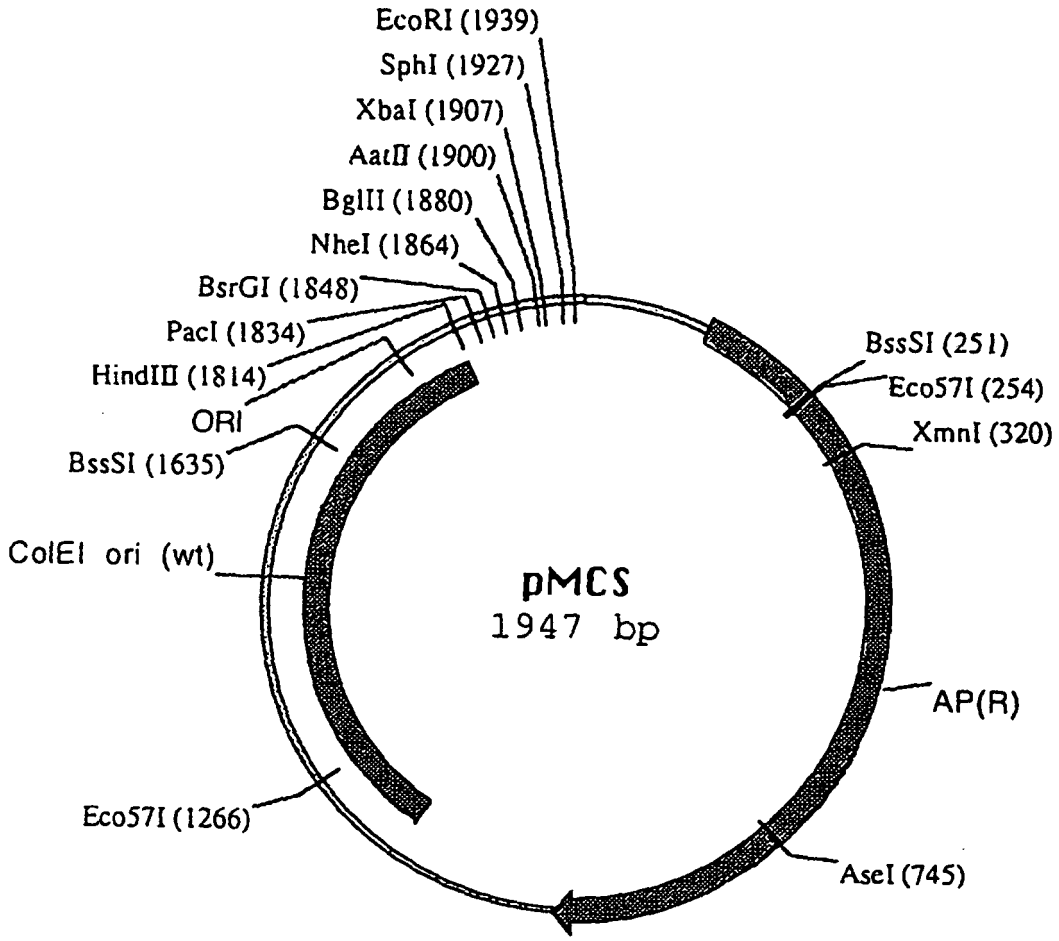
Figure 27: functional map and sequence of MCS module



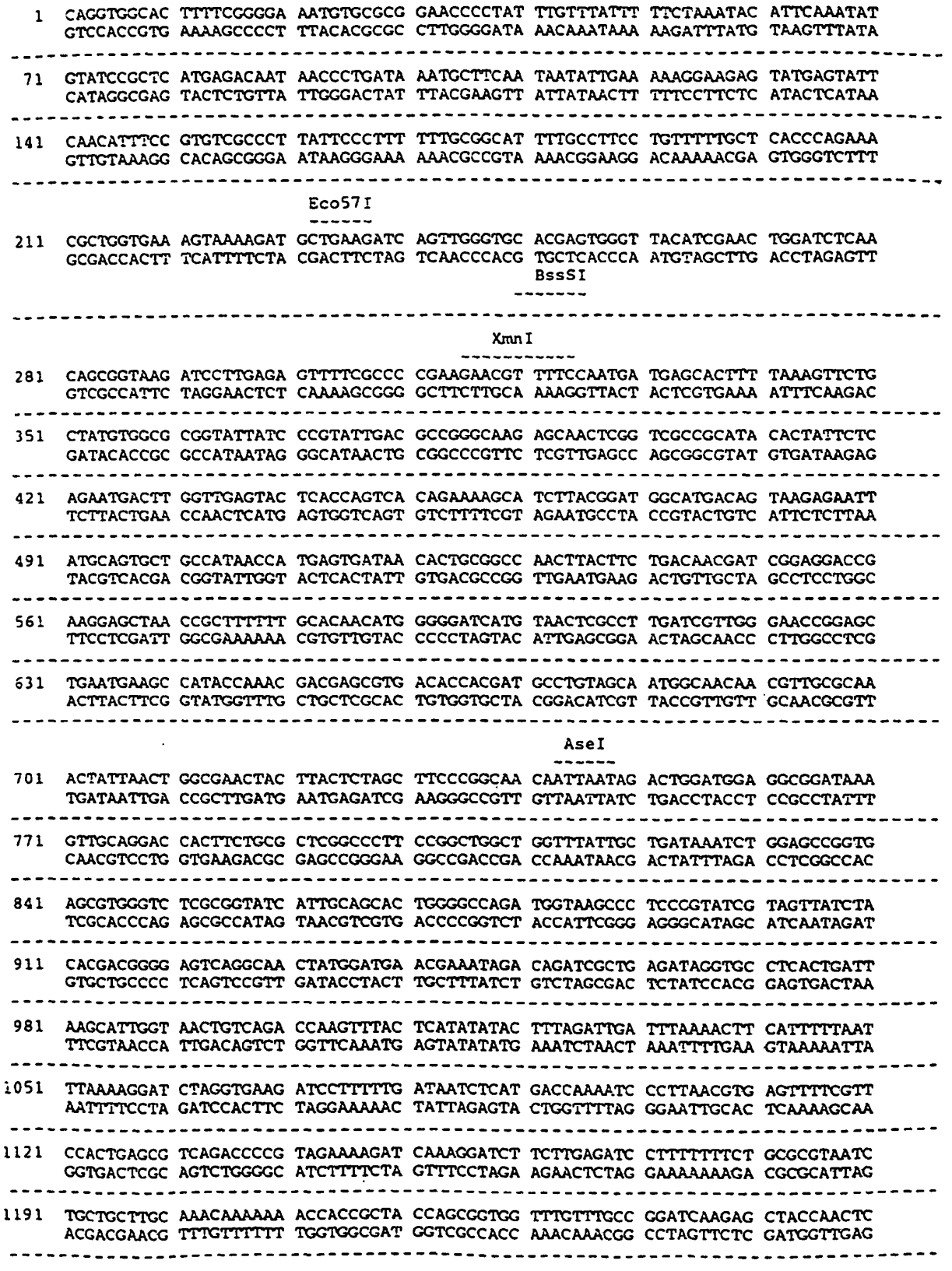
MCS

```
          HindIII             PacI               BsrGI           NheI
          -------             --------           ------          -------
  1  ACATGTAAGC TTCCCCCCCC CCTTAATTAA CCCCCCCCCC TGTACACCCC CCCCCCGCTA GCCCCCCCCC
     TGTACATTCG AAGGGGGGGG GGAATTAATT GGGGGGGGGG ACATGTGGGG GGGGGGCGAT CGGGGGGGGG
-------------------------------------------------------------------------------
                                                                      EcoRI
                                                                      ------
       BglII            AatII         XbaI             SphI                AatII
       ------           -------       -------          -------             ---
 71  CCAGATCTCC CCCCCCCCGA CGTCCCCCCT CTAGACCCCC CCCCCGCATG CCCCCCCCCC CGAATTCGAC
     GGTCTAGAGG GGGGGGGGCT GCAGGGGGGA GATCTGGGGG GGGGGCGTAC GGGGGGGGGG GCTTAAGCTG
-------------------------------------------------------------------------------
     Aat
     ---
141  GTC
     CAG
-------------------------------------------------------------------------------
```

Figure 28: functional map and sequence of pMCS cloning vector



pMCS



pMCS

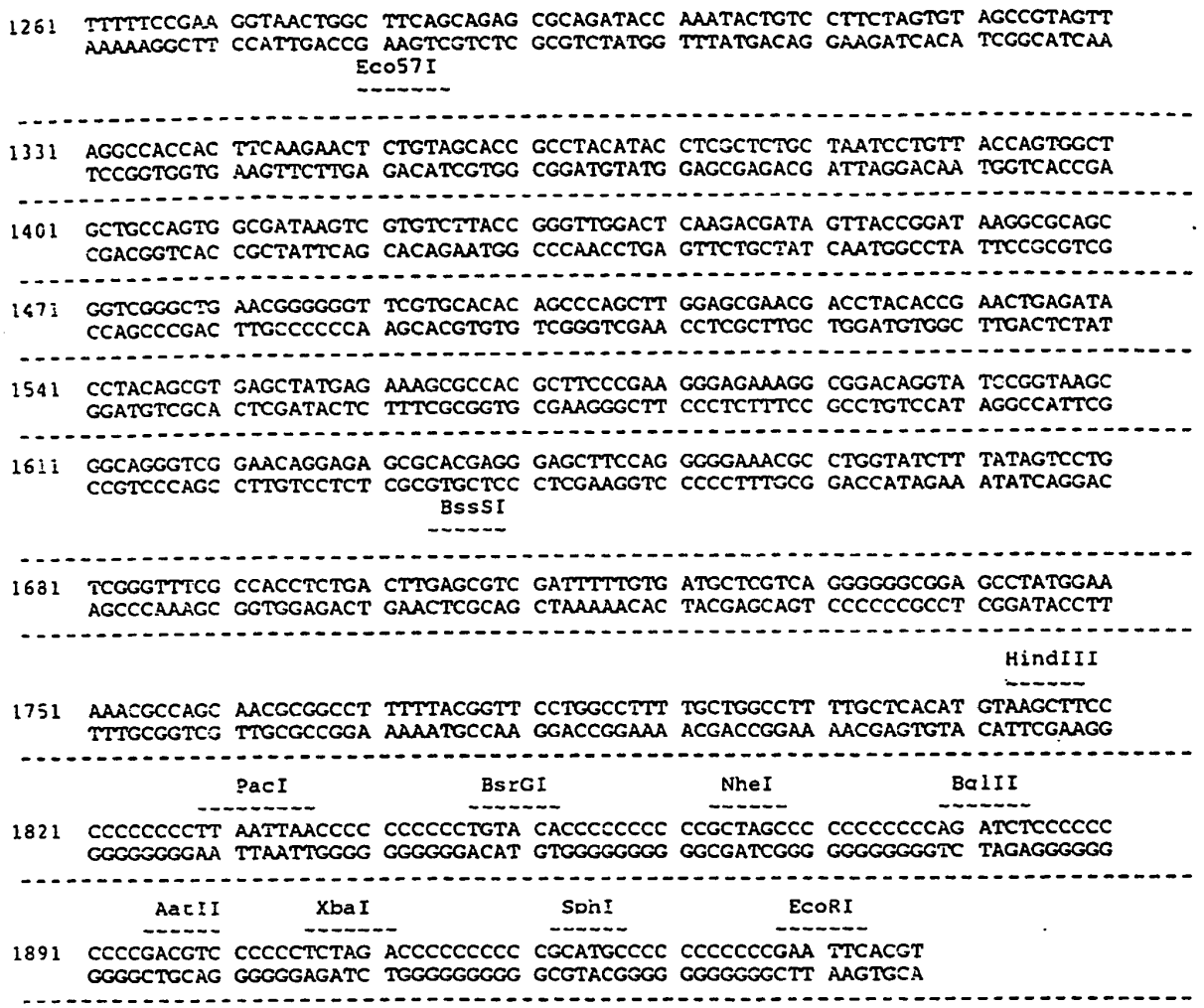
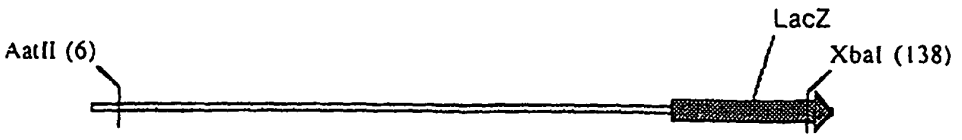

1261 TTTTTCCGAA GGTAACTGGC TTCAGCAGAG CGCAGATACC AAATACTGTC CTTCTAGTGT AGCCGTAGTT
AAAAAGGCTT CCATTGACCG AAGTCGTCTC GCGTCTATGG TTTATGACAG GAAGATCACA TCGGCATCAA
Eco57I
1331 AGGCCACCAC TTCAAGAACT CTGTAGCACC GCCTACATAC CTCGCTCTGC TAATCCTGTT ACCAGTGGCT
TCCGGTGGTG AAGTTCTTGA GACATCGTGG CGGATGTATG GAGCGAGACG ATTAGGACAA TGGTCACCGA
1401 GCTGCCAGTG GCGATAAGTC GTGTCTTACC GGGTTGGACT CAAGACGATA GTTACCGGAT AAGGCGCAGC
CGACGGTCAC CGCTATTCAG CACAGAATGG CCCAACCTGA GTTCTGCTAT CAATGGCCTA TTCCGCGTCG
1471 GGTCGGGCTG AACGGGGGGT TCGTGCACAC AGCCCAGCTT GGAGCGAACG ACCTACACCG AACTGAGATA
CCAGCCCGAC TTGCCCCCCA AGCACGTGTG TCGGGTCGAA CCTCGCTTGC TGGATGTGGC TTGACTCTAT
1541 CCTACAGCGT GAGCTATGAG AAAGCGCCAC GCTTCCCGAA GGGAGAAAGG CGGACAGGTA TCCGGTAAGC
GGATGTCGCA CTCGATACTC TTTCGCGGTG CGAAGGGCTT CCCTCTTTCC GCCTGTCCAT AGGCCATTCG
1611 GGCAGGGTCG GAACAGGAGA GCGCACGAGG GAGCTTCCAG GGGGAAACGC CTGGTATCTT TATAGTCCTG
CCGTCCCAGC CTTGTCCTCT CGCGTGCTCC CTCGAAGGTC CCCCTTTGCG GACCATAGAA ATATCAGGAC
BssSI
1681 TCGGGTTTCG CCACCTCTGA CTTGAGCGTC GATTTTTGTG ATGCTCGTCA GGGGGGCGGA GCCTATGGAA
AGCCCAAAGC GGTGGAGACT GAACTCGCAG CTAAAAACAC TACGAGCAGT CCCCCCGCCT CGGATACCTT
HindIII
1751 AAACGCCAGC AACGCGGCCT TTTTACGGTT CCTGGCCTTT TGCTGGCCTT TTGCTCACAT GTAAGCTTCC
TTTGCGGTCG TTGCGCCGGA AAAATGCCAA GGACCGGAAA ACGACCGGAA AACGAGTGTA CATTCGAAGG
PacI BsrGI NheI BglII
1821 CCCCCCCCTT AATTAACCCC CCCCCCTGTA CACCCCCCCC CCGCTAGCCC CCCCCCCCAG ATCTCCCCCC
GGGGGGGGAA TTAATTGGGG GGGGGGACAT GTGGGGGGGG GGCGATCGGG GGGGGGGGTC TAGAGGGGGG
AatII XbaI SphI EcoRI
1891 CCCCGACGTC CCCCCTCTAG ACCCCCCCCC CGCATGCCCC CCCCCCCGAA TTCACGT
GGGGCTGCAG GGGGGAGATC TGGGGGGGGG GCGTACGGGG GGGGGGGCTT AAGTGCA

Figure 29: functional map and sequence of pCAL module M1



M1
142 bp

M1

```
      AatII
      ------
  1   GACGTCTTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT
      CTGCAGAATT ACACTCAATC GAGTGAGTAA TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA
-------------------------------------------------------------------------------
                                                                          XbaI
                                                                          ----
 71   GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGATTA CGAATTTCTA
      CAACACACCT TAACACTCGC CTATTGTTAA AGTGTGTCCT TTGTCGATAC TGGTACTAAT GCTTAAAGAT
-------------------------------------------------------------------------------
      Xb
      --
141   GA
      CT
-------------------------------------------------------------------------------
```

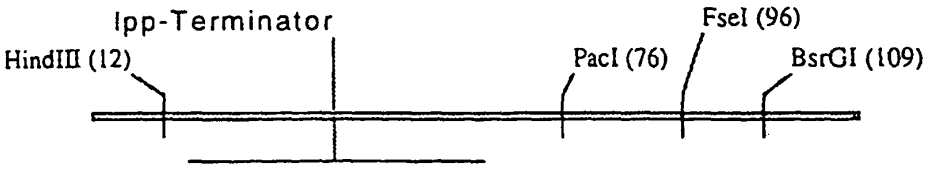
Figure 30: functional map and sequence of pCAL module M7-II

M7-III (ss/myc/TAG)
520 bp

M7-III (ss/myc/TAG)

```
      EcoRI
      ------
  1   GAATTCGAGC AGAAGCTGAT CTCTGAGGAG GATCTGTAGG GTGGTGGCTC TGGTTCCGGT GATTTTGATT
      CTTAAGCTCG TCTTCGACTA GAGACTCCTC CTAGACATCC CACCACCGAG ACCAAGGCCA CTAAAACTAA
-----------------------------------------------------------------------------------
 71   ATGAAAAGAT GGCAAACGCT AATAAGGGGG CTATGACCGA AAATGCCGAT GAAAACGCGC TACAGTCTGA
      TACTTTTCTA CCGTTTGCGA TTATTCCCCC GATACTGGCT TTTACGGCTA CTTTTGCGCG ATGTCAGACT
-----------------------------------------------------------------------------------
141   CGCTAAAGGC AAACTTGATT CTGTCGCTAC TGATTACGGT GCTGCTATCG ATGGTTTCAT TGGTGACGTT
      GCGATTTCCG TTTGAACTAA GACAGCGATG ACTAATGCCA CGACGATAGC TACCAAAGTA ACCACTGCAA
-----------------------------------------------------------------------------------
211   TCCGGCCTTG CTAATGGTAA TGGTGCTACT GGTGATTTTG CTGGCTCTAA TTCCCAAATG GCTCAAGTCG
      AGGCCGGAAC GATTACCATT ACCACGATGA CCACTAAAAC GACCGAGATT AAGGGTTTAC CGAGTTCAGC
-----------------------------------------------------------------------------------
                                  XmnI
                                  ------------
281   GTGACGGTGA TAATTCACCT TTAATGAATA ATTTCCGTCA ATATTTACCT TCCCTCCCTC AATCGGTTGA
      CACTGCCACT ATTAAGTGGA AATTACTTAT TAAAGGCAGT TATAAATGGA AGGGAGGGAG TTAGCCAACT
-----------------------------------------------------------------------------------
351   ATGTCGCCCT TTTGTCTTTG GCGCTGGTAA ACCATATGAA TTTTCTATTG ATTGTGACAA AATAAACTTA
      TACAGCGGGA AAACAGAAAC CGCGACCATT TGGTATACTT AAAAGATAAC TAACACTGTT TTATTTGAAT
-----------------------------------------------------------------------------------
421   TTCCGTGGTG TCTTTGCGTT TCTTTTATAT GTTGCCACCT TTATGTATGT ATTTTCTACG TTTGCTAACA
      AAGGCACCAC AGAAACGCAA AGAAAATATA CAACGGTGGA AATACATACA TAAAAGATGC AAACGATTGT
-----------------------------------------------------------------------------------
                                HindIII
                                ------
491   TACTGCGTAA TAAGGAGTCT TGATAAGCTT
      ATGACGCATT ATTCCTCAGA ACTATTCGAA
-----------------------------------------------------------------------------------
```

Figure 31: functional map and sequence of pCAL module M9-II

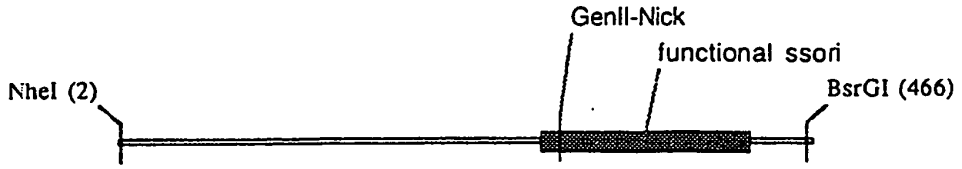


M9-II
123 bp

M9-II

```
               HindIII
               -------
  1  GGGGGGGGGG AAGCTTGACC TGTGAAGTGA AAAATGGCGC AGATTGTGCG ACATTTTTTT TGTCTGCCGT
     CCCCCCCCCC TTCGAACTGG ACACTTCACT TTTTACCGCG TCTAACACGC TGTAAAAAAA ACAGACGGCA
-------------------------------------------------------------------------------
       PacI                FseI              BsrGI
       --------            ---------         -------
 71  TTAATTAAAG GGGGGGGGGG GCCGGCCTGG GGGGGGGTGT ACAGGGGGGG GGG
     AATTAATTTC CCCCCCCCCC CGGCCGGACC CCCCCCCACA TGTCCCCCCC CCC
-------------------------------------------------------------------------------
```

Figure 32: functional map and sequence of pCAL module M11-III

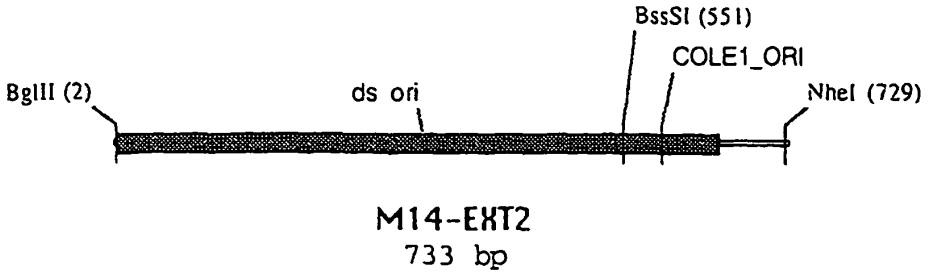


M11-III
470 bp

M11-III

```
     NheI
     ------
  1  GCTAGCACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG TGACCGCTAC
     CGATCGTGCG CGGGACATCG CCGCGTAATT CGCGCCGCCC ACACCACCAA TGCGCGTCGC ACTGGCGATG
-------------------------------------------------------------------------------
 71  ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT
     TGAACGGTCG CGGGATCGCG GGCGAGGAAA GCGAAAGAAG GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA
-------------------------------------------------------------------------------
141  CCCCGTCAAG CTCTAAATCG GGGCATCCCT TTAGGGTTCC GATTTAGTGC TTTACGGCAC CTCGACCCCA
     GGGGCAGTTC GAGATTTAGC CCCGTAGGGA AATCCCAAGG CTAAATCACG AAATGCCGTG GAGCTGGGGT
-------------------------------------------------------------------------------
211  AAAAACTTGA TTAGGGTGAT GGTTCTCGTA GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC
     TTTTTGAACT AATCCCACTA CCAAGAGCAT CACCCGGTAG CGGGACTATC TGCCAAAAAG CGGGAAACTG
-------------------------------------------------------------------------------
281  GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC
     CAACCTCAGG TGCAAGAAAT TATCACCTGA GAACAAGGTT TGACCTTGTT GTGAGTTGGG ATAGAGCCAG
-------------------------------------------------------------------------------
351  TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA AAATGAGCTG ATTTAACAAA
     ATAAGAAAAC TAAATATTCC CTAAAACGGC TAAAGCCGGA TAACCAATTT TTTACTCGAC TAAATTGTTT
-------------------------------------------------------------------------------
                                                    BsrGI
                                                    ------
421  AATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT TTCATGTACA
     TTAAATTGCG CTTAAAATTG TTTTATAATT GCAAATGTTA AAGTACATGT
-------------------------------------------------------------------------------
```

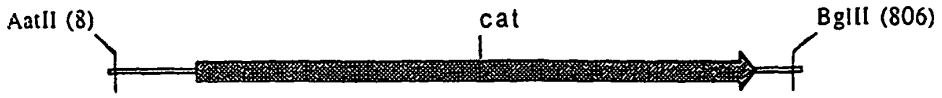
Figure 33: functional map and sequence of pCAL module M14-Ext2



M14-EXT2

```
      BglII
      ------
  1   AGATCTGACC AAAATCCCTT AACGTGAGTT TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA AAAGATCAAA
      TCTAGACTGG TTTTAGGGAA TTGCACTCAA AAGCAAGGTG ACTCGCAGTC TGGGGCATCT TTTCTAGTTT
-----------------------------------------------------------------------------
 71   GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC AAAAAAACCA CCGCTACCAG
      CCTAGAAGAA CTCTAGGAAA AAAAGACGCG CATTAGACGA CGAACGTTTG TTTTTTTGGT GGCGATGGTC
-----------------------------------------------------------------------------
141   CGGTGGTTTG TTTGCCGGAT CAAGAGCTAC CAACTCTTTT TCCGAAGGTA ACTGGCTACA GCAGAGCGCA
      GCCACCAAAC AAACGGCCTA GTTCTCGATG GTTGAGAAAA AGGCTTCCAT TGACCGATGT CGTCTCGCGT
-----------------------------------------------------------------------------
211   GATACCAAAT ACTGTTCTTC TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT
      CTATGGTTTA TGACAAGAAG ATCACATCGG CATCAATCCG GTGGTGAAGT TCTTGAGACA TCGTGGCGGA
-----------------------------------------------------------------------------
281   ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA TAAGTCGTGT CTTACCGGGT
      TGTATGGAGC GAGACGATTA GGACAATGGT CACCGACGAC GGTCACCGCT ATTCAGCACA GAATGGCCCA
-----------------------------------------------------------------------------
351   TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC
      ACCTGAGTTC TGCTATCAAT GGCCTATTCC GCGTCGCCAG CCCGACTTGC CCCCCAAGCA CGTGTGTCGG
-----------------------------------------------------------------------------
421   CAGCTTGGAG CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC TATGAGAAAG CGCCACGCTT
      GTCGAACCTC GCTTGCTGGA TGTGGCTTGA CTCTATGGAT GTCGCACTCG ATACTCTTTC GCGGTGCGAA
-----------------------------------------------------------------------------
491   CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC
      GGGCTTCCCT CTTTCCGCCT GTCCATAGGC CATTCGCCGT CCCAGCCTTG TCCTCTCGCG TGCTCCCTCG
                                                              BssSI
                                                              -------
-----------------------------------------------------------------------------
561   TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT
      AAGGTCCCCC TTTGCGGACC ATAGAAATAT CAGGACAGCC CAAAGCGGTG GAGACTGAAC TCGCAGCTAA
-----------------------------------------------------------------------------
631   TTTGTGATGC TCGTCAGGGG GGCGGAGCCT ATGGAAAAAC GCCAGCAACG CGGCCTTTTT ACGGTTCCTG
      AAACACTACG AGCAGTCCCC CCGCCTCGGA TACCTTTTTG CGGTCGTTGC GCCGGAAAAA TGCCAAGGAC
-----------------------------------------------------------------------------
                                  NheI
                                  -------
701   GCCTTTTGCT GGCCTTTTGC TCACATGGCT AGC
      CGGAAAACGA CCGGAAAACG AGTGTACCGA TCG
-----------------------------------------------------------------------------
```

Figure 34: functional map and sequence of pCAL module M17

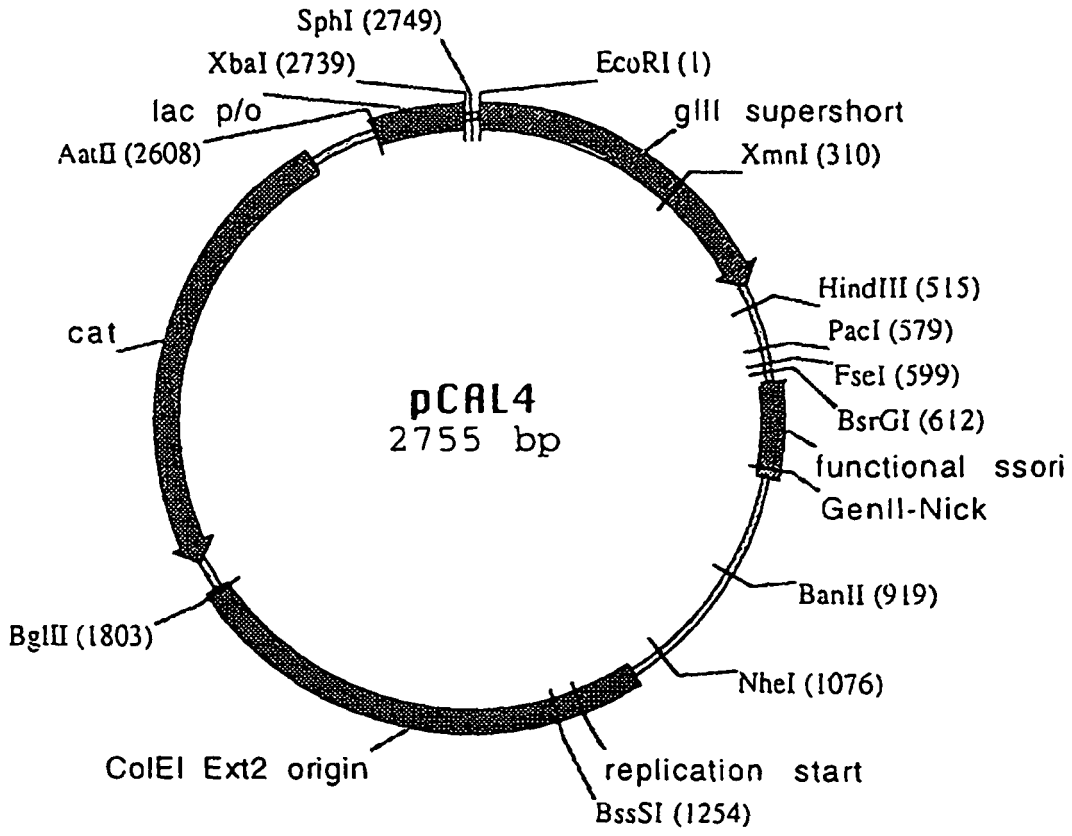


M17
813 bp

M17

```
       AatII
       ------
  1  GGGACGTCGG GTGAGGTTCC AACTTTCACC ATAATGAAAT AAGATCACTA CCGGGCGTAT TTTTTGAGTT
     CCCTGCAGCC CACTCCAAGG TTGAAAGTGG TATTACTTTA TTCTAGTGAT GGCCCGCATA AAAAACTCAA
-----------------------------------------------------------------------------
 71  ATCGAGATTT TCAGGAGCTA AGGAAGCTAA AATGGAGAAA AAAATCACTG GATATACCAC CGTTGATATA
     TAGCTCTAAA AGTCCTCGAT TCCTTCGATT TTACCTCTTT TTTTAGTGAC CTATATGGTG GCAACTATAT
-----------------------------------------------------------------------------
141  TCCCAATGGC ATCGTAAAGA ACATTTTGAG GCATTTCAGT CAGTTGCTCA ATGTACCTAT AACCAGACCG
     AGGGTTACCG TAGCATTTCT TGTAAAACTC CGTAAAGTCA GTCAACGAGT TACATGGATA TTGGTCTGGC
-----------------------------------------------------------------------------
211  TTCAGCTGGA TATTACGGCC TTTTTAAAGA CCGTAAAGAA AAATAAGCAC AAGTTTTATC CGGCCTTTAT
     AAGTCGACCT ATAATGCCGG AAAAATTTCT GGCATTTCTT TTTATTCGTG TTCAAAATAG GCCGGAAATA
-----------------------------------------------------------------------------
281  TCACATTCTT GCCCGCCTGA TGAATGCTCA CCCGGAGTTC CGTATGGCAA TGAAAGACGG TGAGCTGGTG
     AGTGTAAGAA CGGGCGGACT ACTTACGAGT GGGCCTCAAG GCATACCGTT ACTTTCTGCC ACTCGACCAC
-----------------------------------------------------------------------------
351  ATATGGGATA GTGTTCACCC TTGTTACACC GTTTTCCATG AGCAAACTGA AACGTTTTCA TCGCTCTGGA
     TATACCCTAT CACAAGTGGG AACAATGTGG CAAAAGGTAC TCGTTTGACT TTGCAAAAGT AGCGAGACCT
-----------------------------------------------------------------------------
421  GTGAATACCA CGACGATTTC CGGCAGTTTC TACACATATA TTCGCAAGAT GTGGCGTGTT ACGGTGAAAA
     CACTTATGGT GCTGCTAAAG GCCGTCAAAG ATGTGTATAT AAGCGTTCTA CACCGCACAA TGCCACTTTT
-----------------------------------------------------------------------------
491  CCTGGCCTAT TTCCCTAAAG GGTTTATTGA GAATATGTTT TTCGTCTCAG CCAATCCCTG GGTGAGTTTC
     GGACCGGATA AAGGGATTTC CCAAATAACT CTTATACAAA AAGCAGAGTC GGTTAGGGAC CCACTCAAAG
-----------------------------------------------------------------------------
561  ACCAGTTTTG ATTTAAACGT AGCCAATATG GACAACTTCT TCGCCCCCGT TTTCACTATG GGCAAATATT
     TGGTCAAAAC TAAATTTGCA TCGGTTATAC CTGTTGAAGA AGCGGGGGCA AAAGTGATAC CCGTTTATAA
-----------------------------------------------------------------------------
631  ATACGCAAGG CGACAAGGTG CTGATGCCGC TGGCGATTCA GGTTCATCAT GCCGTTTGTG ATGGCTTCCA
     TATGCGTTCC GCTGTTCCAC GACTACGGCG ACCGCTAAGT CCAAGTAGTA CGGCAAACAC TACCGAAGGT
-----------------------------------------------------------------------------
701  TGTCGGCAGA ATGCTTAATG AATTACAACA GTACTGCGAT GAGTGGCAGG GCGGGGCGTA ATTTTTTTAA
     ACAGCCGTCT TACGAATTAC TTAATGTTGT CATGACGCTA CTCACCGTCC CGCCCCGCAT TAAAAAAATT
-----------------------------------------------------------------------------
                                        BglII
                                        ------
771  GGCAGTTATT GGGTGCCCTT AAACGCCTGG TGCTAGATCT TCC
     CCGTCAATAA CCCACGGGAA TTTGCGGACC ACGATCTAGA AGG
-----------------------------------------------------------------------------
```

Figure 35: functional map and sequence of modular vector pCAL4



pCAL4

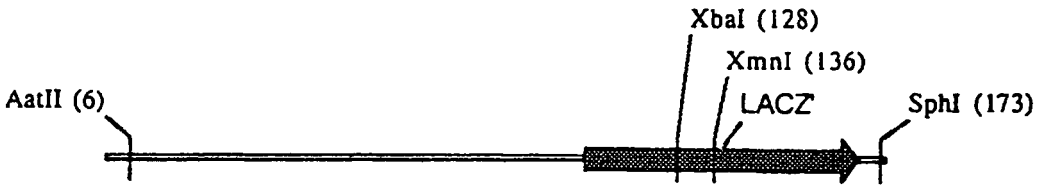
```
      EcoRI
      -----
   1  AATTCGAGCA GAAGCTGATC TCTGAGGAGG ATCTGTAGGG TGGTGGCTCT GGTTCCGGTG ATTTTGATTA
      TTAAGCTCGT CTTCGACTAG AGACTCCTCC TAGACATCCC ACCACCGAGA CCAAGGCCAC TAAAACTAAT
---------------------------------------------------------------------------------
  71  TGAAAAGATG GCAAACGCTA ATAAGGGGGC TATGACCGAA AATGCCGATG AAAACGCGCT ACAGTCTGAC
      ACTTTTCTAC CGTTTGCGAT TATTCCCCCG ATACTGGCTT TTACGGCTAC TTTTGCGCGA TGTCAGACTG
---------------------------------------------------------------------------------
 141  GCTAAAGGCA AACTTGATTC TGTCGCTACT GATTACGGTG CTGCTATCGA TGGTTTCATT GGTGACGTTT
      CGATTTCCGT TTGAACTAAG ACAGCGATGA CTAATGCCAC GACGATAGCT ACCAAAGTAA CCACTGCAAA
---------------------------------------------------------------------------------
 211  CCGGCCTTGC TAATGGTAAT GGTGCTACTG GTGATTTTGC TGGCTCTAAT TCCCAAATGG CTCAAGTCGG
      GGCCGGAACG ATTACCATTA CCACGATGAC CACTAAAACG ACCGAGATTA AGGGTTTACC GAGTTCAGCC
---------------------------------------------------------------------------------
                                  XmnI
                                  ----------
 281  TGACGGTGAT AATTCACCTT TAATGAATAA TTTCCGTCAA TATTTACCTT CCCTCCCTCA ATCGGTTGAA
      ACTGCCACTA TTAAGTGGAA ATTACTTATT AAAGGCAGTT ATAAATGGAA GGGAGGGAGT TAGCCAACTT
---------------------------------------------------------------------------------
 351  TGTCGCCCTT TTGTCTTTGG CGCTGGTAAA CCATATGAAT TTTCTATTGA TTGTGACAAA ATAAACTTAT
      ACAGCGGGAA AACAGAAACC GCGACCATTT GGTATACTTA AAAGATAACT AACACTGTTT TATTTGAATA
---------------------------------------------------------------------------------
 421  TCCGTGGTGT CTTTGCGTTT CTTTTATATG TTGCCACCTT TATGTATGTA TTTTCTACGT TTGCTAACAT
      AGGCACCACA GAAACGCAAA GAAAATATAC AACGGTGGAA ATACATACAT AAAAGATGCA AACGATTGTA
---------------------------------------------------------------------------------
                               HindIII
                               ------
 491  ACTGCGTAAT AAGGAGTCTT GATAAGCTTG ACCTGTGAAG TGAAAAATGG CGCAGATTGT GCGACATTTT
      TGACGCATTA TTCCTCAGAA CTATTCGAAC TGGACACTTC ACTTTTTACC GCGTCTAACA CGCTGTAAAA
---------------------------------------------------------------------------------
                    PacI                FseI              BsrGI
                    ---------           --------          ------
 561  TTTTGTCTGC CGTTTAATTA AAGGGGGGGG GGGGCCGGCC TGGGGGGGGG TGTACATGAA ATTGTAAACG
      AAAACAGACG GCAAATTAAT TTCCCCCCCC CCCCGGCCGG ACCCCCCCCC ACATGTACTT TAACATTTGC
---------------------------------------------------------------------------------
 631  TTAATATTTT GTTAAAATTC GCGTTAAATT TTTGTTAAAT CAGCTCATTT TTTAACCAAT AGGCCGAAAT
      AATTATAAAA CAATTTTAAG CGCAATTTAA AAACAATTTA GTCGAGTAAA AAATTGGTTA TCCGGCTTTA
---------------------------------------------------------------------------------
 701  CGGCAAAATC CCTTATAAAT CAAAAGAATA GACCGAGATA GGGTTGAGTG TTGTTCCAGT TTGGAACAAG
      GCCGTTTTAG GGAATATTTA GTTTTCTTAT CTGGCTCTAT CCCAACTCAC AACAAGGTCA AACCTTGTTC
---------------------------------------------------------------------------------
 771  AGTCCACTAT TAAAGAACGT GGACTCCAAC GTCAAAGGGC GAAAAACCGT CTATCAGGGC GATGGCCCAC
      TCAGGTGATA ATTTCTTGCA CCTGAGGTTG CAGTTTCCCG CTTTTTGGCA GATAGTCCCG CTACCGGGTG
---------------------------------------------------------------------------------
 841  TACGAGAACC ATCACCCTAA TCAAGTTTTT TGGGGTCGAG GTGCCGTAAA GCACTAAATC GGAACCCTAA
      ATGCTCTTGG TAGTGGGATT AGTTCAAAAA ACCCCAGCTC CACGGCATTT CGTGATTTAG CCTTGGGATT
---------------------------------------------------------------------------------
        BanII
        ------
 911  AGGGAGCCCC CGATTTAGAG CTTGACGGGG AAAGCCGGCG AACGTGGCGA GAAAGGAAGG GAAGAAAGCG
      TCCCTCGGGG GCTAAATCTC GAACTGCCCC TTTCGGCCGC TTGCACCGCT CTTTCCTTCC CTTCTTTCGC
---------------------------------------------------------------------------------
 981  AAAGGAGCGG GCGCTAGGGC GCTGGCAAGT GTAGCGGTCA CGCTGCGCGT AACCACCACA CCCGCCGCGC
      TTTCCTCGCC CGCGATCCCG CGACCGTTCA CATCGCCAGT GCGACGCGCA TTGGTGGTGT GGGCGGCGCG
---------------------------------------------------------------------------------
                                  NheI
                                  ------
1051  TTAATGCGCC GCTACAGGGC GCGTGCTAGC CATGTGAGCA AAAGGCCAGC AAAAGGCCAG GAACCGTAAA
      AATTACGCGG CGATGTCCCG CGCACGATCG GTACACTCGT TTTCCGGTCG TTTTCCGGTC CTTGGCATTT
---------------------------------------------------------------------------------
1121  AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA
      TTCCGGCGCA ACGACCGCAA AAAGGTATCC GAGGCGGGGG GACTGCTCGT AGTGTTTTTA GCTGCGAGTT
---------------------------------------------------------------------------------
```

pCAL4

pCAL4

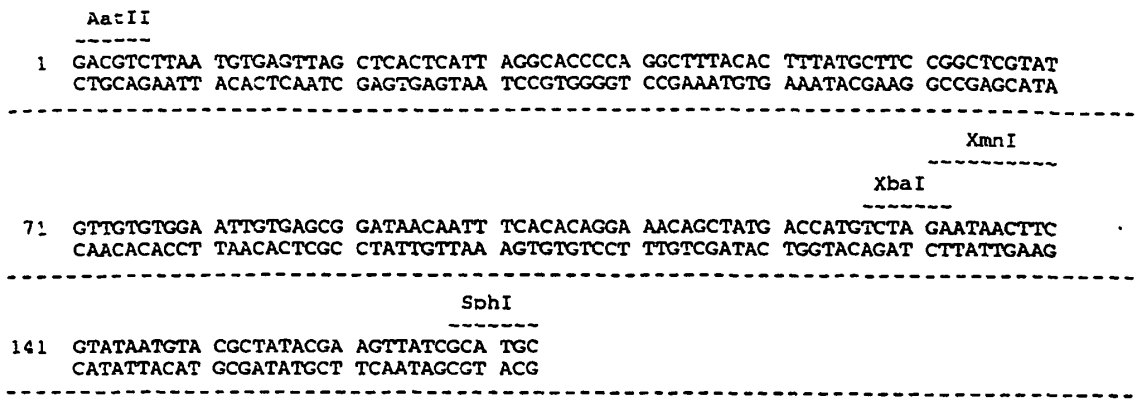
```
2521  TAGCTCCTGA AAATCTCGAT AACTCAAAAA ATACGCCCGG TAGTGATCTT ATTTCATTAT GGTGAAAGTT
      ATCGAGGACT TTTAGAGCTA TTGAGTTTTT TATGCGGGCC ATCACTAGAA TAAAGTAATA CCACTTTCAA

                   AatII
                   ------
2591  GGAACCTCAC CCGACGTCTA ATGTGAGTTA GCTCACTCAT TAGGCACCCC AGGCTTTACA CTTTATGCTT
      CCTTGGAGTG GGCTGCAGAT TACACTCAAT CGAGTGAGTA ATCCGTGGGG TCCGAAATGT GAAATACGAA

2661  CCGGCTCGTA TGTTGTGTGG AATTGTGAGC GGATAACAAT TTCACACAGG AAACAGCTAT GACCATGATT
      GGCCGAGCAT ACAACACACC TTAACACTCG CCTATTGTTA AAGTGTGTCC TTTGTCGATA CTGGTACTAA

             XbaI   SphI   EcoRI
             ------------     -
2731  ACGAATTTCT AGAGCATGCG GGGGG
      TGCTTAAAGA TCTCGTACGC CCCCC
```

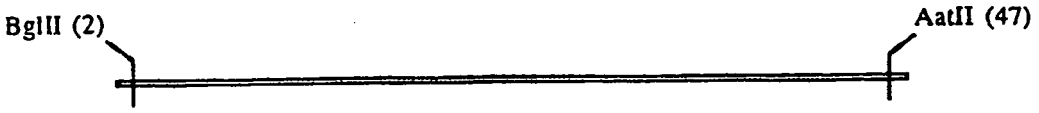
Figure 35a: Functional maps and sequences of additional pCAL vector modules and pCAL vectors



M2
173 bp

M2 Page 1

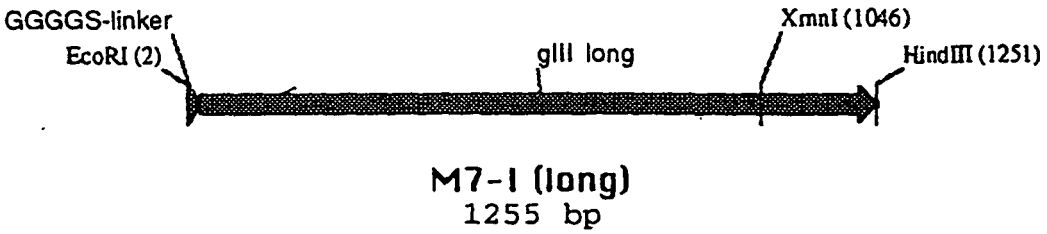

AatII
1 GACGTCTTAA TGTGAGTTAG CTCACTCATT AGGCACCCCA GGCTTTACAC TTTATGCTTC CGGCTCGTAT
CTGCAGAATT ACACTCAATC GAGTGAGTAA TCCGTGGGGT CCGAAATGTG AAATACGAAG GCCGAGCATA
XmnI
XbaI
71 GTTGTGTGGA ATTGTGAGCG GATAACAATT TCACACAGGA AACAGCTATG ACCATGTCTA GAATAACTTC
CAACACACCT TAACACTCGC CTATTGTTAA AGTGTGTCCT TTGTCGATAC TGGTACAGAT CTTATTGAAG
SphI
141 GTATAATGTA CGCTATACGA AGTTATCGCA TGC
CATATTACAT GCGATATGCT TCAATAGCGT ACG



M3
47 bp

M3 Page 1

```
     BglII                                              AatII
     ------                                             ------
  1  AGATCTCATA ACTTCGTATA ATGTATGCTA TACGAAGTTA TGACGTC
     TCTAGAGTAT TGAAGCATAT TACATACGAT ATGCTTCAAT ACTGCAG
```



M7-1 (long)
1255 bp

M7-1 (long) Page 1

```
      EcoRI
      ------
   1  GAATTCGGTG GTGGTGGATC TGCGTGCGCT GAAACGGTTG AAAGTTGTTT AGCAAAATCC CATACAGAAA
      CTTAAGCCAC CACCACCTAG ACGCACGCGA CTTTGCCAAC TTTCAACAAA TCGTTTTAGG GTATGTCTTT
---------------------------------------------------------------------------------
  71  ATTCATTTAC TAACGTCTGG AAAGACGACA AAACTTTAGA TCGTTACGCT AACTATGAGG GCTGTCTGTG
      TAAGTAAATG ATTGCAGACC TTTCTGCTGT TTTGAAATCT AGCAATGCGA TTGATACTCC CGACAGACAC
---------------------------------------------------------------------------------
 141  GAATGCTACA GGCGTTGTAG TTTGTACTGG TGACGAAACT CAGTGTTACG GTACATGGGT TCCTATTGGG
      CTTACGATGT CCGCAACATC AAACATGACC ACTGCTTTGA GTCACAATGC CATGTACCCA AGGATAACCC
---------------------------------------------------------------------------------
 211  CTTGCTATCC CTGAAAATGA GGGTGGTGGC TCTGAGGGTG GCGGTTCTGA GGGTGGCGGT TCTGAGGGTG
      GAACGATAGG GACTTTTACT CCCACCACCG AGACTCCCAC CGCCAAGACT CCCACCGCCA AGACTCCCAC
---------------------------------------------------------------------------------
 281  GCGGTACTAA ACCTCCTGAG TACGGTGATA CACCTATTCC GGGCTATACT TATATCAACC CTCTCGACGG
      CGCCATGATT TGGAGGACTC ATGCCACTAT GTGGATAAGG CCCGATATGA ATATAGTTGG GAGAGCTGCC
---------------------------------------------------------------------------------
 351  CACTTATCCG CCTGGTACTG AGCAAAACCC CGCTAATCCT AATCCTTCTC TTGAGGAGTC TCAGCCTCTT
      GTGAATAGGC GGACCATGAC TCGTTTTGGG GCGATTAGGA TTAGGAAGAG AACTCCTCAG AGTCGGAGAA
---------------------------------------------------------------------------------
 421  AATACTTTCA TGTTTCAGAA TAATAGGTTC CGAAATAGGC AGGGGGCATT AACTGTTTAT ACGGGCACTG
      TTATGAAAGT ACAAAGTCTT ATTATCCAAG GCTTTATCCG TCCCCCGTAA TTGACAAATA TGCCCGTGAC
---------------------------------------------------------------------------------
 491  TTACTCAAGG CACTGACCCC GTTAAAACTT ATTACCAGTA CACTCCTGTA TCATCAAAAG CCATGTATGA
      AATGAGTTCC GTGACTGGGG CAATTTTGAA TAATGGTCAT GTGAGGACAT AGTAGTTTTC GGTACATACT
---------------------------------------------------------------------------------
 561  CGCTTACTGG AACGGTAAAT TCAGAGACTG CGCTTTCCAT TCTGGCTTTA ATGAGGATTT ATTTGTTTGT
      GCGAATGACC TTGCCATTTA AGTCTCTGAC GCGAAAGGTA AGACCGAAAT TACTCCTAAA TAAACAAACA
---------------------------------------------------------------------------------
 631  GAATATCAAG GCCAATCGTC TGACCTGCCT CAACCTCCTG TCAATGCTGG CGGCGGCTCT GGTGGTGGTT
      CTTATAGTTC CGGTTAGCAG ACTGGACGGA GTTGGAGGAC AGTTACGACC GCCGCCGAGA CCACCACCAA
---------------------------------------------------------------------------------
 701  CTGGTGGCGG CTCTGAGGGT GGTGGCTCTG AGGGTGGCGG TTCTGAGGGT GGCGGCTCTG AGGGAGGCGG
      GACCACCGCC GAGACTCCCA CCACCGAGAC TCCCACCGCC AAGACTCCCA CCGCCGAGAC TCCCTCCGCC
---------------------------------------------------------------------------------
 771  TTCCGGTGGT GGCTCTGGTT CCGGTGATTT TGATTATGAA AAGATGGCAA ACGCTAATAA GGGGGCTATG
      AAGGCCACCA CCGAGACCAA GGCCACTAAA ACTAATACTT TTCTACCGTT TGCGATTATT CCCCCGATAC
---------------------------------------------------------------------------------
 841  ACCGAAAATG CCGATGAAAA CGCGCTACAG TCTGACGCTA AAGGCAAACT TGATTCTGTC GCTACTGATT
      TGGCTTTTAC GGCTACTTTT GCGCGATGTC AGACTGCGAT TTCCGTTTGA ACTAAGACAG CGATGACTAA
---------------------------------------------------------------------------------
 911  ACGGTGCTGC TATCGATGGT TTCATTGGTG ACGTTTCCGG CCTTGCTAAT GGTAATGGTG CTACTGGTGA
      TGCCACGACG ATAGCTACCA AAGTAACCAC TGCAAAGGCC GGAACGATTA CCATTACCAC GATGACCACT
---------------------------------------------------------------------------------
                                                                        XmnI
                                                                        ----------
 981  TTTTGCTGGC TCTAATTCCC AAATGGCTCA AGTCGGTGAA GGTGATAATT CACCTTTAAT GAATAATTTC
      AAAACGACCG AGATTAAGGG TTTACCGAGT TCAGCCACTT CCACTATTAA GTGGAAATTA CTTATTAAAG
---------------------------------------------------------------------------------
1051  CGTCAATATT TACCTTCCAT CCCTCAATCG GTTGAATGTC GCCCTTTTGT CTTTGGCGCT GGTAAACCCT
      GCAGTTATAA ATGGAAGGTA GGGAGTTAGC CAACTTACAG CGGGAAAACA GAAACCGCGA CCATTTGGGA
---------------------------------------------------------------------------------
1121  ATGAATTTTC TATTGATTGT GACAAAATAA ACTTATTCCG TGGTGTCTTT GCGTTTCTTT TATATGTTGC
      TACTTAAAAG ATAACTAACA CTGTTTTATT TGAATAAGGC ACCACAGAAA CGCAAAGAAA ATATACAACG
---------------------------------------------------------------------------------
                                                                  HindIII
                                                                  -------
1191  CACCTTTATG TATGTATTTT CTACGTTTGC TAACATACTG CGTAATAAGG AGTCTTGATA AGCTT
      GTGGAAATAC ATACATAAAA GATGCAAACG ATTGTATGAC GCATTATTCC TCAGAACTAT TCGAA
---------------------------------------------------------------------------------
```

M7-11 (ss-TAG)
502 bp

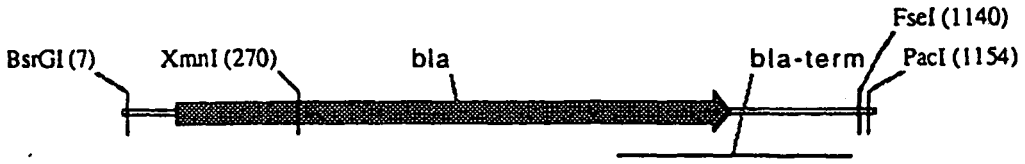
M7-11 (ss-TAG) Page 1

```
        EcoRI
        ------
  1  CGGGAATTCG GAGGCGGTTC CGGTGGTGGC TCTGGTTCCG GTGATTTTGA TTATGAAAAG ATGGCAAACG
     GCCCTTAAGC CTCCGCCAAG GCCACCACCG AGACCAAGGC CACTAAAACT AATACTTTTC TACCGTTTGC
-----------------------------------------------------------------------------
 71  CTAATAAGGG GGCTATGACC GAAAATGCCG ATGAAAACGC GCTACAGTCT GACGCTAAAG GCAAACTTGA
     GATTATTCCC CCGATACTGG CTTTTACGGC TACTTTTGCG CGATGTCAGA CTGCGATTTC CGTTTGAACT
-----------------------------------------------------------------------------
141  TTCTGTCGCT ACTGATTACG GTGCTGCTAT CGATGGTTTC ATTGGTGACG TTTCCGGCCT TGCTAATGGT
     AAGACAGCGA TGACTAATGC CACGACGATA GCTACCAAAG TAACCACTGC AAAGGCCGGA ACGATTACCA
-----------------------------------------------------------------------------
211  AATGGTGCTA CTGGTGATTT TGCTGGCTCT AATTCCCAAA TGGCTCAAGT CGGTGACGGT GATAATTCAC
     TTACCACGAT GACCACTAAA ACGACCGAGA TTAAGGGTTT ACCGAGTTCA GCCACTGCCA CTATTAAGTG
-----------------------------------------------------------------------------
              XmnI
             -----------
281  CTTTAATGAA TAATTTCCGT CAATATTTAC CTTCCCTCCC TCAATCGGTT GAATGTCGCC CTTTTGTCTT
     GAAATTACTT ATTAAAGGCA GTTATAAATG GAAGGGAGGG AGTTAGCCAA CTTACAGCGG GAAAACAGAA
-----------------------------------------------------------------------------
351  TGGCGCTGGT AAACCATATG AATTTTCTAT TGATTGTGAC AAAATAAACT TATTCCGTGG TGTCTTTGCG
     ACCGCGACCA TTTGGTATAC TTAAAAGATA ACTAACACTG TTTTATTTGA ATAAGGCACC ACAGAAACGC
-----------------------------------------------------------------------------
421  TTTCTTTTAT ATGTTGCCAC CTTTATGTAT GTATTTTCTA CGTTTGCTAA CATACTGCGT AATAAGGAGT
     AAAGAAAATA TACAACGGTG GAAATACATA CATAAAAGAT GCAAACGATT GTATGACGCA TTATTCCTCA
-----------------------------------------------------------------------------
         HindIII
         --------
491  CTTGATAAGC TT
     GAACTATTCG AA
-----------------------------------------------------------------------------
```

M8
47 bp

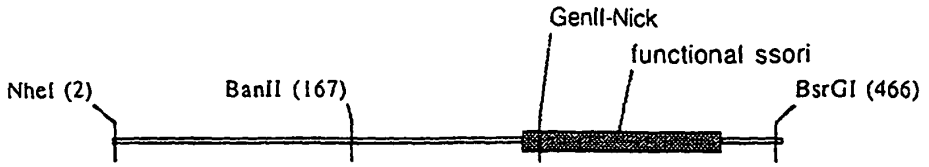
M8 Page 1

```
     SphI                                          HindIII
     ------                                         ------
  1  GCATGCCATA ACTTCGTATA ATGTACGCTA TACGAAGTTA TAAGCTT
     CGTACGGTAT TGAAGCATAT TACATGCGAT ATGCTTCAAT ATTCGAA
```



M10-11
1163 bp

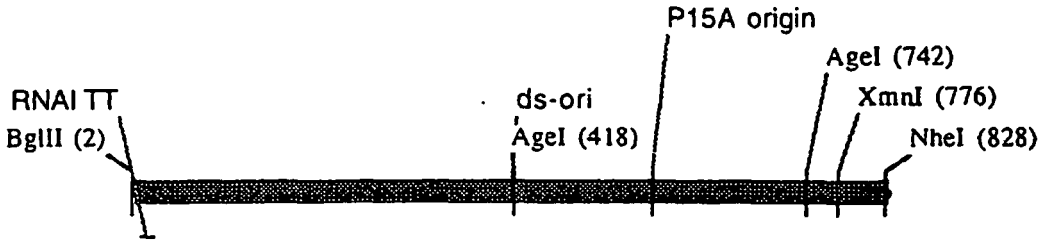
```
          BsrGI
          -------
   1  GGGGGTGTAC ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA TAATATTGAA
      CCCCCACATG TAAGTTTATA CATAGGCGAG TACTCTGTTA TTGGGACTAT TTACGAAGTT ATTATAACTT
---------------------------------------------------------------------------------
  71  AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT TATTCCCTTT TTTGCGGCAT TTTGCCTTCC
      TTTCCTTCTC ATACTCATAA GTTGTAAAGG CACAGCGGGA ATAAGGGAAA AAACGCCGTA AAACGGAAGG
---------------------------------------------------------------------------------
 141  TGTTTTTGCT CACCCAGAAA CGCTGGTGAA AGTAAAAGAT GCTGAGGATC AGTTGGGTGC GCGAGTGGGT
      ACAAAAACGA GTGGGTCTTT GCGACCACTT TCATTTTCTA CGACTCCTAG TCAACCCACG CGCTCACCCA
---------------------------------------------------------------------------------
                                                                    XmnI
                                                                    ----------
 211  TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC CGAAGAACGT TTTCCAATGA
      ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACTCT CAAAAGCGGG GCTTCTTGCA AAAGGTTACT
---------------------------------------------------------------------------------
 281  TGAGCACTTT TAAAGTTCTG CTATGTGGCG CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG
      ACTCGTGAAA ATTTCAAGAC GATACACCGC GCCATAATAG GGCATAACTG CGGCCCGTTC TCGTTGAGCC
---------------------------------------------------------------------------------
 351  TCGCCGCATA CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA TCTTACGGAT
      AGCGGCGTAT GTGATAAGAG TCTTACTGAA CCAACTCATG AGTGGTCAGT GTCTTTTCGT AGAATGCCTA
---------------------------------------------------------------------------------
 421  GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA TGAGTGATAA CACTGCGGCC AACTTACTTC
      CCGTACTGTC ATTCTCTTAA TACGTCACGA CGGTATTGGT ACTCACTATT GTGACGCCGG TTGAATGAAG
---------------------------------------------------------------------------------
 491  TGACAACGAT CGGAGGACCG AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT
      ACTGTTGCTA GCCTCCTGGC TTCCTCGATT GGCGAAAAAA CGTGTTGTAC CCCCTAGTAC ATTGAGCGGA
---------------------------------------------------------------------------------
 561  TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG ACACCACGAT GCCTGTAGCA
      ACTAGCAACC CTTGGCCTCG ACTTACTTCG GTATGGTTTG CTGCTCGCAC TGTGGTGCTA CGGACATCGT
---------------------------------------------------------------------------------
 631  ATGGCAACAA CGTTGCGCAA ACTATTAACT GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAGTTAATAG
      TACCGTTGTT GCAACGCGTT TGATAATTGA CCGCTTGATG AATGAGATCG AAGGGCCGTT GTCAATTATC
---------------------------------------------------------------------------------
 701  ACTGGATGGA GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT GGTTTATTGC
      TGACCTACCT CCGCCTATTT CAACGTCCTG GTGAAGACGC GAGCCGGGAA GGCCGACCGA CCAAATAACG
---------------------------------------------------------------------------------
 771  TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC
      ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG TAACGTCGTG ACCCCGGTCT ACCATTCGGG
---------------------------------------------------------------------------------
 841  TCCCGTATCG TAGTTATCTA CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
      AGGGCATAGC ATCAATAGAT GTGCTGCCCC TCAGTCCGTT GATACCTACT TGCTTTATCT GTCTAGCGAC
---------------------------------------------------------------------------------
 911  AGATAGGTGC CTCACTGATT AAGCATTGGG TAACTGTCAG ACCAAGTTTA CTCATATATA CTTTAGATTG
      TCTATCCACG GAGTGACTAA TTCGTAACCC ATTGACAGTC TGGTTCAAAT GAGTATATAT GAAATCTAAC
---------------------------------------------------------------------------------
 981  ATTTAAAACT TCATTTTTAA TTTAAAAGGA TCTAGGTGAA GATCCTTTTT GATAATCTCA TGACCAAAAT
      TAAATTTTGA AGTAAAAATT AAATTTTCCT AGATCCACTT CTAGGAAAAA CTATTAGAGT ACTGGTTTTA
---------------------------------------------------------------------------------
1051  CCCTTAACGT GAGTTTTCGT TCCACTGAGC GTCAGACCCC GTAGAAAAGA TCAAAGGATC TTCTTGAGAT
      GGGAATTGCA CTCAAAAGCA AGGTGACTCG CAGTCTGGGG CATCTTTTCT AGTTTCCTAG AAGAACTCTA
---------------------------------------------------------------------------------
                    FseI            PacI
                    ---------       ---------
1121  CCTTTTTGAT AATGGCCGGC CCCCCCCCTT AATTAAGGGG GGG
      GGAAAAACTA TTACCGGCCG GGGGGGGGAA TTAATTCCCC CCC
---------------------------------------------------------------------------------
```



M11-II
470 bp

M11-II

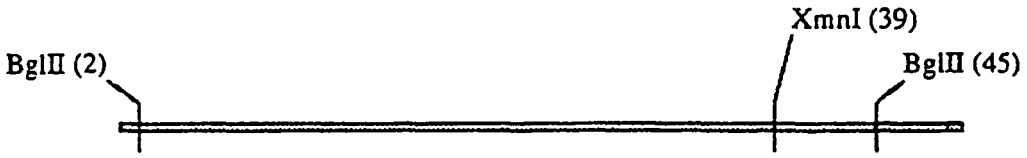
```
      NheI
      ------
  1   GCTAGCACGC GCCCTGTAGC GGCGCATTAA GCGCGGCGGG TGTGGTGGTT ACGCGCAGCG TGACCGCTAC
      CGATCGTGCG CGGGACATCG CCGCGTAATT CGCGCCGCCC ACACCACCAA TGCGCGTCGC ACTGGCGATG
-------------------------------------------------------------------------------
 71   ACTTGCCAGC GCCCTAGCGC CCGCTCCTTT CGCTTTCTTC CCTTCCTTTC TCGCCACGTT CGCCGGCTTT
      TGAACGGTCG CGGGATCGCG GGCGAGGAAA GCGAAAGAAG GGAAGGAAAG AGCGGTGCAA GCGGCCGAAA
-------------------------------------------------------------------------------
                            BanII
                            ------
141   CCCCGTCAAG CTCTAAATCG GGGGCTCCCT TTAGGGTTCC GATTTAGTGC TTTACGGCAC CTCGACCCCA
      GGGGCAGTTC GAGATTTAGC CCCCGAGGGA AATCCCAAGG CTAAATCACG AAATGCCGTG GAGCTGGGGT
-------------------------------------------------------------------------------
211   AAAAACTTGA TTAGGGTGAT GGTTCTCGTA GTGGGCCATC GCCCTGATAG ACGGTTTTTC GCCCTTTGAC
      TTTTTGAACT AATCCCACTA CCAAGAGCAT CACCCGGTAG CGGGACTATC TGCCAAAAAG CGGGAAACTG
-------------------------------------------------------------------------------
281   GTTGGAGTCC ACGTTCTTTA ATAGTGGACT CTTGTTCCAA ACTGGAACAA CACTCAACCC TATCTCGGTC
      CAACCTCAGG TGCAAGAAAT TATCACCTGA GAACAAGGTT TGACCTTGTT GTGAGTTGGG ATAGAGCCAG
-------------------------------------------------------------------------------
351   TATTCTTTTG ATTTATAAGG GATTTTGCCG ATTTCGGCCT ATTGGTTAAA AAATGAGCTG ATTTAACAAA
      ATAAGAAAAC TAAATATTCC CTAAAACGGC TAAAGCCGGA TAACCAATTT TTTACTCGAC TAAATTGTTT
-------------------------------------------------------------------------------
                                                      BsrGI
                                                      ------
421   AATTTAACGC GAATTTTAAC AAAATATTAA CGTTTACAAT TTCATGTACA
      TTAAATTGCG CTTAAAATTG TTTTATAATT GCAAATGTTA AAGTACATGT
-------------------------------------------------------------------------------
```


P15A origin
AgeI (742)
RNAI TT
BglII (2)
ds-ori
AgeI (418)
XmnI (776)
NheI (828)
1

M12
832 bp

M12 Page 1

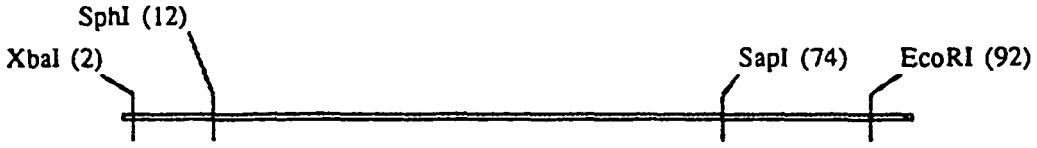
```
      BalII
      ------
  1   AGATCTAATA AGATGATCTT CTTGAGATCG TTTTGGTCTG CGCGTAATCT CTTGCTCTGA AAACGAAAAA
      TCTAGATTAT TCTACTAGAA GAACTCTAGC AAAACCAGAC GCGCATTAGA GAACGAGACT TTTGCTTTTT

 71   ACCGCCTTGC AGGGCGGTTT TTCGTAGGTT CTCTGAGCTA CCAACTCTTT GAACCGAGGT AACTGGCTTG
      TGGCGGAACG TCCCGCCAAA AAGCATCCAA GAGACTCGAT GGTTGAGAAA CTTGGCTCCA TTGACCGAAC

141   GAGGAGCGCA GTCACTAAAA CTTGTCCTTT CAGTTTAGCC TTAACCGGCG CATGACTTCA AGACTAACTC
      CTCCTCGCGT CAGTGATTTT GAACAGGAAA GTCAAATCGG AATTGGCCGC GTACTGAAGT TCTGATTGAG

211   CTCTAAATCA ATTACCAGTG GCTGCTGCCA GTGGTGCTTT TGCATGTCTT TCCGGGTTGG ACTCAAGACG
      GAGATTTAGT TAATGGTCAC CGACGACGGT CACCACGAAA ACGTACAGAA AGGCCCAACC TGAGTTCTGC

281   ATAGTTACCG GATAAGGCGC AGCGGTCGGA CTGAACGGGG GGTTCGTGCA TACAGTCCAG CTTGGAGCGA
      TATCAATGGC CTATTCCGCG TCGCCAGCCT GACTTGCCCC CCAAGCACGT ATGTCAGGTC GAACCTCGCT

                                                                          AgeI
                                                                          ----
351   ACTGCCTACC CGGAACTGAG TGTCAGGCGT GGAATGAGAC AAACGCGGCC ATAACAGCGG AATGACACCG
      TGACGGATGG GCCTTGACTC ACAGTCCGCA CCTTACTCTG TTTGCGCCGG TATTGTCGCC TTACTGTGGC

      AgeI
      --
421   GTAAACCGAA AGGCAGGAAC AGGAGAGCGC AGGAGGGAGC CGCCAGGGGG AAACGCCTGG TATCTTTATA
      CATTTGGCTT TCCGTCCTTG TCCTCTCGCG TCCTCCCTCG GCGGTCCCCC TTTGCGGACC ATAGAAATAT

491   GTCCTGTCGG GTTTCGCCAC CACTGATTTG AGCGTCAGAT TTCGTGATGC TTGTCAGGGG GGCGGAGCCT
      CAGGACAGCC CAAAGCGGTG GTGACTAAAC TCGCAGTCTA AAGCACTACG AACAGTCCCC CCGCCTCGGA

561   ATGGAAAAAC GGCTTTGCCG CGGCCCTCTC ACTTCCCTGT TAAGTATCTT CCTGGCATCT TCCAGGAAAT
      TACCTTTTTG CCGAAACGGC GCCGGGAGAG TGAAGGGACA ATTCATAGAA GGACCGTAGA AGGTCCTTTA

631   CTCCGCCCCG TTCGTAAGCC ATTTCCGCTC GCCGCAGTCG AACGACCGAG CGTAGCGAGT CAGTGAGCGA
      GAGGCGGGGC AAGCATTCGG TAAAGGCGAG CGGCGTCAGC TTGCTGGCTC GCATCGCTCA GTCACTCGCT

                                                  AgeI
                                                  ------
701   GGAAGCGGAA TATATCCTGT ATCACATATT CTGCTGACGC ACCGGTGCAG CCTTTTTTCT CCTGCCACAT
      CCTTCGCCTT ATATAGGACA TAGTGTATAA GACGACTGCG TGGCCACGTC GGAAAAAAGA GGACGGTGTA

        XmnI                                                      NheI
        ----------                                                --------
771   GAAGCACTTC ACTGACACCC TCATCAGTGC CAACATAGTA AGCCAGTATA CACTCCGCTA GC
      CTTCGTGAAG TGACTGTGGG AGTAGTCACG GTTGTATCAT TCGGTCATAT GTGAGGCGAT CG
```



M13
49 bp

M13 Page 1

```
     BglII                                  XmnI      BglII
     ------                                 ----------------
  1  AGATCTCATA ACTTCGTATA ATGTATGCTA TACGAAGTTA TTCAGATCT
     TCTAGAGTAT TGAAGCATAT TACATACGAT ATGCTTCAAT AAGTCTAGA
```



M19
96 bp

M19 Page1

```
     XbaI  SphI                                                          SapI
     ------------                                                        -----
  1  TCTAGAGCAT GCGTAGGAGA AAATAAAATG AAACAAAGCA CTATTGCACT GGCACTCTTA CCGTTGCTCT
     AGATCTCGTA CGCATCCTCT TTTATTTTAC TTTGTTTCGT GATAACGTGA CCGTGAGAAT GGCAACGAGA

     SapI                    EcoRI
     --                      ------
 71  TCACCCCTGT TACCAAAGCC GAATTC
     AGTGGGGACA ATGGTTTCGG CTTAAG
```

M20
120 bp

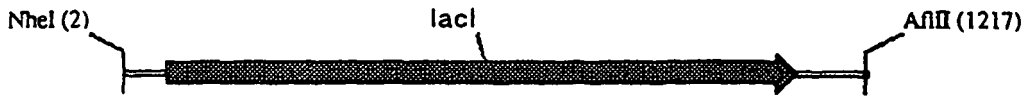
M20 Page 1

```
     XbaI  SphI                                                                SapI
     ------------                                                              -----
  1  TCTAGAGCAT GCGTAGGAGA AAATAAAATG AAACAAAGCA CTATTGCACT GGCACTCTTA CCGTTGCTCT
     AGATCTCGTA CGCATCCTCT TTTATTTTAC TTTGTTTCGT GATAACGTGA CCGTGAGAAT GGCAACGAGA

     SapI                                      MunI   EcoRI
     --                                        ------------
 71  TCACCCCTGT TACCAAAGCC GACTACAAAG ATGAAGTGCA ATTGGAATTC
     AGTGGGGACA ATGGTTTCGG CTGATGTTTC TACTTCACGT TAACCTTAAG
```

M21
96 bp

M21 Page 1

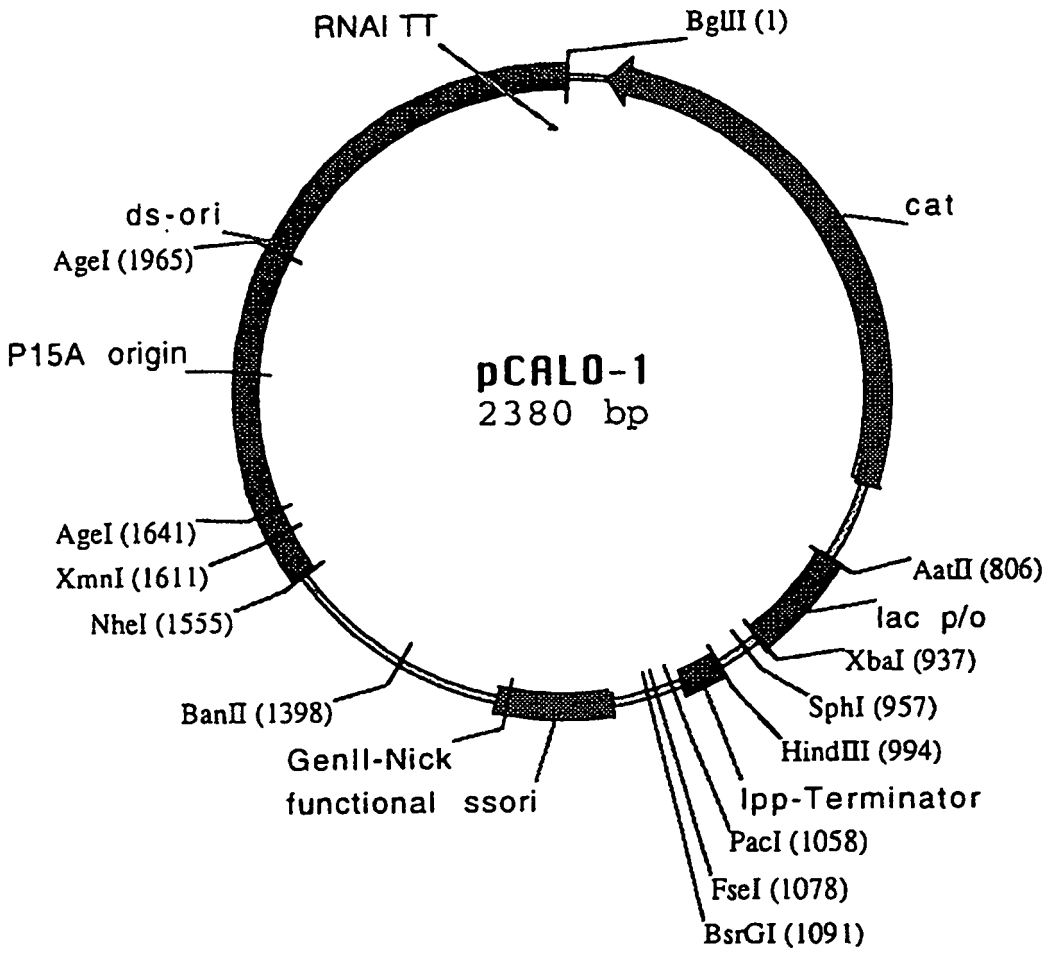
```
     XbaI
     ------
  1  TCTAGAGGTT GAGGTGATTT TATGAAAAAG AATATCGCAT TTCTTCTTGC ATCTATGTTC GTTTTTTCTA
     AGATCTCCAA CTCCACTAAA ATACTTTTTC TTATAGCGTA AAGAAGAACG TAGATACAAG CAAAAAAGAT
-------------------------------------------------------------------------------
                NsiI        EcoRI
               -------      ------
 71  TTGCTACAAA TGCATACGCT GAATTC
     AACGATGTTT ACGTATGCGA CTTAAG
-------------------------------------------------------------------------------
```



M41
1221 bp

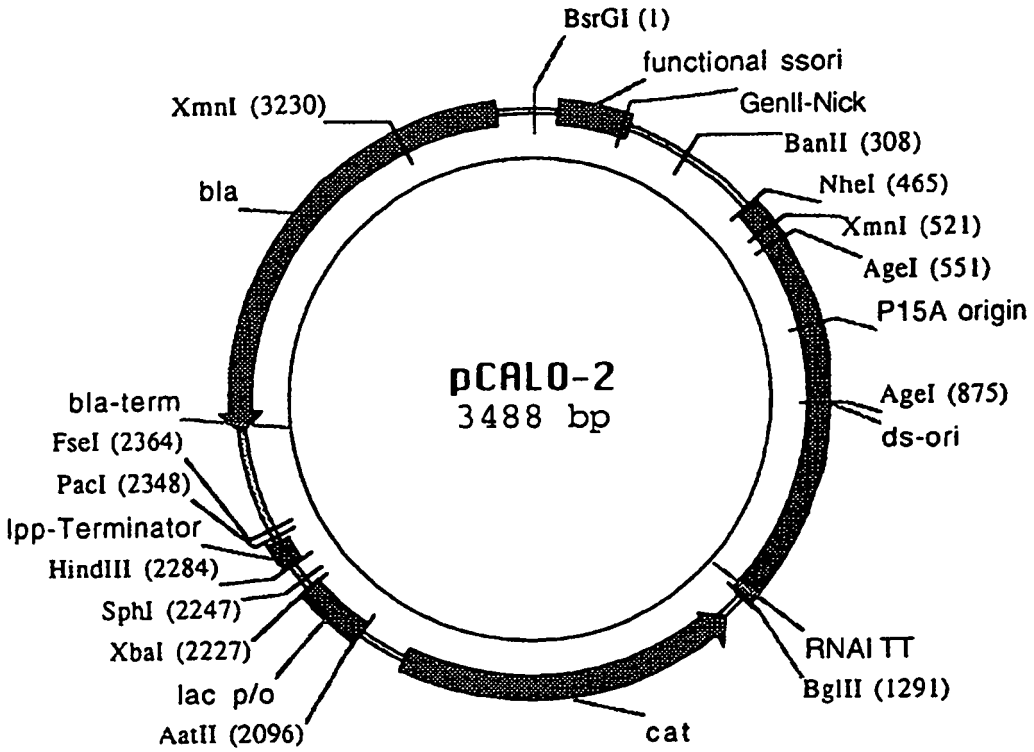
M41 Page 1

```
      NheI
      ------
   1  GCTAGCATCG AATGGCGCAA AACCTTTCGC GGTATGGCAT GATAGCGCCC GGAAGAGAGT CAATTCAGGG
      CGATCGTAGC TTACCGCGTT TTGGAAAGCG CCATACCGTA CTATCGCGGG CCTTCTCTCA GTTAAGTCCC
---------------------------------------------------------------------------------
  71  TGGTGAATGT GAAACCAGTA ACGTTATACG ATGTCGCAGA GTATGCCGGT GTCTCTTATC AGACCGTTTC
      ACCACTTACA CTTTGGTCAT TGCAATATGC TACAGCGTCT CATACGGCCA CAGAGAATAG TCTGGCAAAG
---------------------------------------------------------------------------------
 141  CCGCGTGGTG AACCAGGCCA GCCACGTTTC TGCGAAAACG CGGGAAAAAG TGGAAGCGGC GATGGCGGAG
      GGCGCACCAC TTGGTCCGGT CGGTGCAAAG ACGCTTTTGC GCCCTTTTTC ACCTTCGCCG CTACCGCCTC
---------------------------------------------------------------------------------
 211  CTGAATTACA TTCCTAACCG CGTGGCACAA CAACTGGCGG GCAAACAGTC GTTGCTGATT GGCGTTGCCA
      GACTTAATGT AAGGATTGGC GCACCGTGTT GTTGACCGCC CGTTTGTCAG CAACGACTAA CCGCAACGGT
---------------------------------------------------------------------------------
 281  CCTCCAGTCT GGCCCTGCAC GCGCCGTCGC AAATTGTCGC GGCGATTAAA TCTCGCGCCG ATCAACTGGG
      GGAGGTCAGA CCGGGACGTG CGCGGCAGCG TTTAACAGCG CCGCTAATTT AGAGCGCGGC TAGTTGACCC
---------------------------------------------------------------------------------
 351  TGCCAGCGTG GTCGTGTCGA TGGTAGAACG AAGCGGCGTC GAAGCCTGTA AAGCGGCGGT GCACAATCTT
      ACGGTCGCAC CAGCACAGCT ACCATCTTGC TTCGCCGCAG CTTCGGACAT TTCGCCGCCA CGTGTTAGAA
---------------------------------------------------------------------------------
 421  CTCGCGCAAC GTGTCAGTGG GCTGATTATT AACTATCCGC TGGATGACCA GGATGCTATT GCTGTGGAAG
      GAGCGCGTTG CACAGTCACC CGACTAATAA TTGATAGGCG ACCTACTGGT CCTACGATAA CGACACCTTC
---------------------------------------------------------------------------------
 491  CTGCCTGCAC TAATGTTCCG GCGTTATTTC TTGATGTCTC TGACCAGACA CCCATCAACA GTATTATTTT
      GACGGACGTG ATTACAAGGC CGCAATAAAG AACTACAGAG ACTGGTCTGT GGGTAGTTGT CATAATAAAA
---------------------------------------------------------------------------------
 561  CTCCCATGAG GACGGTACGC GACTGGGCGT GGAGCATCTG GTCGCATTGG GCCACCAGCA AATCGCGCTG
      GAGGGTACTC CTGCCATGCG CTGACCCGCA CCTCGTAGAC CAGCGTAACC CGGTGGTCGT TTAGCGCGAC
---------------------------------------------------------------------------------
 631  TTAGCTGGCC CATTAAGTTC TGTCTCGGCG CGTCTGCGTC TGGCTGGCTG GCATAAATAT CTCACTCGCA
      AATCGACCGG GTAATTCAAG ACAGAGCCGC GCAGACGCAG ACCGACCGAC CGTATTTATA GAGTGAGCGT
---------------------------------------------------------------------------------
 701  ATCAAATTCA GCCGATAGCG GAACGGGAAG GCGACTGGAG TGCCATGTCC GGTTTTCAAC AAACCATGCA
      TAGTTTAAGT CGGCTATCGC CTTGCCCTTC CGCTGACCTC ACGGTACAGG CCAAAAGTTG TTTGGTACGT
---------------------------------------------------------------------------------
 771  AATGCTGAAT GAGGGCATCG TTCCCACTGC GATGCTGGTT GCCAACGATC AGATGGCGCT GGGCGCAATG
      TTACGACTTA CTCCCGTAGC AAGGGTGACG CTACGACCAA CGGTTGCTAG TCTACCGCGA CCCGCGTTAC
---------------------------------------------------------------------------------
 841  CGTGCCATTA CCGAGTCCGG GCTGCGCGTT GGTGCGGACA TCTCGGTAGT GGGATACGAC GATACCGAGG
      GCACGGTAAT GGCTCAGGCC CGACGCGCAA CCACGCCTGT AGAGCCATCA CCCTATGCTG CTATGGCTCC
---------------------------------------------------------------------------------
 911  ACAGCTCATG TTATATCCCG CCGCTGACCA CCATCAAACA GGATTTTCGC CTGCTGGGGC AAACCAGCGT
      TGTCGAGTAC AATATAGGGC GGCGACTGGT GGTAGTTTGT CCTAAAAGCG GACGACCCCG TTTGGTCGCA
---------------------------------------------------------------------------------
 981  GGACCGCTTG CTGCAACTCT CTCAGGGCCA GGCGGTGAAG GGCAATCAGC TGTTGCCCGT CTCACTGGTG
      CCTGGCGAAC GACGTTGAGA GAGTCCCGGT CCGCCACTTC CCGTTAGTCG ACAACGGGCA GAGTGACCAC
---------------------------------------------------------------------------------
1051  AAAAGAAAAA CCACCCTGGC TCCCAATACG CAAACCGCCT CTCCCCGCGC GTTGGCCGAT TCACTGATGC
      TTTTCTTTTT GGTGGGACCG AGGGTTATGC GTTTGGCGGA GAGGGGCGCG CAACCGGCTA AGTGACTACG
---------------------------------------------------------------------------------
1121  AGCTGGCACG ACAGGTTTCC CGACTGGAAA GCGGGCAGTG AGGCTACCCG ATAAAAGCGG CTTCCTGACA
      TCGACCGTGC TGTCCAAAGG GCTGACCTTT CGCCCGTCAC TCCGATGGGC TATTTTCGCC GAAGGACTGT
---------------------------------------------------------------------------------
                                  AflII
                                  -------
1191  GGAGGCCGTT TTGTTTTGCA GCCCACTTAA G
      CCTCCGGCAA AACAAAACGT CGGGTGAATT C
---------------------------------------------------------------------------------
```


RNAI TT
BglII (1)
cat
ds-ori
AgeI (1965)
P15A origin
pCALO-1
2380 bp
AgeI (1641)
XmnI (1611)
NheI (1555)
AatII (806)
lac p/o
XbaI (937)
SphI (957)
HindIII (994)
BanII (1398)
GenII-Nick
functional ssori
lpp-Terminator
PacI (1058)
FseI (1078)
BsrGI (1091)

```
     BglII
     -----
   1 GATCTAGCAC CAGGCGTTTA AGGGCACCAA TAACTGCCTT AAAAAAATTA CGCCCCGCCC TGCCACTCAT
     CTAGATCGTG GTCCGCAAAT TCCCGTGGTT ATTGACGGAA TTTTTTTAAT GCGGGGCGGG ACGGTGAGTA
------------------------------------------------------------------------------
  71 CGCAGTACTG TTGTAATTCA TTAAGCATTC TGCCGACATG GAAGCCATCA CAAACGGCAT GATGAACCTG
     GCGTCATGAC AACATTAAGT AATTCGTAAG ACGGCTGTAC CTTCGGTAGT GTTTGCCGTA CTACTTGGAC
------------------------------------------------------------------------------
 141 AATCGCCAGC GGCATCAGCA CCTTGTCGCC TTGCGTATAA TATTTGCCCA TAGTGAAAAC GGGGGCGAAG
     TTAGCGGTCG CCGTAGTCGT GGAACAGCGG AACGCATATT ATAAACGGGT ATCACTTTTG CCCCCGCTTC
------------------------------------------------------------------------------
 211 AAGTTGTCCA TATTGGCTAC GTTTAAATCA AAACTGGTGA AACTCACCCA GGGATTGGCT GAGACGAAAA
     TTCAACAGGT ATAACCGATG CAAATTTAGT TTTGACCACT TTGAGTGGGT CCCTAACCGA CTCTGCTTTT
------------------------------------------------------------------------------
 281 ACATATTCTC AATAAACCCT TTAGGGAAAT AGGCCAGGTT TTCACCGTAA CACGCCACAT CTTGCGAATA
     TGTATAAGAG TTATTTGGGA AATCCCTTTA TCCGGTCCAA AAGTGGCATT GTGCGGTGTA GAACGCTTAT
------------------------------------------------------------------------------
 351 TATGTGTAGA AACTGCCGGA AATCGTCGTG GTATTCACTC CAGAGCGATG AAAACGTTTC AGTTTGCTCA
     ATACACATCT TTGACGGCCT TTAGCAGCAC CATAAGTGAG GTCTCGCTAC TTTTGCAAAG TCAAACGAGT
------------------------------------------------------------------------------
 421 TGGAAAACGG TGTAACAAGG GTGAACACTA TCCCATATCA CCAGCTCACC GTCTTTCATT GCCATACGGA
     ACCTTTTGCC ACATTGTTCC CACTTGTGAT AGGGTATAGT GGTCGAGTGG CAGAAAGTAA CGGTATGCCT
------------------------------------------------------------------------------
 491 ACTCCGGGTG AGCATTCATC AGGCGGGCAA GAATGTGAAT AAAGGCCGGA TAAAACTTGT GCTTATTTTT
     TGAGGCCCAC TCGTAAGTAG TCCGCCCGTT CTTACACTTA TTTCCGGCCT ATTTTGAACA CGAATAAAAA
------------------------------------------------------------------------------
 561 CTTTACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG GTCTGGTTAT AGGTACATTG AGCAACTGAC
     GAAATGCCAG AAATTTTTCC GGCATTATAG GTCGACTTGC CAGACCAATA TCCATGTAAC TCGTTGACTG
------------------------------------------------------------------------------
 631 TGAAATGCCT CAAAATGTTC TTTACGATGC CATTGGGATA TATCAACGGT GGTATATCCA GTGATTTTTT
     ACTTTACGGA GTTTTACAAG AAATGCTACG GTAACCCTAT ATAGTTGCCA CCATATAGGT CACTAAAAAA
------------------------------------------------------------------------------
 701 TCTCCATTTT AGCTTCCTTA GCTCCTGAAA ATCTCGATAA CTCAAAAAAT ACGCCCGGTA GTGATCTTAT
     AGAGGTAAAA TCGAAGGAAT CGAGGACTTT TAGAGCTATT GAGTTTTTTA TGCGGGCCAT CACTAGAATA
------------------------------------------------------------------------------
                                     AatII
                                     ------
 771 TTCATTATGG TGAAAGTTGG AACCTCACCC GACGTCTAAT GTGAGTTAGC TCACTCATTA GGCACCCCAG
     AAGTAATACC ACTTTCAACC TTGGAGTGGG CTGCAGATTA CACTCAATCG AGTGAGTAAT CCGTGGGGTC
------------------------------------------------------------------------------
 841 GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA
     CGAAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT AACACTCGCC TATTGTTAAA GTGTGTCCTT
------------------------------------------------------------------------------
                                XbaI                SphI
                                -------             ------
 911 ACAGCTATGA CCATGATTAC GAATTTCTAG ACCCCCCCCC CGCATGCCAT AACTTCGTAT AATGTACGCT
     TGTCGATACT GGTACTAATG CTTAAAGATC TGGGGGGGGG GCGTACGGTA TTGAAGCATA TTACATGCGA
------------------------------------------------------------------------------
               HindIII
               -------
 981 ATACGAAGTT ATAAGCTTGA CCTGTGAAGT GAAAAATGGC GCAGATTGTG CGACATTTTT TTTGTCTGCC
     TATGCTTCAA TATTCGAACT GGACACTTCA CTTTTTACCG CGTCTAACAC GCTGTAAAAA AAACAGACGG
------------------------------------------------------------------------------
        PacI                FseI                BsrGI
        --------            --------            -------
1051 GTTTAATTAA AGGGGGGGGG GGGCCGGCCT GGGGGGGGGT GTACATGAAA TTGTAAACGT TAATATTTTG
     CAAATTAATT TCCCCCCCCC CCCGGCCGGA CCCCCCCCCA CATGTACTTT AACATTTGCA ATTATAAAAC
------------------------------------------------------------------------------
1121 TTAAAATTCG CGTTAAATTT TTGTTAAATC AGCTCATTTT TTAACCAATA GGCCGAAATC GGCAAAATCC
     AATTTTAAGC GCAATTTAAA AACAATTTAG TCGAGTAAAA AATTGGTTAT CCGGCTTTAG CCGTTTTAGG
------------------------------------------------------------------------------
1191 CTTATAAATC AAAAGAATAG ACCGAGATAG GGTTGAGTGT TGTTCCAGTT TGGAACAAGA GTCCACTATT
     GAATATTTAG TTTTCTTATC TGGCTCTATC CCAACTCACA ACAAGGTCAA ACCTTGTTCT CAGGTGATAA
------------------------------------------------------------------------------
```

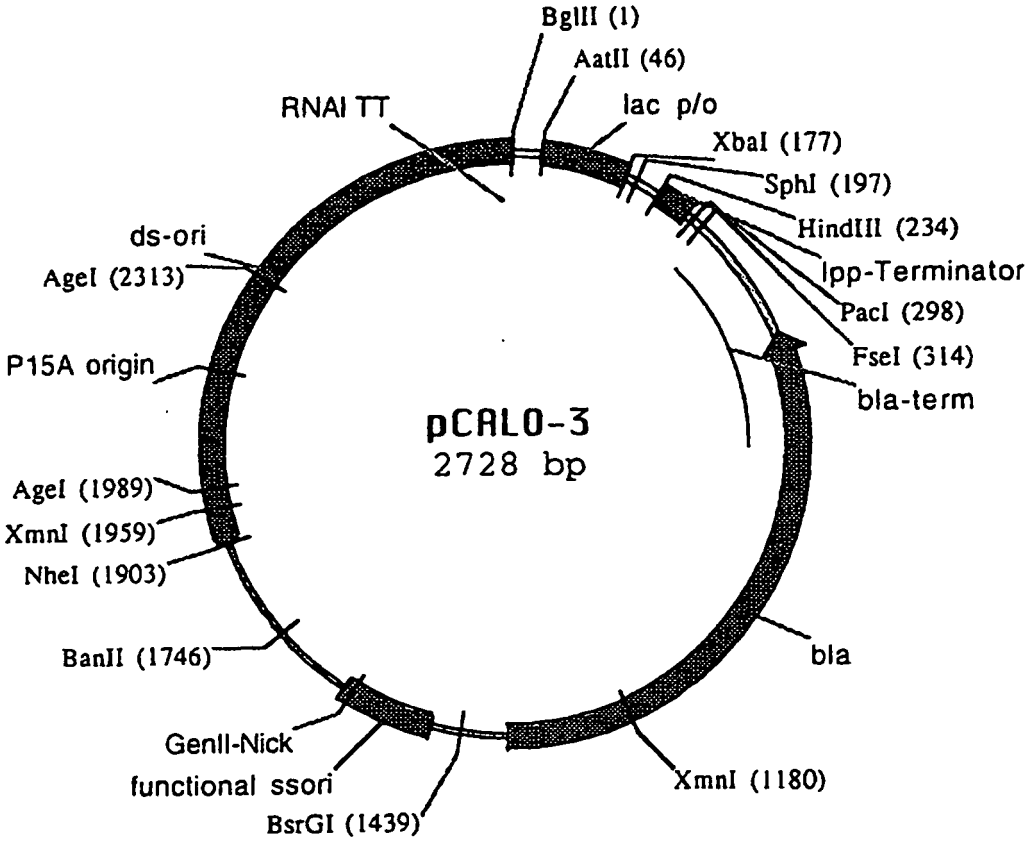
```
1261  AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC TATCAGGGCG ATGGCCCACT ACGAGAACCA
      TTTCTTGCAC CTGAGGTTGC AGTTTCCCGC TTTTTGGCAG ATAGTCCCGC TACCGGGTGA TGCTCTTGGT
-------------------------------------------------------------------------------
                                                                        BanII
                                                                        ------
1331  TCACCCTAAT CAAGTTTTTT GGGGTCGAGG TGCCGTAAAG CACTAAATCG GAACCCTAAA GGGAGCCCCC
      AGTGGGATTA GTTCAAAAAA CCCCAGCTCC ACGGCATTTC GTGATTTAGC CTTGGGATTT CCCTCGGGGG
-------------------------------------------------------------------------------
1401  GATTTAGAGC TTGACGGGGA AAGCCGGCGA ACGTGGCGAG AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG
      CTAAATCTCG AACTGCCCCT TTCGGCCGCT TGCACCGCTC TTTCCTTCCC TTCTTTCGCT TTCCTCGCCC
-------------------------------------------------------------------------------
1471  CGCTAGGGCG CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA ACCACCACAC CCGCCGCGCT TAATGCGCCG
      GCGATCCCGC GACCGTTCAC ATCGCCAGTG CGACGCGCAT TGGTGGTGTG GGCGGCGCGA ATTACGCGGC
-------------------------------------------------------------------------------
                    NheI                                                XmnI
                    ------                                              -----
1541  CTACAGGGCG CGTGCTAGCG GAGTGTATAC TGGCTTACTA TGTTGGCACT GATGAGGGTG TCAGTGAAGT
      GATGTCCCGC GCACGATCGC CTCACATATG ACCGAATGAT ACAACCGTGA CTACTCCCAC AGTCACTTCA
-------------------------------------------------------------------------------
      XmnI                          AgeI
      -----                         -------
1611  GCTTCATGTG GCAGGAGAAA AAAGGCTGCA CCGGTGCGTC AGCAGAATAT GTGATACAGG ATATATTCCG
      CGAAGTACAC CGTCCTCTTT TTTCCGACGT GGCCACGCAG TCGTCTTATA CACTATGTCC TATATAAGGC
-------------------------------------------------------------------------------
1681  CTTCCTCGCT CACTGACTCG CTACGCTCGG TCGTTCGACT GCGGCGAGCG GAAATGGCTT ACGAACGGGG
      GAAGGAGCGA GTGACTGAGC GATGCGAGCC AGCAAGCTGA CGCCGCTCGC CTTTACCGAA TGCTTGCCCC
-------------------------------------------------------------------------------
1751  CGGAGATTTC CTGGAAGATG CCAGGAAGAT ACTTAACAGG GAAGTGAGAG GGCCGCGGCA AAGCCGTTTT
      GCCTCTAAAG GACCTTCTAC GGTCCTTCTA TGAATTGTCC CTTCACTCTC CCGGCGCCGT TTCGGCAAAA
-------------------------------------------------------------------------------
1821  TCCATAGGCT CCGCCCCCCT GACAAGCATC ACGAAATCTG ACGCTCAAAT CAGTGGTGGC GAAACCCGAC
      AGGTATCCGA GGCGGGGGGA CTGTTCGTAG TGCTTTAGAC TGCGAGTTTA GTCACCACCG CTTTGGGCTG
-------------------------------------------------------------------------------
1891  AGGACTATAA AGATACCAGG CGTTTCCCCC TGGCGGCTCC CTCCTGCGCT CTCCTGTTCC TGCCTTTCGG
      TCCTGATATT TCTATGGTCC GCAAAGGGGG ACCGCCGAGG GAGGACGCGA GAGGACAAGG ACGGAAAGCC
-------------------------------------------------------------------------------
        AgeI
        ------
1961  TTTACCGGTG TCATTCCGCT GTTATGGCCG CGTTTGTCTC ATTCCACGCC TGACACTCAG TTCCGGGTAG
      AAATGGCCAC AGTAAGGCGA CAATACCGGC GCAAACAGAG TAAGGTGCGG ACTGTGAGTC AAGGCCCATC
-------------------------------------------------------------------------------
2031  GCAGTTCGCT CCAAGCTGGA CTGTATGCAC GAACCCCCCG TTCAGTCCGA CCGCTGCGCC TTATCCGGTA
      CGTCAAGCGA GGTTCGACCT GACATACGTG CTTGGGGGGC AAGTCAGGCT GGCGACGCGG AATAGGCCAT
-------------------------------------------------------------------------------
2101  ACTATCGTCT TGAGTCCAAC CCGGAAAGAC ATGCAAAAGC ACCACTGGCA GCAGCCACTG GTAATTGATT
      TGATAGCAGA ACTCAGGTTG GGCCTTTCTG TACGTTTTCG TGGTGACCGT CGTCGGTGAC CATTAACTAA
-------------------------------------------------------------------------------
2171  TAGAGGAGTT AGTCTTGAAG TCATGCGCCG GTTAAGGCTA AACTGAAAGG ACAAGTTTTA GTGACTGCGC
      ATCTCCTCAA TCAGAACTTC AGTACGCGGC CAATTCCGAT TTGACTTTCC TGTTCAAAAT CACTGACGCG
-------------------------------------------------------------------------------
2241  TCCTCCAAGC CAGTTACCTC GGTTCAAAGA GTTGGTAGCT CAGAGAACCT ACGAAAAACC GCCCTGCAAG
      AGGAGGTTCG GTCAATGGAG CCAAGTTTCT CAACCATCGA GTCTCTTGGA TGCTTTTTGG CGGGACGTTC
-------------------------------------------------------------------------------
                                                                        BglII
                                                                        -
2311  GCGGTTTTTT CGTTTTCAGA GCAAGAGATT ACGCGCAGAC CAAAACGATC TCAAGAAGAT CATCTTATTA
      CGCCAAAAAA GCAAAAGTCT CGTTCTCTAA TGCGCGTCTG GTTTTGCTAG AGTTCTTCTA GTAGAATAAT
-------------------------------------------------------------------------------
```


BsrGI (1)
functional ssori
GenII-Nick
BanII (308)
NheI (465)
XmnI (521)
AgeI (551)
P15A origin
AgeI (875)
ds-ori
RNAI TT
BglII (1291)
cat
AatII (2096)
lac p/o
XbaI (2227)
SphI (2247)
HindIII (2284)
lpp-Terminator
PacI (2348)
FseI (2364)
bla-term
bla
XmnI (3230)
pCALO-2
3488 bp

```
     BsrGI
     -----
  1  GTACATGAAA TTGTAAACGT TAATATTTTG TTAAAATTCG CGTTAAATTT TTGTTAAATC AGCTCATTTT
     CATGTACTTT AACATTTGCA ATTATAAAAC AATTTTAAGC GCAATTTAAA AACAATTTAG TCGAGTAAAA
-----------------------------------------------------------------------------
 71  TTAACCAATA GGCCGAAATC GGCAAAATCC CTTATAAATC AAAAGAATAG ACCGAGATAG GGTTGAGTGT
     AATTGGTTAT CCGGCTTTAG CCGTTTTAGG GAATATTTAG TTTTCTTATC TGGCTCTATC CCAACTCACA
-----------------------------------------------------------------------------
141  TGTTCCAGTT TGGAACAAGA GTCCACTATT AAAGAACGTG GACTCCAACG TCAAAGGGCG AAAAACCGTC
     ACAAGGTCAA ACCTTGTTCT CAGGTGATAA TTTCTTGCAC CTGAGGTTGC AGTTTCCCGC TTTTTGGCAG
-----------------------------------------------------------------------------
211  TATCAGGGCG ATGGCCCACT ACGAGAACCA TCACCCTAAT CAAGTTTTTT GGGGTCGAGG TGCCGTAAAG
     ATAGTCCCGC TACCGGGTGA TGCTCTTGGT AGTGGGATTA GTTCAAAAAA CCCCAGCTCC ACGGCATTTC
-----------------------------------------------------------------------------
                            BanII
                            ------
281  CACTAAATCG GAACCCTAAA GGGAGCCCCC GATTTAGAGC TTGACGGGGA AAGCCGGCGA ACGTGGCGAG
     GTGATTTAGC CTTGGGATTT CCCTCGGGGG CTAAATCTCG AACTGCCCCT TTCGGCCGCT TGCACCGCTC
-----------------------------------------------------------------------------
351  AAAGGAAGGG AAGAAAGCGA AAGGAGCGGG CGCTAGGGCG CTGGCAAGTG TAGCGGTCAC GCTGCGCGTA
     TTTCCTTCCC TTCTTTCGCT TTCCTCGCCC GCGATCCCGC GACCGTTCAC ATCGCCAGTG CGACGCGCAT
-----------------------------------------------------------------------------
                                                  NheI
                                                  ------
421  ACCACCACAC CCGCCGCGCT TAATGCGCCG CTACAGGGCG CGTGCTAGCG GAGTGTATAC TGGCTTACTA
     TGGTGGTGTG GGCGGCGCGA ATTACGCGGC GATGTCCCGC GCACGATCGC CTCACATATG ACCGAATGAT
-----------------------------------------------------------------------------
                                  XmnI                                  AgeI
                                  ----------                            -------
491  TGTTGGCACT GATGAGGGTG TCAGTGAAGT GCTTCATGTG GCAGGAGAAA AAAGGCTGCA CCGGTGCGTC
     ACAACCGTGA CTACTCCCAC AGTCACTTCA CGAAGTACAC CGTCCTCTTT TTTCCGACGT GGCCACGCAG
-----------------------------------------------------------------------------
561  AGCAGAATAT GTGATACAGG ATATATTCCG CTTCCTCGCT CACTGACTCG CTACGCTCGG TCGTTCGACT
     TCGTCTTATA CACTATGTCC TATATAAGGC GAAGGAGCGA GTGACTGAGC GATGCGAGCC AGCAAGCTGA
-----------------------------------------------------------------------------
631  GCGGCGAGCG GAAATGGCTT ACGAACGGGG CGGAGATTTC CTGGAAGATG CCAGGAAGAT ACTTAACAGG
     CGCCGCTCGC CTTTACCGAA TGCTTGCCCC GCCTCTAAAG GACCTTCTAC GGTCCTTCTA TGAATTGTCC
-----------------------------------------------------------------------------
701  GAAGTGAGAG GGCCGCGGCA AAGCCGTTTT TCCATAGGCT CCGCCCCCCT GACAAGCATC ACGAAATCTG
     CTTCACTCTC CCGGCGCCGT TTCGGCAAAA AGGTATCCGA GGCGGGGGGA CTGTTCGTAG TGCTTTAGAC
-----------------------------------------------------------------------------
771  ACGCTCAAAT CAGTGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC TGGCGGCTCC
     TGCGAGTTTA GTCACCACCG CTTTGGGCTG TCCTGATATT TCTATGGTCC GCAAAGGGGG ACCGCCGAGG
-----------------------------------------------------------------------------
                                  AgeI
                                  ------
841  CTCCTGCGCT CTCCTGTTCC TGCCTTTCGG TTTACCGGTG TCATTCCGCT GTTATGGCCG CGTTTGTCTC
     GAGGACGCGA GAGGACAAGG ACGGAAAGCC AAATGGCCAC AGTAAGGCGA CAATACCGGC GCAAACAGAG
-----------------------------------------------------------------------------
911  ATTCCACGCC TGACACTCAG TTCCGGGTAG GCAGTTCGCT CCAAGCTGGA CTGTATGCAC GAACCCCCCG
     TAAGGTGCGG ACTGTGAGTC AAGGCCCATC CGTCAAGCGA GGTTCGACCT GACATACGTG CTTGGGGGGC
-----------------------------------------------------------------------------
981  TTCAGTCCGA CCGCTGCGCC TTATCCGGTA ACTATCGTCT TGAGTCCAAC CCGGAAAGAC ATGCAAAAGC
     AAGTCAGGCT GGCGACGCGG AATAGGCCAT TGATAGCAGA ACTCAGGTTG GGCCTTTCTG TACGTTTTCG
-----------------------------------------------------------------------------
1051 ACCACTGGCA GCAGCCACTG GTAATTGATT TAGAGGAGTT AGTCTTGAAG TCATGCGCCG GTTAAGGCTA
     TGGTGACCGT CGTCGGTGAC CATTAACTAA ATCTCCTCAA TCAGAACTTC AGTACGCGGC CAATTCCGAT
-----------------------------------------------------------------------------
1121 AACTGAAAGG ACAAGTTTTA GTGACTGCGC TCCTCCAAGC CAGTTACCTC GGTTCAAAGA GTTGGTAGCT
     TTGACTTTCC TGTTCAAAAT CACTGACGCG AGGAGGTTCG GTCAATGGAG CCAAGTTTCT CAACCATCGA
-----------------------------------------------------------------------------
1191 CAGAGAACCT ACGAAAAACC GCCCTGCAAG GCGGTTTTTT CGTTTTCAGA GCAAGAGATT ACGCGCAGAC
     GTCTCTTGGA TGCTTTTTGG CGGGACGTTC CGCCAAAAAA GCAAAAGTCT CGTTCTCTAA TGCGCGTCTG
-----------------------------------------------------------------------------
```

```
                                        BglII
                                        -------
1261  CAAAACGATC TCAAGAAGAT CATCTTATTA GATCTAGCAC CAGGCGTTTA AGGGCACCAA TAACTGCCTT
      GTTTTGCTAG AGTTCTTCTA GTAGAATAAT CTAGATCGTG GTCCGCAAAT TCCCGTGGTT ATTGACGGAA
-------------------------------------------------------------------------------

1331  AAAAAAATTA CGCCCCGCCC TGCCACTCAT CGCAGTACTG TTGTAATTCA TTAAGCATTC TGCCGACATG
      TTTTTTTAAT GCGGGGCGGG ACGGTGAGTA GCGTCATGAC AACATTAAGT AATTCGTAAG ACGGCTGTAC
-------------------------------------------------------------------------------

1401  GAAGCCATCA CAAACGGCAT GATGAACCTG AATCGCCAGC GGCATCAGCA CCTTGTCGCC TTGCGTATAA
      CTTCGGTAGT GTTTGCCGTA CTACTTGGAC TTAGCGGTCG CCGTAGTCGT GGAACAGCGG AACGCATATT
-------------------------------------------------------------------------------

1471  TATTTGCCCA TAGTGAAAAC GGGGGCGAAG AAGTTGTCCA TATTGGCTAC GTTTAAATCA AAACTGGTGA
      ATAAACGGGT ATCACTTTTG CCCCCGCTTC TTCAACAGGT ATAACCGATG CAAATTTAGT TTTGACCACT
-------------------------------------------------------------------------------

1541  AACTCACCCA GGGATTGGCT GAGACGAAAA ACATATTCTC AATAAACCCT TTAGGGAAAT AGGCCAGGTT
      TTGAGTGGGT CCCTAACCGA CTCTGCTTTT TGTATAAGAG TTATTTGGGA AATCCCTTTA TCCGGTCCAA
-------------------------------------------------------------------------------

1611  TTCACCGTAA CACGCCACAT CTTGCGAATA TATGTGTAGA AACTGCCGGA AATCGTCGTG GTATTCACTC
      AAGTGGCATT GTGCGGTGTA GAACGCTTAT ATACACATCT TTGACGGCCT TTAGCAGCAC CATAAGTGAG
-------------------------------------------------------------------------------

1681  CAGAGCGATG AAAACGTTTC AGTTTGCTCA TGGAAAACGG TGTAACAAGG GTGAACACTA TCCCATATCA
      GTCTCGCTAC TTTTGCAAAG TCAAACGAGT ACCTTTTGCC ACATTGTTCC CACTTGTGAT AGGGTATAGT
-------------------------------------------------------------------------------

1751  CCAGCTCACC GTCTTTCATT GCCATACGGA ACTCCGGGTG AGCATTCATC AGGCGGGCAA GAATGTGAAT
      GGTCGAGTGG CAGAAAGTAA CGGTATGCCT TGAGGCCCAC TCGTAAGTAG TCCGCCCGTT CTTACACTTA
-------------------------------------------------------------------------------

1821  AAAGGCCGGA TAAAACTTGT GCTTATTTTT CTTTACGGTC TTTAAAAAGG CCGTAATATC CAGCTGAACG
      TTTCCGGCCT ATTTTGAACA CGAATAAAAA GAAATGCCAG AAATTTTTCC GGCATTATAG GTCGACTTGC
-------------------------------------------------------------------------------

1891  GTCTGGTTAT AGGTACATTG AGCAACTGAC TGAAATGCCT CAAAATGTTC TTTACGATGC CATTGGGATA
      CAGACCAATA TCCATGTAAC TCGTTGACTG ACTTTACGGA GTTTTACAAG AAATGCTACG GTAACCCTAT
-------------------------------------------------------------------------------

1961  TATCAACGGT GGTATATCCA GTGATTTTTT TCTCCATTTT AGCTTCCTTA GCTCCTGAAA ATCTCGATAA
      ATAGTTGCCA CCATATAGGT CACTAAAAAA AGAGGTAAAA TCGAAGGAAT CGAGGACTTT TAGAGCTATT
-------------------------------------------------------------------------------

                                                                    AatII
                                                                    ------
2031  CTCAAAAAAT ACGCCCGGTA GTGATCTTAT TTCATTATGG TGAAAGTTGG AACCTCACCC GACGTCTAAT
      GAGTTTTTTA TGCGGGCCAT CACTAGAATA AAGTAATACC ACTTTCAACC TTGGAGTGGG CTGCAGATTA
-------------------------------------------------------------------------------

2101  GTGAGTTAGC TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA
      CACTCAATCG AGTGAGTAAT CCGTGGGGTC CGAAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT
-------------------------------------------------------------------------------

                                                          XbaI
                                                          -------
2171  TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA CCATGATTAC GAATTTCTAG ACCCCCCCCC
      AACACTCGCC TATTGTTAAA GTGTGTCCTT TGTCGATACT GGTACTAATG CTTAAAGATC TGGGGGGGGG
-------------------------------------------------------------------------------

       SphI                                       HindIII
       ------                                     ------
2241  CGCATGCCAT AACTTCGTAT AATGTACGCT ATACGAAGTT ATAAGCTTGA CCTGTGAAGT GAAAAATGGC
      GCGTACGGTA TTGAAGCATA TTACATGCGA TATGCTTCAA TATTCGAACT GGACACTTCA CTTTTTACCG
-------------------------------------------------------------------------------

                                        PacI          PseI
                                        --------      ---------
2311  GCAGATTGTG CGACATTTTT TTTGTCTGCC GTTTAATTAA GGGGGGGGGC CGGCCATTAT CAAAAAGGAT
      CGTCTAACAC GCTGTAAAAA AAACAGACGG CAAATTAATT CCCCCCCCCG GCCGGTAATA GTTTTTCCTA
-------------------------------------------------------------------------------

2381  CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT
      GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG CAATTCCCTA
-------------------------------------------------------------------------------

2451  TTTGGTCATG AGATTATCAA AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA
      AAACCAGTAC TCTAATAGTT TTTCCTAGAA GTGGATCTAG GAAAATTTAA TTTTTACTTC AAAATTTAGT
-------------------------------------------------------------------------------
```

```
2521  ATCTAAAGTA TATATGAGTA AACTTGGTCT GACAGTTACC CAATGCTTAA TCAGTGAGGC ACCTATCTCA
      TAGATTTCAT ATATACTCAT TTGAACCAGA CTGTCAATGG GTTACGAATT AGTCACTCCG TGGATAGAGT
-------------------------------------------------------------------------------

2591  GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG
      CGCTAGACAG ATAAAGCAAG TAGGTATCAA CGGACTGAGG GGCAGCACAT CTATTGATGC TATGCCCTCC
-------------------------------------------------------------------------------

2661  GCTTACCATC TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
      CGAATGGTAG ACCGGGGTCA CGACGTTACT ATGGCGCTCT GGGTGCGAGT GGCCGAGGTC TAAATAGTCG
-------------------------------------------------------------------------------

2731  AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT
      TTATTTGGTC GGTCGGCCTT CCCGGCTCGC GTCTTCACCA GGACGTTGAA ATAGGCGGAG GTAGGTCAGA
-------------------------------------------------------------------------------

2801  ATTAACTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG
      TAATTGACAA CGGCCCTTCG ATCTCATTCA TCAAGCGGTC AATTATCAAA CGCGTTGCAA CAACGGTAAC
-------------------------------------------------------------------------------

2871  CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG
      GATGTCCGTA GCACCACAGT GCGAGCAGCA AACCATACCG AAGTAAGTCG AGGCCAAGGG TTGCTAGTTC
-------------------------------------------------------------------------------

2941  GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA
      CGCTCAATGT ACTAGGGGGT ACAACACGTT TTTTCGCCAA TCGAGGAAGC CAGGAGGCTA GCAACAGTCT
-------------------------------------------------------------------------------

3011  AGTAAGTTGG CCGCAGTGTT ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT
      TCATTCAACC GGCGTCACAA TAGTGAGTAC CAATACCGTC GTGACGTATT AAGAGAATGA CAGTACGGTA
-------------------------------------------------------------------------------

3081  CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC
      GGCATTCTAC GAAAAGACAC TGACCACTCA TGAGTTGGTT CAGTAAGACT CTTATCACAT ACGCCGCTGG
-------------------------------------------------------------------------------

3151  GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC
      CTCAACGAGA ACGGGCCGCA GTTATGCCCT ATTATGGCGC GGTGTATCGT CTTGAAATTT TCACGAGTAG
-------------------------------------------------------------------------------

            XmnI
            ----------
3221  ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC
      TAACCTTTTG CAAGAAGCCC CGCTTTTGAG AGTTCCTAGA ATGGCGACAA CTCTAGGTCA AGCTACATTG
-------------------------------------------------------------------------------

3291  CCACTCGCGC ACCCAACTGA TCCTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG
      GGTGAGCGCG TGGGTTGACT AGGAGTCGTA GAAAATGAAA GTGGTCGCAA AGACCCACTC GTTTTTGTCC
-------------------------------------------------------------------------------

3361  AAGGCAAAAT GCCGCAAAAA AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT
      TTCCGTTTTA CGGCGTTTTT TCCCTTATTC CCGCTGTGCC TTTACAACTT ATGAGTATGA GAAGGAAAAA
-------------------------------------------------------------------------------

                                                              BsrGI
                                                              ~
3431  CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAAT
      GTTATAATAA CTTCGTAAAT AGTCCCAATA ACAGAGTACT CGCCTATGTA TAAACTTA
-------------------------------------------------------------------------------
```


BglII (1)
AatII (46)
lac p/o
XbaI (177)
SphI (197)
HindIII (234)
lpp-Terminator
PacI (298)
FseI (314)
bla-term
bla
XmnI (1180)
GenII-Nick
functional ssori
BsrGI (1439)
BanII (1746)
NheI (1903)
XmnI (1959)
AgeI (1989)
P15A origin
AgeI (2313)
ds-ori
RNAI TT
pCALO-3
2728 bp

```
     BglII                                                AatII
     -----                                                ------
   1 GATCTCATAA CTTCGTATAA TGTATGCTAT ACGAAGTTAT GACGTCTAAT GTGAGTTAGC TCACTCATTA
     CTAGAGTATT GAAGCATATT ACATACGATA TGCTTCAATA CTGCAGATTA CACTCAATCG AGTGAGTAAT
-----------------------------------------------------------------------------
  71 GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT
     CCGTGGGGTC CGAAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT AACACTCGCC TATTGTTAAA
-----------------------------------------------------------------------------
                                          XbaI             SphI
                                          -------          ------
 141 CACACAGGAA ACAGCTATGA CCATGATTAC GAATTTCTAG ACCCCCCCCC CGCATGCCAT AACTTCGTAT
     GTGTGTCCTT TGTCGATACT GGTACTAATG CTTAAAGATC TGGGGGGGGG GCGTACGGTA TTGAAGCATA
-----------------------------------------------------------------------------
                           HindIII
                           ------
 211 AATGTACGCT ATACGAAGTT ATAAGCTTGA CCTGTGAAGT GAAAAATGGC GCAGATTGTG CGACATTTTT
     TTACATGCGA TATGCTTCAA TATTCGAACT GGACACTTCA CTTTTTACCG CGTCTAACAC GCTGTAAAAA
-----------------------------------------------------------------------------
                   PacI            FseI
                   --------        ----------
 281 TTTGTCTGCC GTTTAATTAA GGGGGGGGGC CGGCCATTAT CAAAAAGGAT CTCAAGAAGA TCCTTTGATC
     AAACAGACGG CAAATTAATT CCCCCCCCCG GCCGGTAATA GTTTTTCCTA GAGTTCTTCT AGGAAACTAG
-----------------------------------------------------------------------------
 351 TTTTCTACGG GGTCTGACGC TCAGTGGAAC GAAAACTCAC GTTAAGGGAT TTTGGTCATG AGATTATCAA
     AAAAGATGCC CCAGACTGCG AGTCACCTTG CTTTTGAGTG CAATTCCCTA AAACCAGTAC TCTAATAGTT
-----------------------------------------------------------------------------
 421 AAAGGATCTT CACCTAGATC CTTTTAAATT AAAAATGAAG TTTTAAATCA ATCTAAAGTA TATATGAGTA
     TTTCCTAGAA GTGGATCTAG GAAAATTTAA TTTTTACTTC AAAATTTAGT TAGATTTCAT ATATACTCAT
-----------------------------------------------------------------------------
 491 AACTTGGTCT GACAGTTACC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC
     TTGAACCAGA CTGTCAATGG GTTACGAATT AGTCACTCCG TGGATAGAGT CGCTAGACAG ATAAAGCAAG
-----------------------------------------------------------------------------
 561 ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
     TAGGTATCAA CGGACTGAGG GGCAGCACAT CTATTGATGC TATGCCCTCC CGAATGGTAG ACCGGGGTCA
-----------------------------------------------------------------------------
 631 GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC AATAAACCAG CCAGCCGGAA
     CGACGTTACT ATGGCGCTCT GGGTGCGAGT GGCCGAGGTC TAAATAGTCG TTATTTGGTC GGTCGGCCTT
-----------------------------------------------------------------------------
 701 GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC CATCCAGTCT ATTAACTGTT GCCGGGAAGC
     CCCGGCTCGC GTCTTCACCA GGACGTTGAA ATAGGCGGAG GTAGGTCAGA TAATTGACAA CGGCCCTTCG
-----------------------------------------------------------------------------
 771 TAGAGTAAGT AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA
     ATCTCATTCA TCAAGCGGTC AATTATCAAA CGCGTTGCAA CAACGGTAAC GATGTCCGTA GCACCACAGT
-----------------------------------------------------------------------------
 841 CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA
     GCGAGCAGCA AACCATACCG AAGTAAGTCG AGGCCAAGGG TTGCTAGTTC CGCTCAATGT ACTAGGGGGT
-----------------------------------------------------------------------------
 911 TGTTGTGCAA AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT
     ACAACACGTT TTTTCGCCAA TCGAGGAAGC CAGGAGGCTA GCAACAGTCT TCATTCAACC GGCGTCACAA
-----------------------------------------------------------------------------
 981 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG CTTTTCTGTG
     TAGTGAGTAC CAATACCGTC GTGACGTATT AAGAGAATGA CAGTACGGTA GGCATTCTAC GAAAAGACAC
-----------------------------------------------------------------------------
1051 ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT
     TGACCACTCA TGAGTTGGTT CAGTAAGACT CTTATCACAT ACGCCGCTGG CTCAACGAGA ACGGGCCGCA
-----------------------------------------------------------------------------
                                                                  XmnI
                                                                  ----------
1121 CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC GTTCTTCGGG
     GTTATGCCCT ATTATGGCGC GGTGTATCGT CTTGAAATTT TCACGAGTAG TAACCTTTTG CAAGAAGCCC
-----------------------------------------------------------------------------
1191 GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGCGC ACCCAACTGA
     CGCTTTTGAG AGTTCCTAGA ATGGCGACAA CTCTAGGTCA AGCTACATTG GGTGAGCGCG TGGGTTGACT
-----------------------------------------------------------------------------
```

```
1261  TCCTCAGCAT CTTTTACTTT CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA
      AGGAGTCGTA GAAAATGAAA GTGGTCGCAA AGACCCACTC GTTTTTGTCC TTCCGTTTTA CGGCGTTTTT

1331  AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA
      TCCCTTATTC CCGCTGTGCC TTTACAACTT ATGAGTATGA GAAGGAAAAA GTTATAATAA CTTCGTAAAT

                                          BsrGI
                                          -------
1401  TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ACATGAAATT GTAAACGTTA ATATTTTGTT
      AGTCCCAATA ACAGAGTACT CGCCTATGTA TAAACTTACA TGTACTTTAA CATTTGCAAT TATAAAACAA

1471  AAAATTCGCG TTAAATTTTT GTTAAATCAG CTCATTTTTT AACCAATAGG CCGAAATCGG CAAAATCCCT
      TTTTAAGCGC AATTTAAAAA CAATTTAGTC GAGTAAAAAA TTGGTTATCC GGCTTTAGCC GTTTTAGGGA

1541  TATAAATCAA AAGAATAGAC CGAGATAGGG TTGAGTGTTG TTCCAGTTTG GAACAAGAGT CCACTATTAA
      ATATTTAGTT TTCTTATCTG GCTCTATCCC AACTCACAAC AAGGTCAAAC CTTGTTCTCA GGTGATAATT

1611  AGAACGTGGA CTCCAACGTC AAAGGGCGAA AAACCGTCTA TCAGGGCGAT GGCCCACTAC GAGAACCATC
      TCTTGCACCT GAGGTTGCAG TTTCCCGCTT TTTGGCAGAT AGTCCCGCTA CCGGGTGATG CTCTTGGTAG

                                                                        BanII
                                                                        ------
1681  ACCCTAATCA AGTTTTTTGG GGTCGAGGTG CCGTAAAGCA CTAAATCGGA ACCCTAAAGG GAGCCCCCGA
      TGGGATTAGT TCAAAAAACC CCAGCTCCAC GGCATTTCGT GATTTAGCCT TGGGATTTCC CTCGGGGGCT

1751  TTTAGAGCTT GACGGGGAAA GCCGGCGAAC GTGGCGAGAA AGGAAGGGAA GAAAGCGAAA GGAGCGGGCG
      AAATCTCGAA CTGCCCCTTT CGGCCGCTTG CACCGCTCTT TCCTTCCCTT CTTTCGCTTT CCTCGCCCGC

1821  CTAGGGCGCT GGCAAGTGTA GCGGTCACGC TGCGCGTAAC CACCACACCC GCCGCGCTTA ATGCGCCGCT
      GATCCCGCGA CCGTTCACAT CGCCAGTGCG ACGCGCATTG GTGGTGTGGG CGGCGCGAAT TACGCGGCGA

                  NheI                                                  XmnI
                  ------                                                -------
1891  ACAGGGCGCG TGCTAGCGGA GTGTATACTG GCTTACTATG TTGGCACTGA TGAGGGTGTC AGTGAAGTGC
      TGTCCCGCGC ACGATCGCCT CACATATGAC CGAATGATAC AACCGTGACT ACTCCCACAG TCACTTCACG

      XmnI                        AgeI
      ---                         -------
1961  TTCATGTGGC AGGAGAAAAA AGGCTGCACC GGTGCGTCAG CAGAATATGT GATACAGGAT ATATTCCGCT
      AAGTACACCG TCCTCTTTTT TCCGACGTGG CCACGCAGTC GTCTTATACA CTATGTCCTA TATAAGGCGA

2031  TCCTCGCTCA CTGACTCGCT ACGCTCGGTC GTTCGACTGC GGCGAGCGGA AATGGCTTAC GAACGGGGCG
      AGGAGCGAGT GACTGAGCGA TGCGAGCCAG CAAGCTGACG CCGCTCGCCT TTACCGAATG CTTGCCCCGC

2101  GAGATTTCCT GGAAGATGCC AGGAAGATAC TTAACAGGGA AGTGAGAGGG CCGCGGCAAA GCCGTTTTTC
      CTCTAAAGGA CCTTCTACGG TCCTTCTATG AATTGTCCCT TCACTCTCCC GGCGCCGTTT CGGCAAAAAG

2171  CATAGGCTCC GCCCCCCTGA CAAGCATCAC GAAATCTGAC GCTCAAATCA GTGGTGGCGA AACCCGACAG
      GTATCCGAGG CGGGGGGACT GTTCGTAGTG CTTTAGACTG CGAGTTTAGT CACCACCGCT TTGGGCTGTC

2241  GACTATAAAG ATACCAGGCG TTTCCCCCTG GCGGCTCCCT CCTGCGCTCT CCTGTTCCTG CCTTTCGGTT
      CTGATATTTC TATGGTCCGC AAAGGGGGAC CGCCGAGGGA GGACGCGAGA GGACAAGGAC GGAAAGCCAA

        AgeI
        ------
2311  TACCGGTGTC ATTCCGCTGT TATGGCCGCG TTTGTCTCAT TCCACGCCTG ACACTCAGTT CCGGGTAGGC
      ATGGCCACAG TAAGGCGACA ATACCGGCGC AAACAGAGTA AGGTGCGGAC TGTGAGTCAA GGCCCATCCG

2381  AGTTCGCTCC AAGCTGGACT GTATGCACGA ACCCCCCGTT CAGTCCGACC GCTGCGCCTT ATCCGGTAAC
      TCAAGCGAGG TTCGACCTGA CATACGTGCT TGGGGGGCAA GTCAGGCTGG CGACGCGGAA TAGGCCATTG

2451  TATCGTCTTG AGTCCAACCC GGAAAGACAT GCAAAAGCAC CACTGGCAGC AGCCACTGGT AATTGATTTA
      ATAGCAGAAC TCAGGTTGGG CCTTTCTGTA CGTTTTCGTG GTGACCGTCG TCGGTGACCA TTAACTAAAT
```

```
2521  GAGGAGTTAG TCTTGAAGTC ATGCGCCGGT TAAGGCTAAA CTGAAAGGAC AAGTTTTAGT GACTGCGCTC
      CTCCTCAATC AGAACTTCAG TACGCGGCCA ATTCCGATTT GACTTTCCTG TTCAAAATCA CTGACGCGAG
-----------------------------------------------------------------------------

2591  CTCCAAGCCA GTTACCTCGG TTCAAAGAGT TGGTAGCTCA GAGAACCTAC GAAAAACCGC CCTGCAAGGC
      GAGGTTCGGT CAATGGAGCC AAGTTTCTCA ACCATCGAGT CTCTTGGATG CTTTTTGGCG GGACGTTCCG
-----------------------------------------------------------------------------

                                                                      BalII
                                                                        -
2661  GGTTTTTTCG TTTTCAGAGC AAGAGATTAC GCGCAGACCA AAACGATCTC AAGAAGATCA TCTTATTA
      CCAAAAAAGC AAAAGTCTCG TTCTCTAATG CGCGTCTGGT TTTGCTAGAG TTCTTCTAGT AGAATAAT
-----------------------------------------------------------------------------
```

Figure 35b: List of oligonucleotides used for synthesis of modules

M1: PCR using template

NoVspAatII
TAGACGTC

M2: synthesis

BloxA-A
TATGAGATCTCATAACTTCGTATAATGTACGCTATACGAAGTTAT

BloxA-B
TAATAACTTCGTATAGCATACATTATACGAAGTTATGAGATCTCA

M3: PCR, NoVspAatII as second oligo

XloxS-muta
CATTTTTTGCCCTCGTTATCTACGCATGCGATAACTTCGTATAGCGTACATTATACGA
AGTTATTCTAGACATGGTCATAGCTGTTTCCTG

M7-I: PCR

gIIINEW-fow
GGGGGGAATTCGGTGGTGGTGGATCTGCGTGCGCTGAAACGGTTGAAAGTTG

gIIINEW-rev
CCCCCCCAAGCTTATCAAGACTCCTTATTACG

M7-II: PCR

gIIIss-fow
GGGGGGGGAATTCGGAGGCGGTTCCGGTGGTGGC

Figure 35b: List of oligonucleotides used for synthesis of modules

M7-III: PCR

gIIIsupernew-fow
GGGGGGGGAATTCGAGCAGAAGCTGATCTCTGAGGAGGATCTGTAGGGTGGTGG
CTCTGGTTCCGGTGATTTTG

M8: synthesis

lox514-A
CCATAACTTCGTATAATGTACGCTATACGAAGTTATA

lox514-B
AGCTTATAACTTCGTATAGCGTACATTATACGAAGTTATGGCATG

M9II: synthesis

M9II-fow
AGCTTGACCTGTGAAGTGAAAAATGGCGCAGATTGTGCGACATTTTTTTTGTCTGCC
GTTTAATTAAAGGGGGGGT

M9II-rev
GTACACCCCCCCCCAGGCCGGCCCCCCCCCCCCTTTAATTAAACGGCAGACAAAAA
AAATGTCGCACAATCTGCG

M10II: assembly PCR with template

bla-fow:
GGGGGGGTGTACATTCAAATATGTATCCGCTCATG

bla-seq4
GGGTTACATCGAACTGGATCTC

bla1-muta
CCAGTTCGATGTAACCCACTCGCGCACCCAACTGATCCTCAGCATCTTTTACTTTCA
CC

Figure 35b: List of oligonucleotides used for synthesis of modules

blaII-muta
ACTCTAGCTTCCCGGCAACAGTTAATAGACTGGATGGAGGCGG

bla-NEW
CTGTTGCCGGGAAGCTAGAGTAAG

bla-rev
CCCCCCCTTAATTAAGGGGGGGGGCCGGCCATTATCAAAAAGGATCTCAAGAAGAT
CC

M11II/III: PCR, site-directed mutagenesis

f1-fow
GGGGGGGGCTAGCACGCGCCCTGTAGCGGCGCATTAA

f1-rev
CCCCCCCTGTACATGAAATTGTAAACGTTAATATTTTG

f1-t133.muta
GGGCGATGGCCCACTACGAGAACCATCACCCTAATC

M12: assembly PCR using template

p15-fow
GGGGGGAGATCTAATAAGATGATCTTCTTGAG

p15-NEWI
GAGTTGGTAGCTCAGAGAACCTACGAAAAACCGCCCTGCAAGGCG

p15-NEWII
GTAGGTTCTCTGAGCTACCAACTC

p15-NEWIII
GTTTCCCCCTGGCGGCTCCCTCCTGCGCTCTCCTGTTCCTGCC

p15-NEWIV
AGGAGGGAGCCGCCAGGGGGAAAC

Figure 35b: List of oligonucleotides used for synthesis of modules

p15-rev
GACATCAGCGCTAGCGGAGTGTATAC

M13: synthesis

BloxXB-A
GATCTCATAACTTCGTATAATGTATGCTATACGAAGTTATTCA

BloxXB-B
GATCTGAATAACTTCGTATAGCATACATTATACGAAGTTATGAGA

M14-Ext2: PCR, site-directed mutagenesis

ColEXT2-fow
GGGGGGGAGATCTGACCAAAATCCCTTAACGTGAG

Col-mutaI
GGTATCTGCGCTCTGCTGTAGCCAGTTACCTTCGG

Col-rev
CCCCCCCGCTAGCCATGTGAGCAAAAGGCCAGCAA

M17: assembly PCR using template

CAT-1
GGGACGTCGGGTGAGGTTCCAAC

CAT-2
CCATACGGAACTCCGGGTGAGCATTCATC

CAT-3
CCGGAGTTCCGTATGG

CAT-4
ACGTTTAAATCAAAACTGG
CAT-5

Figure 35b: List of oligonucleotides used for synthesis of modules

CCAGTTTTGATTTAAACGTAGCCAATATGGACAACTTCTTCGCCCCCGTTTTCACTAT
GGGCAAATATT

CAT-6
GGAAGATCTAGCACCAGGCGTTTAAG

M41: assembly PCR using template

LAC1
GAGGCCGGCCATCGAATGGCGCAAAAC

LAC2
CGCGTACCGTCCTCATGGGAGAAAATAATAC

LAC3
CCATGAGGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCC
GCTGTTAGCTGGCCCATTAAG

LAC4
GTCAGCGGCGGGATATAACATGAGCTGTCCTCGGTATCGTCG

LAC5
GTTATATCCCGCCGCTGACCACCATCAAAC

LAC6
CATCAGTGAATCGGCCAACGCGCGGGGAGAGGCGGTTTGCGT4TTGGGAGCCAGGGTGGTTT
TTC

LAC7
GGTTAATTAACCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATCAGTGA
ATCGGCCAAC

M41-MCS-fow
CTAGACTAGTGTTTAAACCGGACCGGGGGGGGGCTTAAGGGGGGGGGGGG

M41-MCS-rev
CTAGCCCCCCCCCCCCTTAAGCCCCCCCCCGGTCCGGTTTAAACACTAGT

Figure 35b: List of oligonucleotides used for synthesis of modules

M41-fow
CTAGACTAGTGTTTAAACCGGACCGGGGGGGGGCTTAAGGGGGGGGGGGG

M41-rev
CCCCCCCTTAAGTGGGCTGCAAAACAAAACGGCCTCCTGTCAGGAAGCCGCTTTTA
TCGGGTAGCCTCACTGCCCGCTTTCC

M41-A2
GTTGTTGTGCCACGCGGTTAGGAATGTAATTCAGCTCCGC

M41-B1
AACCGCGTGGCACAACAAC

M41-B2
CTTCGTTCTACCATCGACACGACCACGCTGGCACCCAGTTG

M41-C1
GTGTCGATGGTAGAACGAAG

M41-CII
CCACAGCAATAGCATCCTGGTCATCCAGCGGATAGTTAATAATCAGCCCACTGACA
CGTTGCGCGAG

M41-DI
GACCAGGATGCTATTGCTGTGG

M41-DII
CAGCGCGATTTGCTGGTGGCCCAATGCGACCAGATGC

M41-EI
CACCAGCAAATCGCGCTG

M41-EII
CCCGGACTCGGTAATGGCACGCATTGCGCCCAGCGCC

M41-FI
GCCATTACCGAGTCCGGG

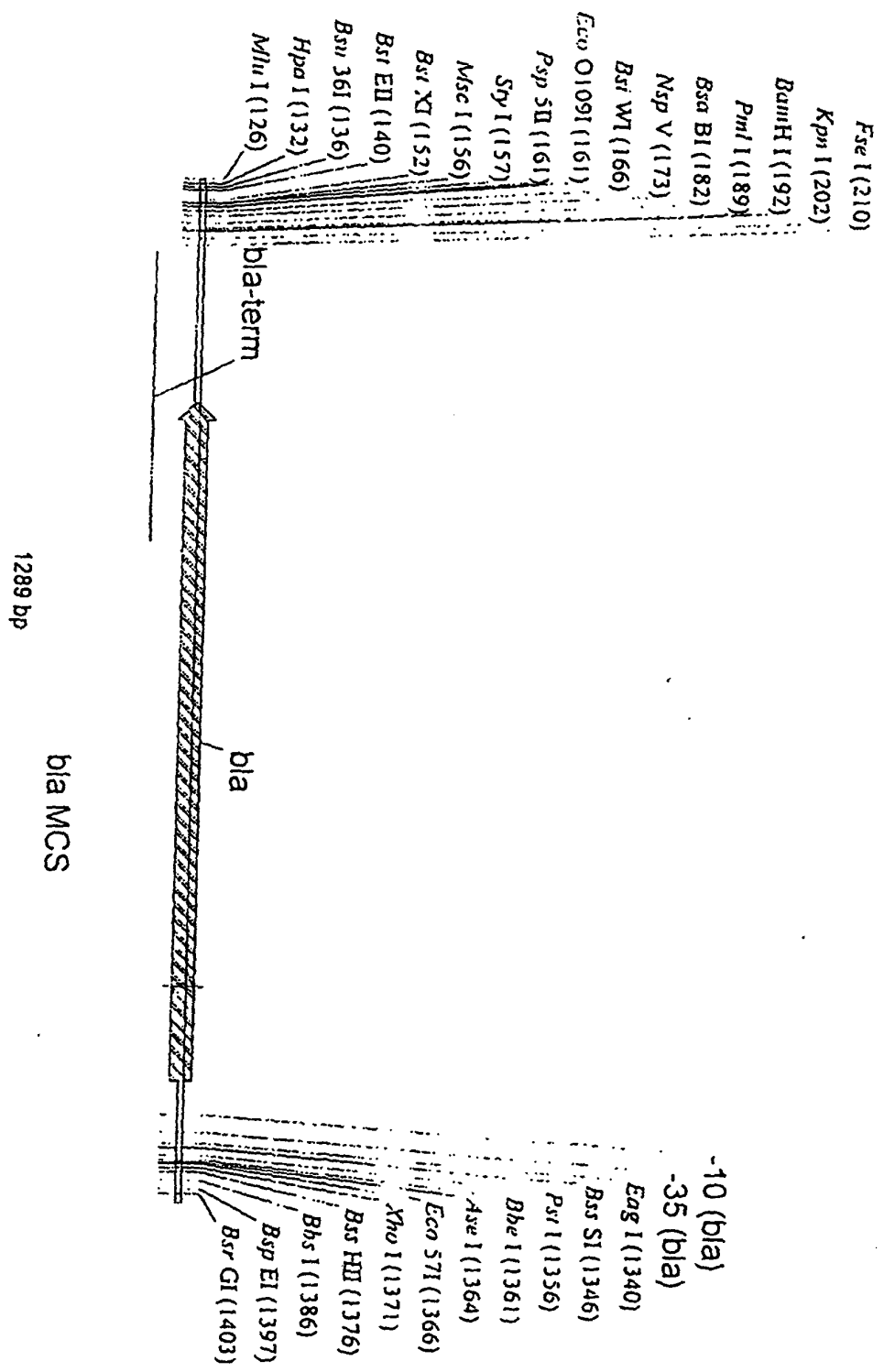
Figure 35b: List of oligonucleotides used for synthesis of modules

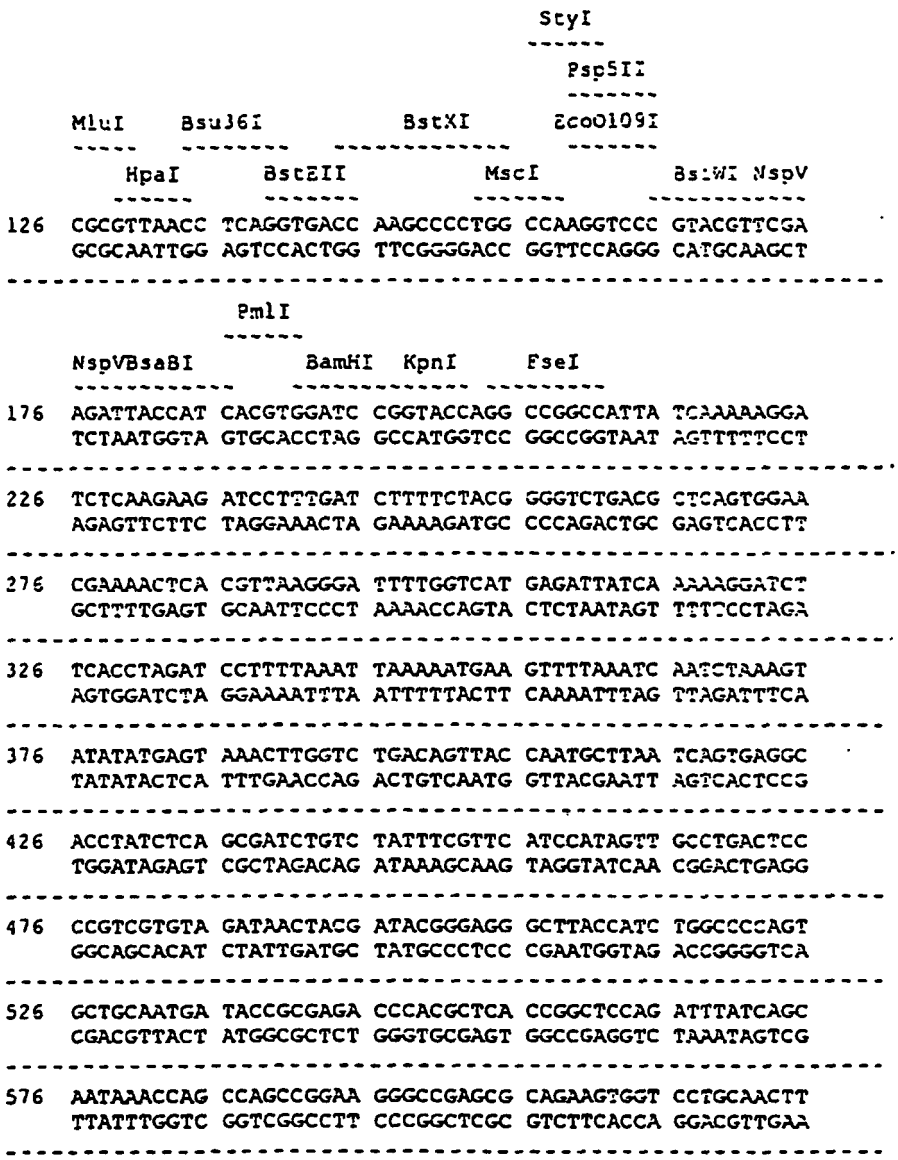
M42: synthesis

Eco-H5-Hind-fow
AATTCCACCATCATCACCATTGACGTCTA

Eco-H5-Hind-rev
AGCTTAGACGTCAATGGTGATGATGGTGG

Figure 36: functional map and sequence of ß-lactamase-MCS module




StyI
PspSII
MluI
Bsu36I
BstXI
Eco0109I
HpaI
BstEII
MscI
BsiWI NspV
PmlI
NspVBsaBI
BamHI
KpnI
FseI
126 CGCGTTAACC TCAGGTGACC AAGCCCCTGG CCAAGGTCCC GTACGTTCGA
GCGCAATTGG AGTCCACTGG TTCGGGGACC GGTTCCAGGG CATGCAAGCT
176 AGATTACCAT CACGTGGATC CGGTACCAGG CCGGCCATTA TCAAAAAGGA
TCTAATGGTA GTGCACCTAG GCCATGGTCC GGCCGGTAAT AGTTTTTCCT
226 TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA
AGAGTTCTTC TAGGAAACTA GAAAAGATGC CCCAGACTGC GAGTCACCTT
276 CGAAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT
GCTTTTGAGT GCAATTCCCT AAAACCAGTA CTCTAATAGT TTTTCCTAGA
326 TCACCTAGAT CCTTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT
AGTGGATCTA GGAAAATTTA ATTTTTACTT CAAAATTTAG TTAGATTTCA
376 ATATATGAGT AAACTTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC
TATATACTCA TTTGAACCAG ACTGTCAATG GTTACGAATT AGTCACTCCG
426 ACCTATCTCA GCGATCTGTC TATTTCGTTC ATCCATAGTT GCCTGACTCC
TGGATAGAGT CGCTAGACAG ATAAAGCAAG TAGGTATCAA CGGACTGAGG
476 CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC TGGCCCCAGT
GGCAGCACAT CTATTGATGC TATGCCCTCC CGAATGGTAG ACCGGGGTCA
526 GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
CGACGTTACT ATGGCGCTCT GGGTGCGAGT GGCCGAGGTC TAAATAGTCG
576 AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT
TTATTTGGTC GGTCGGCCTT CCCGGCTCGC GTCTTCACCA GGACGTTGAA

```
 626  TATCCGCCTC CATCCAGTCT ATTAACTGTT GCCGGGAAGC TAGAGTAAGT
      ATAGGCGGAG GTAGGTCAGA TAATTGACAA CGGCCCTTCG ATCTCATTCA
 ------------------------------------------------------------

 676  AGTTCGCCAG TTAATAGTTT GCGCAACGTT GTTGCCATTG CTACAGGCAT
      TCAAGCGGTC AATTATCAAA CGCGTTGCAA CAACGGTAAC GATGTCCGTA
 ------------------------------------------------------------

 726  CGTGGTGTCA CGCTCGTCGT TTGGTATGGC TTCATTCAGC TCCGGTTCCC
      GCACCACAGT GCGAGCAGCA AACCATACCG AAGTAAGTCG AGGCCAAGGG
 ------------------------------------------------------------

 776  AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA AAAAGCGGTT
      TTGCTAGTTC CGCTCAATGT ACTAGGGGGT ACAACACGTT TTTTCGCCAA
 ------------------------------------------------------------

 826  AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT
      TCGAGGAAGC CAGGAGGCTA GCAACAGTCT TCATTCAACC GGCGTCACAA
 ------------------------------------------------------------

 876  ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT
      TAGTGAGTAC CAATACCGTC GTGACGTATT AAGAGAATGA CAGTACGGTA
 ------------------------------------------------------------

 926  CCGTAAGATG CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA
      GGCATTCTAC GAAAAGACAC TGACCACTCA TGAGTTGGTT CAGTAAGACT
 ------------------------------------------------------------

 976  GAATAGTGTA TGCGGCGACC GAGTTGCTCT TGCCCGGCGT CAATACGGGA
      CTTATCACAT ACGCCGCTGG CTCAACGAGA ACGGGCCGCA GTTATGCCCT
 ------------------------------------------------------------

1026  TAATACCGCG CCACATAGCA GAACTTTAAA AGTGCTCATC ATTGGAAAAC
      ATTATGGCGC GGTGTATCGT CTTGAAATTT TCACGAGTAG TAACCTTTTG
 ------------------------------------------------------------

1076  GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT GAGATCCAGT
      CAAGAAGCCC CGCTTTTGAG AGTTCCTAGA ATGGCGACAA CTCTAGGTCA
 ------------------------------------------------------------

                   BssSI
                  ------
1126  TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT
      AGCTACATTG GGTGAGCACG TGGGTTGACT AGAAGTCGTA GAAAATGAAA
                                      Eco57I
                                      ------
 ------------------------------------------------------------

1176  CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA
      GTGGTCGCAA AGACCCACTC GTTTTTGTCC TTCCGTTTTA CGGCGTTTTT
 ------------------------------------------------------------
```

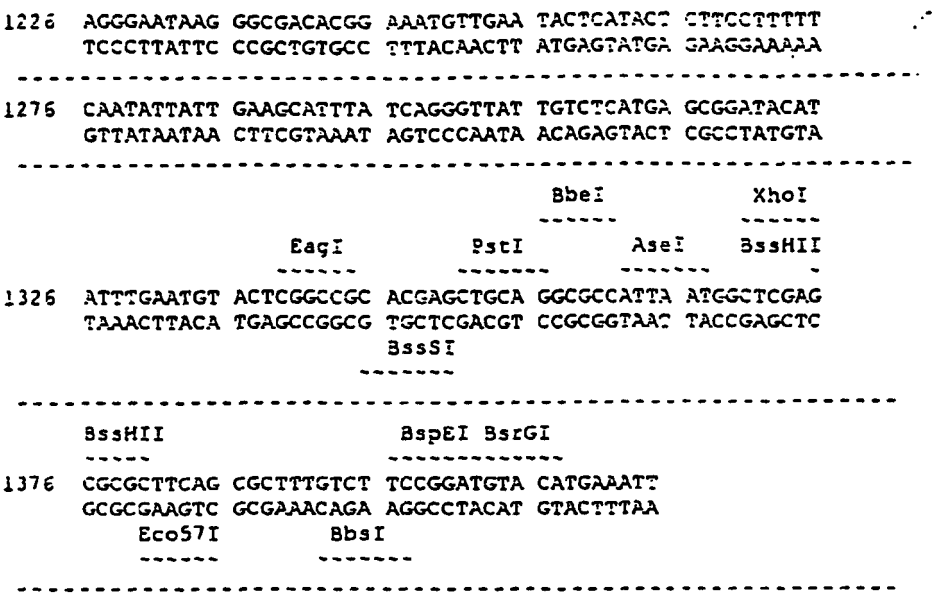

1226 AGGGAATAAG GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT
TCCCTTATTC CCGCTGTGCC TTTACAACTT ATGAGTATGA GAAGGAAAAA
1276 CAATATTATT GAAGCATTTA TCAGGGTTAT TGTCTCATGA GCGGATACAT
GTTATAATAA CTTCGTAAAT AGTCCCAATA ACAGAGTACT CGCCTATGTA
BbeI
XhoI
EagI
PstI
AseI
BssHII
1326 ATTTGAATGT ACTCGGCCGC ACGAGCTGCA GGCGCCATTA ATGGCTCGAG
TAAACTTACA TGAGCCGGCG TGCTCGACGT CCGCGGTAAT TACCGAGCTC
BssSI
BssHII
BspEI BsrGI
1376 CGCGCTTCAG CGCTTTGTCT TCCGGATGTA CATGAAATT
GCGCGAAGTC GCGAAACAGA AGGCCTACAT GTACTTTAA
Eco57I
BbsI

Figure 37: Oligo and primer design for Vκ CDR3 libraries

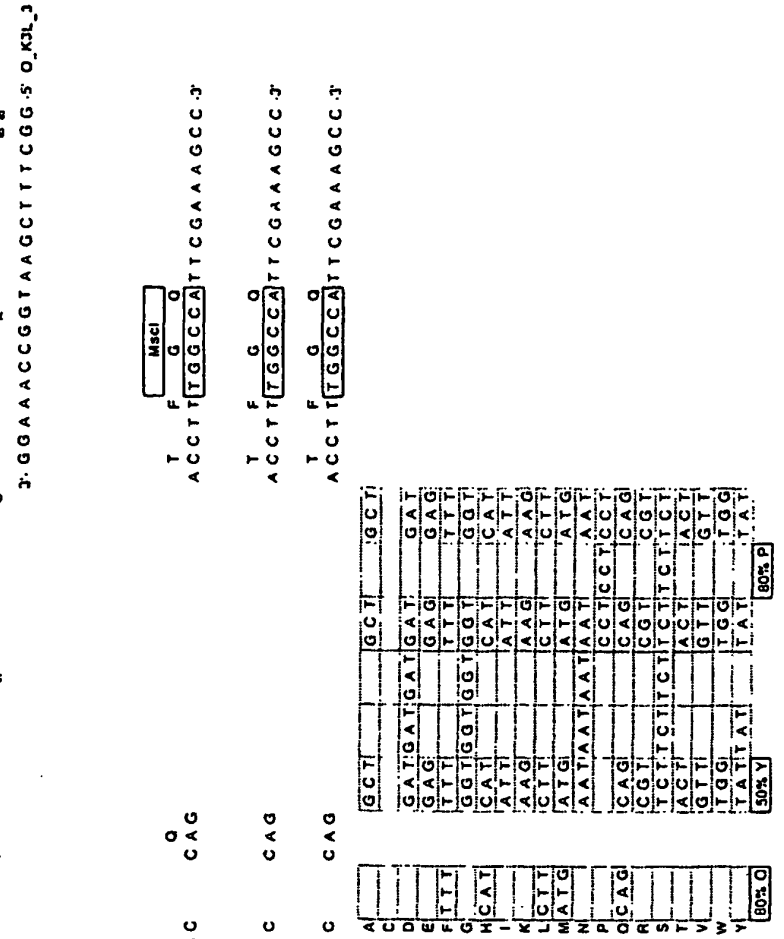


Figure 38: Oligo and primer design for Vλ CDR3 libraries

Figure 39: functional map of expression vector series pBS13

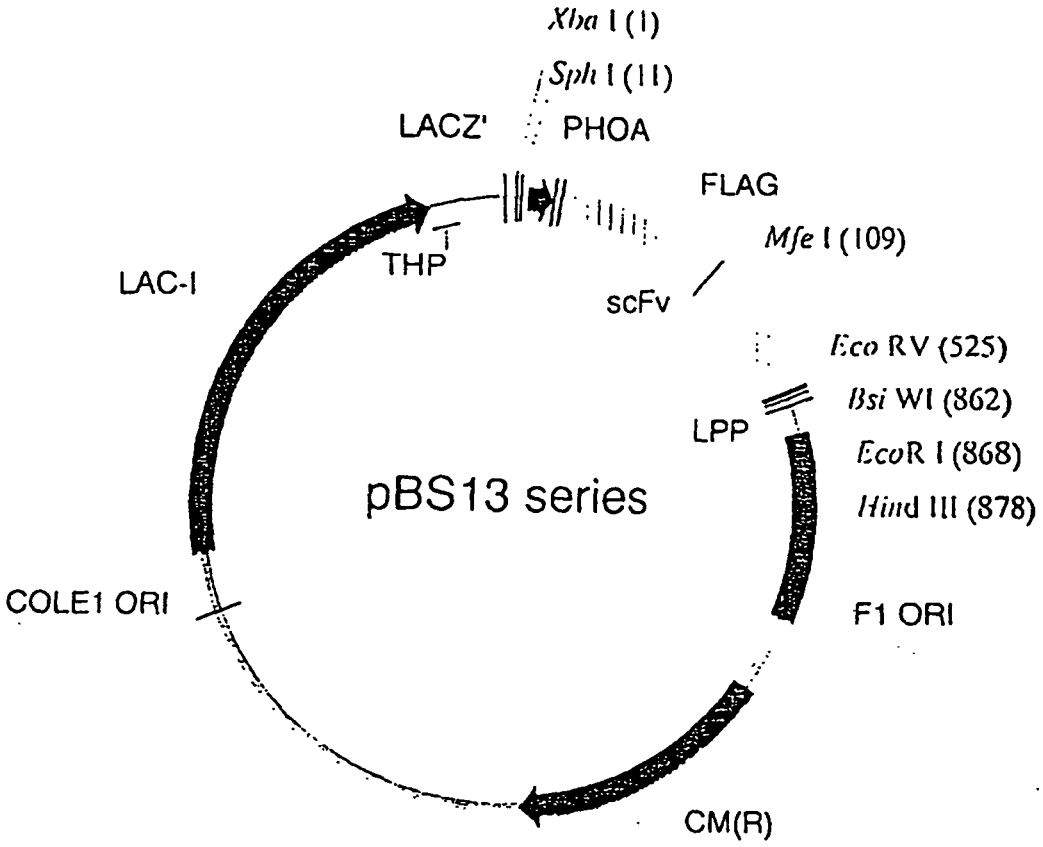


Figure 40: Expression data for HuCAL scFvs (pBS13, 30°C)

% soluble	κ1	κ2	κ3	κ4	λ1	λ2	λ3
H1A	61%	58%	52%	42%	90%	61%	60%
H1B	39%	48%	66%	48%	47%	39%	36%
H2	47%	57%	46%	49%	37%	36%	45%
H3	85%	67%	76%	61%	80%	71%	83%
H4	69%	52%	51%	44%	45%	33%	42%
H5	49%	49%	46%	67%	54%	46%	47%
H6	90%	58%	54%	47%	45%	50%	51%

Total amount compared to H3κ2	κ1	κ2	κ3	κ4	λ1	λ2	λ3
H1A	289%	94%	166%	272%	20%	150%	78%
H1B	219%	122%	89%	139%	117%	158%	101%
H2	186%	223%	208%	182%	126%	60%	97%
H3	50%		71%	54%	59%	130%	47%
H4	37%	55%	60%	77%	195%	107%	251%
H5	98%	201%	167%	83%	93%	128%	115%
H6	65%	117%	89%	109%	299%	215%	278%

Soluble amount compared to H3κ2	κ1	κ2	κ3	κ4	λ1	λ2	λ3
H1A	191%	88%	121%	122%	26%	211%	76%
H1B	124%	95%	83%	107%	79%	142%	59%
H2	126%	204%	139%	130%	66%	50%	70%
H3	63%	-	81%	49%	69%	143%	61%
H4	40%	47%	49%	54%	95%	55%	125%
H5	69%	158%	116%	80%	72%	84%	84%
H6	85%	122%	87%	77%	162%	162%	212%

	McPC
soluble	38%
%H3k2 total	117%
%H3k2 soluble	69%

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- EP 0368684 B1 **[0003]**
- EP 9603647 W **[0152]**
- EP 95113021 A **[0152]**

Non-patent literature cited in the description

- **JOHN DE KRUIF et al.** *J. Mol. Biol,* April 1995, vol. 248, 97-105 **[0006]**
- **BARBAS III, C.F. ; BAIN, J.D. ; HOEKSTRA, D.M. ; LERNER, R.A.** *PNAS,* 1992, vol. 89, 4457-4461 **[0145]**
- **BETTER, M. ; CHANG, P. ; ROBINSON, R. ; HORWITZ, A.H.** *Science,* 1988, vol. 240, 1041-1043 **[0145]**
- **BLAKE, M.S. ; JOHNSTON, K.H. ; RUSSEL-JONES, G.J. ; GOTSCHLICH, E.C.** *Anal. Biochem.,* 1984, vol. 136, 175-179 **[0145]**
- **CARTER, P. ; KELLY, R.F. ; RODRIGUES, M.L. ; SNEDECOR, B. ; COVRRUBIAS, M. ; VELLIGAN, M.D. ; WONG, W.L.T. ; ROWLAND, A.M. ; KOTTS, C.E. ; CARVER, M.E.** *Bio/Technology,* 1992, vol. 10, 163-167 **[0145]**
- **CHOTHIA, C. ; LESK, A.M.** *J. Biol. Chem.,* 1987, vol. 196, 910-917 **[0145]**
- **CHOTHIA, C. ; LESK, A.M. ; GHERARDI, E. ; TOMLINSON, I.A. ; WALTER, G. ; MARKS, J.D. ; LLEWELYN, M.B. ; WINTER, G.** *J. Mol. Biol.,* 1992, vol. 227, 799-817 **[0145]**
- **CHOTHIA, C. ; LESK, A.M. ; TRAMONTANO, A. ; LEVITT, M. ; SMITH-GILL, S.J. ; AIR, G. ; SHERIFF, S. ; PADLAN, E.A. ; DAVIES, D. ; TULIP, W.R.** *Nature,* 1989, vol. 342, 877-883 **[0145]**
- **CHUCHANA, P. ; BLANCHER, A. ; BROCKLY, F. ; ALEXANDRE, D. ; LEFRANC, G ; LEFRANC, M.-P.** *Eur. J. Immunol.,* 1990, vol. 20, 1317-1325 **[0145]**
- **COX, J.P.L. ; TOMLINSON, I.M. ; WINTER, G.** *Eur. J. Immunol.,* 1994, vol. 24, 827-836 **[0145]**
- **GE, L. ; KNAPPIK, A. ; PACK, P. ; FREUND, C. ; PLÜCKTHUN, A.** Antibody Engineering. Borrebaeck. Oxford University Press, 1995, 229-266 **[0145]**
- **GILL, S.C. ; VON HIPPEL, P.H.** *Anal. Biochem.,* 1989, vol. 182, 319.326 **[0145]**
- **HOCHULI, E. ; BANNWARTH, W. ; DÖBELI, H. ; GENTZ, R. ; STÜBER, D.** *Bio/Technology,* 1988, vol. 6, 1321-1325 **[0145]**
- **HOPP, T.P. ; PRICKETT, K.S. ; PRICE, V.L. ; LIBBY, R.T. ; MARCH, C.J. ; CERRETTI, D.P. ; URDAL, D.L. ; CONLON, P.J.** *Bio/Technology,* 1988, vol. 6, 1204-1210 **[0145]**
- **KABAT, E.A. ; WU, T.T. ; PERRY, H.M. ; GOTTESMANN, K.S. ; FOELLER, C.** *Sequences of proteins of immunological interest,* 1991 **[0145]**
- **KNAPPIK, A. ; PLÜCKTHUN, A.** *Biotechniques,* 1994, vol. 17, 754-761 **[0145]**
- **KNAPPIK, A. ; PLÜCKTHUN, A.** *Protein Engineering,* 1995, vol. 8, 81-89 **[0145]**
- **KUNKEL, T.A. ; BEBENEK, K. ; MCCLARY, J.** *Methods in Enzymol.,* 1991, vol. 204, 125-39 **[0145]**
- **LINDNER, P. ; GUTH, B. ; WÜLFING, C. ; KREBBER, C. ; STEIPE, B. ; MÜLLER, F. ; PLÜCKTHUN, A.** *Methods: A Companion to Methods Enzymol.,* 1992, vol. 4, 41-56 **[0145]**
- **LOWMAN, H.B. ; BASS, S.H. ; SIMPSON, N. ; WELLS, J.A.** *Biochemistry,* 1991, vol. 30, 10832-10838 **[0145]**
- **PACK, P. ; PLÜCKTHUN, A.** *Biochemistry,* 1992, vol. 31, 1579-1584 **[0145]**
- **PACK, P. ; KUJAU, M. ; SCHROECKH, V. ; KNÜPFER, U. ; WENDEROTH, R. ; RIESENBERG D. ; PLÜCKTHUN, A.** *Bio/Technology,* 1993, vol. 11, 1271-1277 **[0145]**
- **PACK, P.** Ph.D. thesis. *Ludwig-Maximilians-Universitat München,* 1994 **[0145]**
- **PERLAK, F. J.** *Nuc. Acids Res,* 1990, vol. 18, 7457-7458 **[0145]**
- **PLÜCKTHUN, A. ; KREBBER, A. ; KREBBER, C. ; HORN, U. ; KNÜPFER, U. ; WENDEROTH, R. ; NIEBA, L. ; PROBA, K. ; RIESENBERG, D.** A practical approach. Antibody Engineering. IRL Press, 1996, 203-252 **[0145]**
- **PRODROMOU, C. ; PEARL, L.H.** *Protein Engineering,* 1992, vol. 5, 827-829 **[0145]**
- **ROSENBERG, S.A. ; LOTZE, M.T.** *Ann. Rev. Immunol,* 1986, vol. 4, 681-709 **[0145]**
- **SKERRA, A. ; PLUCKTHUN, A.** *Science,* 1988, vol. 240, 1038-1041 **[0145]**
- **SKERRA, A. ; PFITZINGER, I. ; PLÜCKTHUN, A.** *Bio/Technology,* 1991, vol. 9, 273-278 **[0145]**
- **SUTHERLAND, L. ; DAVIDSON, J. ; GLASS, L.L. ; JACOBS, H.T.** *BioTechniques,* 1995, vol. 18, 458-464 **[0145]**

- **TOMLINSON, I.M. ; WALTER, G. ; MARKS, J.D. ; LLEWELYN, M.B. ; WINTER, G.** *J. Mol. Biol.,* 1992, vol. 227, 776-798 **[0145]**
- **ULLRICH, H.D. ; PATTEN, P.A. ; YANG, P.L. ; ROMESBERG, F.E. ; SCHULTZ, P.G.** *Proc. Natl. Acad. Sci. USA,* 1995, vol. 92, 11907-11911 **[0145]**
- **VAN DIJK, K.W. ; MORTARI, F. ; KIRKHAM, P.M. ; SCHROEDER JR., H.W. ; MILNER, E.C.B.** *Eur. J. Immunol.,* 1993, vol. 23, 832-839 **[0145]**
- **VIRNEKÄS, B. ; GE, L. ; PLÜCKTHUN, A. ; SCHNEIDER, K.C. ; WELLNHOFER, G. ; MORONEY, S.E.** *Nucleic Acids Research,* 1994, vol. 22, 5600-5607 **[0145]**
- **VITETTA, E.S. ; THORPE, P.E. ; UHR, J.** *Immunol. Today,* 1993, vol. 14, 253-259 **[0145]**
- **WILLIAMS, S.C. ; WINTER, G.** *Eur. J. Immunol.,* 1993, vol. 23, 1456-1481 **[0145]**
- **WINTER, G. ; GRIFFITHS, A.D. ; HAWKINS, R.E. ; HOOGENBOOM, H.R.** *Ann. Rev. Immunol,* 1994, vol. 12, 433-455 **[0145]**
- **ALESCIO-ZONTA, L. ; BAGLIONI, C.** *Eur.J.Biochem.,* 1970, vol. 15, 450-463 **[0146]**
- **ANDREWS, D.W. ; CAPRA, J.D.** *Biochemistry,* 1981, vol. 20, 5816-5822 **[0146]**
- **ANDRIS, J.S. ; EHRLICH, P.H. ; OSTBERG, L. ; CAPRA, J.D.** *J.Immunol.,* 1992, vol. 149, 4053-4059 **[0146] [0147]**
- **ATKINSON, P.M. ; LAMPMAN, G.W. ; FURIE, B.C. ; NAPARSTEK, Y. ; SCHWARTZ, R.S. ; STOLLAR, B.D. ; FURIE, B.** *J.Clin.Invest.,* 1985, vol. 75, 1138-1143 **[0146] [0148]**
- **AUCOUTURIER, P. ; BAUWENS, M. ; KHAMLICHI, A.A. ; DENOROY, L ; SPINELLI, S. ; TOUCHARD, G. ; PREUD'HOMME, J.-L. ; COGNE, M.** *J.Immunol.,* 1993, vol. 150, 3561-3568 **[0146]**
- **AVILA, M.A. ; VAZQUES, J. ; DANIELSSON, L. ; FERNANDEZ DE COSSIO, M.E. ; BORREBAECK, C.A.K.** *Gene,* 1993, vol. 127, 273-274 **[0146] [0148]**
- **BARBAS, C.F., III ; CROWE, JR., J.E. ; CABABA, D. ; JONES, T.M. ; ZEBEDEE, S.L. ; MURPHY, B.R. ; CHANOCK, R.M. ; BURTON, D.R.** *Proc.Natl.Acad.Sci.USA,* 1992, vol. 89, 10164-10168 **[0146] [0148]**
- **BARBAS, C.F., III et al.** *J-Mol-Biol.,* 1993, vol. 230, 812-23 **[0146]**
- **BENTLEY, D.L. ; RABBITTS, T.H.** *Nature,* 1980, vol. 288, 730-733 **[0146]**
- **BENTLEY, D.L. ; RABBITTS, T.H.** *Cell,* 1983, vol. 32, 181-189 **[0146]**
- **BENTLEY, D.L.** *Nature,* 1984, vol. 307, 77-80 **[0146]**
- **BHAT, N.M. ; BIEBER, M.M. ; CHAPMAN, C.J. ; STEVENSON, F.K. ; TENG, N.N.H.** *J.Immunol.,* 1993, vol. 151, 5011-5021 **[0146] [0147]**
- **BLAISON, G. ; KUNTZ, J.-L. ; PASQUALI, J.-L.** *Eur.J.Immunol.,* 1991, vol. 21, 1221-1227 **[0146]**
- **BRAUN, H. ; LEIBOLD, W. ; BARNIKOL, H.U. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1971, vol. 352, 647-651 **[0146]**
- *Z.Physiol.Chem.,* 1972, vol. 353, 1284-1306 **[0146]**
- **CAPRA, J.D. ; KEHOE, J.M.** *Adv.Immunology,* 1975, vol. 20, 1-40 **[0146] [0146]**
- **ANDREWS, D.W. ; CAPRA. J.D.** *Proc.Nat.Acad.Sci.USA,* 1981, vol. 78, 3799-3803 **[0146]**
- **LEDFORD. D.K. ; GONI, F. ; PIZZOLATO, M. ; FRANKLIN, E.C. ; SOLOMON, A. ; FRANGIONE, B.** *J.Immunol.,* 1983, vol. 131, 1322-1325 **[0146]**
- **CHASTAGNER, P. ; THEZE, J. ; ZOUALI, M.** *Gene,* 1991, vol. 101, 305-306 **[0146]**
- **CHEN, P.P. ; ROBBINS, D.L. ; JIRIK, F.R. ; KIPPS, T.J. ; CARSON, D.A.** *J.Exp.Med,* 1987, vol. 166, 1900-1905 **[0146] [0146]**
- **LIU, M.-F. ; ROBBINS, D.L. ; CROWLEY, J.J. ; SINHA, S. ; KOZIN, F. ; KIPPS, T.J. ; CARSON, D.A. ; CHEN.P.P.** *J.Immunol.,* 1989, vol. 142, 688-694 **[0146]**
- **CHERSI, A. ; NATALI, P.G.** *Immunochemistry,* 1978, vol. 15, 585-589 **[0146]**
- **CO, M.S. ; DESCHAMPS, M. ; WHITLEY, R.J. ; QUEEN, C.** *Proc.Natl.Acad.Sci.USA,* 1991, vol. 88, 2869-2873 **[0146]**
- **CUISINIER, A.-M. ; FUMOUX, F. ; FOUGEREAU, M. ; TONNELLE, C.** *Mol.Immunol.,* 1992, vol. 29, 1363-1373 **[0146] [0147] [0148]**
- **DAVIDSON, A. ; MANHEIMER-LORY, A. ; ARANOW, C. ; PETERSON, R. ; HANNIGAN, N. ; DIAMOND, B.** *J.Clin.Invest.,* 1990, vol. 85, 1401-1409 **[0146]**
- **DENOMME, G.A. ; MAHMOUDI, M. ; EDWARDS, J.Y. ; MASSICOTTE, H. ; CAIRNS, E. ; BELL, D.A.** *Hum.Antibod.Hybridomas,* 1993, vol. 4, 98-103 **[0146]**
- **DERSIMONIAN, H. ; MCADAM, K.P.W.J. ; MACKWORTH-YOUNG, C. ; STOLLAR, B.D.** *J.Iinmunol.,* 1989, vol. 142, 4027-4033 **[0146]**
- **DREYER, W.J. ; GRAY, W.R. ; HOOD, L.** Symp. Quantitative Biol. Cold Spring Harbor, 1967, vol. 32, 353-367 **[0146]**
- **EBELING, S.B. ; SCHUTTE, M.E.M. ; LOGTENBERG, T.** *Eur.J.Immunol.,* 1993, vol. 23, 1405-1408 **[0146]**
- **EULITZ, M. ; KLEY, H.-P.** *Immunochem,* 1977, vol. 14, 289-297 **[0146]**
- **EULITZ, M. ; LINKE, R.P.** *Z.Physiol.Chem.,* 1982, vol. 363, 1347-1358 **[0146]**
- **EULITZ, M. ; BREUER, M. ; EBLEN, A. ; WEISS, D.T. ; SOLOMON, A.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0146]**
- **EULITZ, M. ; GOTZE, D. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1972, vol. 353, 487-491 **[0146]**
- **EULITZ, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1974, vol. 355, 842-866 **[0146]**
- **EULITZ, M. ; KLEY, H.P. ; ZEITLER, H.J.** *Z.Physiol.Chem.,* 1979, vol. 360, 725-734 **[0146]**

- **EZAKI, I. ; KANDA, H. ; SAKAI, K. ; FUKUI, N. ; SHINGU, M. ; NOBUNAGA, M. ; WATANABE, T.** *Arthritis And Rheumatism,* 1991, vol. 34, 343-350 **[0146]**
- **FELGENHAUER, M. ; KOHL, J. ; RUKER, F.** *Nucl.Acids Res.,* 1990, vol. 18, 4927 **[0146] [0148]**
- **FERRI. G. ; STOPPINI, M. ; IADAROLA, P. ; BELLOTTI, V. ; MERLINI, G.** *Biochim.Biophys.Acta,* 1989, vol. 995, 103-108 **[0146]**
- **GILLIES, S.D. ; DORAI, H. ; WESOLOWSKI, J. ; MAJEAU, G. ; YOUNG, D. ; BOYD, J. ; GARDNER, J. ; JAMES, K.** *Bio/Tech.,* 1989, vol. 7, 799-804 **[0146] [0148]**
- **GONI, F. ; FRANGIONE, B.** *Proc.Nat.Acad.Sci.USA,* 1983, vol. 80, 4837-4841 **[0146] [0148]**
- **GONI, F.R. ; CHEN, P.P. ; MCGINNIS, D. ; ARJONILLA, M.L. ; FERNANDEZ, J. ; CARSON, D. ; SOLOMON, A. ; MENDEZ, E. ; FRANGIONE, B.** *J.Immunol.,* 1989, vol. 142, 3158-3163 **[0146]**
- **GORMAN. S.D. ; CLARK, M.R. ; ROULEDGE, E.G. ; COBBOLD, S.P. ; WALDMANN, H.** *Proc.Natl.Acad.Sci.USA,* 1991, vol. 88, 4181-4185 **[0146]**
- **GOTTLIEB, P.D. ; CUNNINGHAM, B.A. ; RUTISHAUSER, U. ; EDELMAN, G.M.** *Biochemistry,* 1970, vol. 9, 3155-3161 **[0146]**
- **GRIFFITHS, A.D. ; MALMQVIST, M. ; MARKS, J.D. ; BYE, J.M. ; EMBLETON, M.J. ; MCCAFFERTY, J. ; BAIER, M. ; HOLLIGER, K.P. ; GORICK, B.D. ; HUGHES-JONES, N.C.** *Embo J.,* 1993, vol. 12, 725-734 **[0146] [0147] [0148]**
- **HIETER, P.A. ; MAX, E.E. ; SEIDMAN, J.G. ; MAIZEL, J.V., JR. ; LEDER, P.** *Cell,* 1980, vol. 22, 197-207 **[0146]**
- **KLOBECK, H.G ; MEINDL, A. ; COMBRIATO, G. ; SOLOMON, A. ; ZACHAU, H.G.** *Nucl.Acids Res.,* 1985, vol. 13, 6499-6513 **[0146] [0146]**
- **HILSCHMANN, N. ; CRAIG, L.C.** *Proc.Nat.Acad.Sci.USA,* 1965, vol. 53, 1403-1409 **[0146] [0146]**
- **HILSCHMANN, N.** *Z.Pliysiol.Chem.,* 1967, vol. 348, 1077-1080 **[0146]**
- **HILSCHMANN, N.** *Z.Physiol.Chem.,* 1967, vol. 348, 1718-1722 **[0146]**
- **HILSCHMANN, N.** *Naturwissenschaften,* 1969, vol. 56, 195-205 **[0146]**
- **HIRABAYASHI, Y. ; MUNAKATA, Y. ; SASAKI, T. ; SANO, H.** *Nucl.Acids Res.,* 1992, vol. 20, 2601 **[0146] [0148]**
- **JAENICHEN, H.-R. ; PECH, M. ; LINDENMAIER, W. ; WILDGRUBER, N. ; ZACHAU, H.G.** *Nuc.Acids Res.,* 1984, vol. 12, 5249-5263 **[0146]**
- **JIRIK. F.R. ; SORGE, J. ; FONG, S. ; HEITZMANN, J.G. ; CURD, J.G. ; CHEN, P.P. ; GOLDFIEN, R ; CARSON, D.A.** *Proc.Nat.Acad.Sci.USA,* 1986, vol. 83, 2195-2199 **[0146]**
- **KAPLAN, A.P. ; METZGER, H.** *Biochemistry,* 1969, vol. 8, 3944-3951 **[0146] [0146]**
- **KLAPPER, D.G. ; CAPRA, J.D.** *Ann.Immunol.(Inst.Pasteur,* 1976, vol. 127c, 261-271 **[0146] [0146]**
- **KENNEDY, M.A.** *J.Exp.Med.,* 1991, vol. 173, 1033-1036 **[0146]**
- **KIM, H.S. ; DEUTSCH, H.F.** *Immunol.,* 1988, vol. 64, 573-579 **[0146]**
- **KIPPS, T.J. ; TOMHAVE, E. ; CHEN, P.P. ; CARSON, D.A.** *J.Exp.Med.,* 1988, vol. 167, 840-852 **[0146]**
- **KIPPS, T.J. ; TOMHAVE, E. ; CHEN, P.P. ; FOX, R.I.** *J.Immunol.,* 1989, vol. 142, 4261-4268 **[0146]**
- **KLEIN, U. ; KUPPERS, R. ; RAJEWSKY, K.** *Eur.J.Immunol.,* 1993, vol. 23, 3272-3277 **[0146]**
- **KLOBECK, H.G. ; BORNKAMMM, G.W. ; COMBRIATO, G. ; MOCIKAT. R. ; POHLENZ, H.D. ; ZACHAU, H.G.** *Nucl.Acids Res.,* 1985, vol. 13, 6515-6529 **[0146]**
- **KLOBECK, H.G. ; COMBRIATO, G. ; ZACHAU, H.G.** *Nuc.Acids Res.,* 1984, vol. 12, 6995-7006 **[0146]**
- **KLOBECK, H.G. ; SOLOMON, A. ; ZACHAU, H.G.** *Nature,* 1984, vol. 309, 73-76 **[0146]**
- **KNIGHT, G.B. ; AGNELLO, V. ; BONAGURA, V. ; BARNES, J.L. ; PANKA, D.J. ; ZHANG, Q-X.** *J.Exp.Med.,* 1993, vol. 178, 1903-1911 **[0146]**
- **KOHLER, H. ; SHIMIZU, A. ; PAUL, C. ; PUTNAM, F.W.** *Science,* 1970, vol. 169, 56-59 **[0146]**
- **KRATZIN, H. ; YANG, C.Y. ; KRUSCHE, J.U. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1980, vol. 361, 1591-1598 **[0146]**
- **KUNICKI, T.J. ; ANNIS, D.S. ; GORSKI, J. ; NUGENT, D.J.** *J.Autoimmunity,* 1991, vol. 4, 433-446 **[0146] [0148]**
- **LARRICK, J.W. ; WALLACE, E.F. ; COLOMA, M.J. ; BRUDERER, U. ; LANG, A.B. ; FRY, K.E.** *Immunological Reviews,* 1992, vol. 130, 69-85 **[0146] [0148]**
- **LAURE, C.J. ; WATANABE, S. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1973, vol. 354, 1503-1504 **[0146]**
- **LEDFORD, D.K. ; GONI, F. ; PIZZOLATO, M. ; FRANKLIN, E.C. ; SOLOMON, A. ; FRANGIONE, B.** *J.Immunol.,* 1983, vol. 131, 1322-1325 **[0146] [0146] [0146]**
- **PONS-ESTEL, B. ; GONI, F. ; SOLOMON, A. ; FRANGIONE, B.** *J.Exp.Med.,* 1984, vol. 160, 893 **[0146]**
- **LEVY. S. ; MENDEL, E. ; KON, S. ; AVNUR, Z. ; LEVY, R.** *J.Exp.Med.,* 1988, vol. 168, 475-489 **[0146]**
- **LIEPNIEKS, J.J. ; DWULET, F.E. ; BENSON, M.D.** *Mol.Immunol.,* 1990, vol. 27, 481-485 **[0146]**
- **MANHEIMER-LORY, A. ; KATZ, J.B. ; PILLINGER, M. ; GHOSSEIN, C. ; SMITH, A. ; DIAMOND, B.** *J.Exp.Med.,* 1991, vol. 174, 1639-1652 **[0146]**

- **MANTOVANI, L. ; WILDER, R.L. ; CASALI, P.** *J.Inununol.,* 1993, vol. 151, 473-488 **[0146]**
- **MARIETTE, X. ; TSAPIS, A. ; BROUET, J.-C.** *Eur.J.Immunol.,* 1993, vol. 23, 846-851 **[0146] [0148]**
- **MARKS, J.D. ; HOOGENBOOM, H.R. ; BONNERT, T.P. ; MCCAFFERTY, J. ; GRIFFITHS, A.D. ; WINTER. G.** *J.Mol.Biol.,* 1991, vol. 222, 581-597 **[0146]**
- **MARSH, P. ; MILLS, F. ; GOULD, H.** *Nuc.Acids Res.,* 1985, vol. 13, 6531-6544 **[0146]**
- **MIDDAUGH, C.R. ; LITMAN, G.W.** *J.Biol.Chem.,* 1987, vol. 262, 3671-3673 **[0146]**
- **MILSTEIN, C. ; DEVERSON, E.V.** *Biochem.J.,* 1971, vol. 123, 945-958 **[0146]**
- **MILSTEIN, C.** *Febs Letters,* 1969, vol. 2, 301-304 **[0146] [0146]**
- **MILSTEIN, C.P. ; DEVERSON, E.V.** *Eur.J.Biochem.,* 1974, vol. 49, 377-391 **[0146]**
- **MORAN, M.J. ; ANDRIS, J.S. ; MATSUMATO, Y.-I. ; CAPRA, J.D. ; HERSH, E.M.** *Mol.Immunol.,* 1993, vol. 30, 1543-1551 **[0146] [0148]**
- **NAKATANI, T. ; NOMURA, N. ; HORIGOME, K. ; OHTSUKA, H. ; NOGUCHI, H.** *Bio/Tech.,* 1989, vol. 7, 805-810 **[0146]**
- **NEWKIRK, M. ; CHEN, P.P. ; CARSON, D. ; POSNETT, D. ; CAPRA, J.D.** *Mol.Inimunol.,* 1986, vol. 23, 239-244 **[0146]**
- **NEWKIRK, M.M. ; GRAM, H. ; HEINRICH, G.F. ; OSTBERG, L. ; CAPRA, J.D. ; WASSERMAN, R.L.** *J.Clin.Invest.,* 1988, vol. 81, 1511-1518 **[0146] [0148]**
- **NEWKIRK, M.M. ; MAGEED, R.A. ; JEFFERIS, R. ; CHEN, P.P. ; CAPRA, J.D.** *J.Exp.Med.,* 1987, vol. 166, 550-564 **[0146] [0148]**
- **OLEE, B.T. ; LU, E.W. ; HUANG, D.-F. ; SOTO-GIL, R.W. ; DEFTOS, M. ; KOZIN, F. ; CARSON, D.A. ; CHEN, P.P.** *J.Exp.Med.,* 1992, vol. 175, 831-842 **[0146] [0147] [0148]**
- **PALIN, W. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1973, vol. 354, 1651-1654 **[0146]**
- *Z.Physiol.Chem.,* 1975, vol. 354, 1651-1654 **[0146]**
- *Z.Physiol.Chem.,* vol. 356, 167-191 **[0146]**
- **PASCUAL, V. ; VICTOR, K. ; LELSZ, D. ; SPELLERBERG, M.B. ; HAMBLIN, T.J. ; THOMPSON, K.M. ; RANDEN, I. ; NATVIG, J. ; CAPRA, J.D. ; STEVENSON, F.K.** *J.Immunol.,* 1991, vol. 146, 4385-4391 **[0146] [0148]**
- **PASCUAL, V. ; VICTOR, K. ; RANDEN, I. ; THOMPSON, K. ; STEINITZ, M. ; FORRE, O. ; FU, S.-M. ; NATVIG, J.B. ; CAPRA, J.D.** *Scand.J.Immunol.,* 1992, vol. 36, 349-362 **[0146] [0147]**
- **PECH, M. ; ZACHAU, H.G.** *Nuc.Acids Res.,* 1984, vol. 12, 9229-9236 **[0146] [0149]**
- **PECH, M. ; JAENICHEN, H.-R. ; POHLENZ, H.-D. ; NEUMAIER, P.S. ; KLOBECK, H.-G. ; ZACHAU, H.G.** *J.Mol.Biol.,* 1984, vol. 176, 189-204 **[0146] [0149]**
- **PONS-ESTEL, B. ; GONI, F. ; SOLOMON, A. ; FRANGIONE, B.** *J.Exp.Med.,* 1984, vol. 160, 893-904 **[0146]**
- **PORTOLANO, S. ; MCLACHLAN, S.M. ; RAPOPORT, B.** *J.Immunol.,* 1993, vol. 151, 2839-2851 **[0146]**
- **PORTOLANO, S. ; SETO, P. ; CHAZENBALK, G.D. ; NAGAYAMA, Y. ; MCLACHLAN, S.M. ; RAPOPORT, B.** *Biochem.Biophys.Res.Commun.,* 1991, vol. 179, 372-377 **[0146] [0148]**
- **PRATT, L.F. ; RASSENTI, L. ; LARRICK, J. ; ROBBINS, B. ; BANKS, P.M. ; KIPPS, T.J.** *J.Immunol.,* 1989, vol. 143, 699-705 **[0146]**
- **PRELLI, F. ; TUMMOLO, D. ; SOLOMON, A. ; FRANGIONE, B.** *J.Immunol.,* 1986, vol. 136, 4169-4173 **[0146]**
- **PUTNAM, F.W. ; WHITLEY, E.J., JR. ; PAUL, C. ; DAVIDSON, J.N.** *Biochemistry,* 1973, vol. 12, 3763-3780 **[0146]**
- **RANDEN, I. ; PASCUAL, V. ; VICTOR, K. ; THOMPSON, K.M. ; FORRE, O. ; CAPRA, J.D. ; NATVIG, J.B.** *Eur.J.Lnmunol.,* 1993, vol. 23, 1220-1225 **[0146]**
- **RASSENTI, L.Z. ; PRATT, L.F. ; CHEN, P.P. ; CARSON, D.A. ; KIPPS, T.J.** *J.Immunol.,* 1991, vol. 147, 1060-1066 **[0146]**
- **REIDL, L.S. ; FRIEDMAN. D.F. ; GOLDMAN, J. ; HARDY, RR. ; JEFFERIES, L.C. ; SILBERSTEIN, L.E.** *J.Immunol.,* 1991, vol. 147, 3623-3631 **[0146]**
- **RIECHMANN, L. ; CLARK, M. ; WALDMANN, H. ; WINTER, G.** *Nature,* 1988, vol. 332, 323-327 **[0146]**
- **RIESEN, W. ; RUDIKOFF, S. ; ORIOL, R. ; POTTER, M.** *Biochemistry,* 1975, vol. 14, 1052-1057 **[0146]**
- **RIESEN, W.F. ; BRAUN, D.G. ; JATON, J.C.** *Proc.Nat.Acad.Sci.USA,* 1976, vol. 73, 2096-2100 **[0146]**
- **RIESEN, W.F. ; JATON, J.C.** *Biochemistry,* 1976, vol. 15, 3829 **[0146]**
- **RODILLA SALA, E. ; KRATZIN, D.H. ; PICK, A.I. ; HILSCHMANN, N.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0146]**
- **SCHIECHL, H. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1971, vol. 352, 111-115 **[0146]**
- *Z.Physiol.Chem.,* 1972, vol. 353, 345-370 **[0146]**
- **SCHNEIDER, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1974, vol. 355, 1164-1168 **[0146]**
- **SHEARMAN, C.W. ; POLLOCK, D. ; WHITE, G. ; HEHIR, K. ; MOORE, G.P. ; KANZY, E.J. ; KURRLE, R.** *J.Immunol.,* 1991, vol. 147, 4366-4373 **[0146]**
- **SHINODA, T.** *J.Biochem.,* 1973, vol. 73, 433-446 **[0146]**
- **SHINODA, T.** *J.Biochein.,* 1975, vol. 77, 1277-1296 **[0146]**
- **SHINODA, T. ; TAKENAWA, T. ; HOSHI, A. ; ISOBE, T.** Amyloid And Amyloidosis. Kluwer Academic Publishers, 1990, 157 **[0146]**

- **SILBERSTEIN, L.E. ; LITWIN, S. ; CARMACK. C.E.** *J.Exp.Med.,* 1989, vol. 169, 1631-1643 **[0146]**
- **SIMS, M.J. ; HASSAL, D.G. ; BRETT, S. ; ROWAN, W. ; LOCKYER, M.J. ; ANGEL, A. ; LEWIS, A.P. ; HALE, G. ; WALDMANN, H. ; CROWE, J.S.** *J.Immunol.,* 1993, vol. 151, 2296-2308 **[0146]**
- **SPATZ, L.A. ; WONG, K.K. ; WILLIAMS, M. ; DESAI, R. ; GOLIER, J. ; BERMAN, J.E. ; ALT, F.W. ; LATOV, N.** *J.Immunol.,* 1990, vol. 144, 2821-2828 **[0146] [0148]**
- **STAVNEZER, J. ; KEKISH, O. ; BATTER, D. ; GRENIER, J. ; BALAZS, I. ; HENDERSON, E. ; ZEGERS, B.J.M.** *Nucl.Acids Res.,* 1985, vol. 13, 3495-3514 **[0146] [0149]**
- **STRAUBINGER, B. ; THIEBE, R. ; PECH, M. ; ZACHAU, H.G.** *Gene,* 1988, vol. 69, 209-214 **[0146]**
- **SUTER, L. ; BARNIKOL, H.U. ; WATANABE, S. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1969, vol. 350, 275-278 **[0146]**
- *Z.Physiol.Chem.,* 1972, vol. 353, 189-208 **[0146]**
- **TEMPEST, P.R. ; BREMNER, P. ; LAMBERT, M. ; TAYLOR, G. ; FURZE, J.M. ; CARR, F.J. ; HARRIS, W.J.** *Bio/Tech.,* 1991, vol. 9, 266-271 **[0146]**
- **TITANI, K. ; SHINODA, T. ; PUTNAM, F.W.** *J.Biol.Chem.,* 1969, vol. 244, 3550-3560 **[0146]**
- **TOFT, K.G. ; OLSTAD, O.K. ; SLETTEN, K. ; WESTERMARK, P.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0146]**
- **VAN ES, J.H. ; AANSTOOT, H. ; GINELIG-MEYLING, F.H.J. ; DERKSEN, R.H.W.M. ; LOGTENBERG, T.** *J.Immunol.,* 1992, vol. 149, 2234-2240 **[0146]**
- **VICTOR, K.D. ; PASCUAL, V. ; LEFVERT, A.K. ; CAPRA, J.D.** *Mol.Immunol.,* 1992, vol. 29, 1501-1506 **[0146] [0148]**
- **VICTOR, K.D. ; PASCUAL, V. ; WILLIAMS, C.L. ; LENNON, V.A. ; CAPRA, J.D.** *Eur.J.Immunol.,* 1992, vol. 22, 2231-2236 **[0146] [0148]**
- **VICTOR, K.D. ; RANDEN, I. ; THOMPSON, K. ; FORRE, O. ; NATVIG, J.B. ; FU, S.M. ; CAPRA, J.D.** *J.Clin.Invest.,* 1991, vol. 87, 1603-1613 **[0146] [0147]**
- **WAGNER, S.D. ; LUZZATTO, L.** *Eur.J.Immunol.,* 1993, vol. 23, 391-397 **[0146]**
- **WATANABE, S. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1970, vol. 351, 1291-1295 **[0146]**
- **WEISBART, R.H. ; WONG, A.L. ; NORITAKE, D. ; KACENA, A. ; CHAN, G. ; RULAND, C. ; CHIN, E. ; CHEN, I.S.Y. ; ROSENBLATT, J.D.** *J.Immunol.,* 1991, vol. 147, 2795-2801 **[0146]**
- **WENG, N.-P. ; YU-LEE, L.-Y. ; SANZ, I. ; PATTEN, B.M. ; MARCUS, D.M.** *J.Immunol.,* 1992, vol. 149, 2518-2529 **[0146] [0147]**
- **WINKLER, T.H. ; FEHR, H. ; KALDEN, J.R.** *Eur.J.Immunol.,* 1992, vol. 22, 1719-1728 **[0146] [0147] [0148] [0151]**
- **ALEXANDRE, D. ; CHUCHANA, P. ; BROCKLY, F. ; BLANCHER, A. ; LEFRANC, G. ; LEFRANC, M.-P.** *Nuc.Acids Res.,* 1989, vol. 17, 3975 **[0147]**
- **ANDERSON, M.L.M. ; BROWN, L. ; MCKENZIE, E. ; KELLOW, J.E. ; YOUNG, B.D.** *Nuc.Acids Res.,* 1985, vol. 13, 2931-2941 **[0147]**
- **ANDRIS, J.S. ; BRODEUR, B.R. ; CAPRA, J.D.** *Mol.Immunol.,* 1993, vol. 30, 1601-1616 **[0147] [0148]**
- **BACZKO, K. ; BRAUN, D.G. ; HESS, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1970, vol. 351, 763-767 **[0147]**
- **BACZKO, K. ; BRAUN, D.G. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1974, vol. 355, 131-154 **[0147]**
- **BERINSTEIN, N. ; LEVY, S. ; LEVY, R.** *Science,* 1989, vol. 244, 337-339 **[0147]**
- **CAIRNS, E. ; KWONG, P.C. ; MISENER, V. ; IP, P. ; BELL, D.A. ; SIMINOVITCH, K.A.** *J.Immunol.,* 1989, vol. 143, 685-691 **[0147] [0148]**
- **CARROLL, W.L. ; YU, M. ; LINK, M.P. ; KORSMEYER, S.J.** *J.Immunol.,* 1989, vol. 143, 692-698 **[0147] [0148]**
- **CHEN, B.L. ; POLJAK, R.J.** *Biochemistry,* 1974, vol. 13, 1295-1302 **[0147]**
- **CHEN, B.L. ; CHIU, Y.Y.H. ; HUMPHREY, R.L. ; POLJAK, R.J.** *Biochim.Biophys.Acta,* 1978, vol. 537, 9-21 **[0147]**
- **COMBRIATO, G. ; KLOBECK, H.G.** *Eur.J.Immunol.,* 1991, vol. 21, 1513-1522 **[0147] [0150]**
- **DWULET, F.E. ; STRAKO, K. ; BENSON, M.D.** *Scand.J.Immunol.,* 1985, vol. 22, 653-660 **[0147]**
- **ELAHNA, P. ; LIVNEH, A. ; MANHEIMER-LORY, A.J. ; DIAMOND, B.** *J.Immunol.,* 1991, vol. 147, 2771-2776 **[0147]**
- **ENGELHARD, M. ; HESS, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1974, vol. 355, 85-88 **[0147]**
- **ENGELHARD, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1975, vol. 356, 1413-1444 **[0147]**
- **EULITZ, M.** *Eur.J.Biochem.,* 1974, vol. 50, 49-69 **[0147]**
- **EULITZ, M. ; BREUER, M. ; LINKE, R.P.** *Biol.Che.Hoppe-Seyler,* 1987, vol. 368, 863-870 **[0147]**
- **EULITZ, M. ; MURPHY, C. ; WEISS, D.T. ; SOLOMON, A.** *J.Immunol.,* 1991, vol. 146, 3091-3096 **[0147]**
- **FETT, J.W. ; DEUTSCH, H.F.** *Biochemistry,* 1974, vol. 13, 4102-4114 **[0147]**
- **FETT, J.W. ; DEUTSCH, H.F.** *Immunochem.,* 1976, vol. 13, 149-155 **[0147]**
- **JABUSCH, J.R. ; DEUTSCH, H.F.** *Mol.Immunol.,* 1982, vol. 19, 901-906 **[0147]**
- **FUREY, W. JR. ; WANG, B.C. ; YOO, C.S. ; SAX, M.** *J.Mol.Biol.,* 1983, vol. 167, 661-692 **[0147]**
- **FYKSE, E.-M. ; SLETTEN, K. ; HUSBY, G. ; CORNWELL, G.G., III.** *Biochem.J.,* 1988, vol. 256, 973-980 **[0147]**

- **GARVER, F.A. ; HILSCHMANN, N.** *Febs Letters,* 1971, vol. 16, 128-132 **[0147]**
- *Eur.J.Biochem.,* 1972, vol. 26, 10-32 **[0147]**
- **GAWINOWICZ, M.A. ; MERLINI, G. ; BIRKEN, S. ; OSSERMAN, E.F. ; KABAT, E.A.** *J.Immunol.,* 1991, vol. 147, 915-920 **[0147] [0148]**
- **GHISO, J. ; SOLOMON, A. ; FRANGIONE, B.** *J.Immunol.,* 1986, vol. 136, 716-719 **[0147]**
- **GULLASKEN, N. ; IDSO, H. ; NILSEN, R. ; SLETTEN, K. ; HUSBY, G. ; CORNWELL, G.G.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0147]**
- **HARINDRANATH, N. ; GOLDFARB, I.S. ; IKEMATSU, H. ; BURASTERO, S.E. ; WILDER, R.L. ; NOTKINS, A.L. ; CASALI, P.** *Int.Immunol.,* 1991, vol. 3, 865-875 **[0147] [0148]**
- **HOLM, E. ; SLETTEN, K. ; HUSBY, G.** *Biochem.J.,* 1986, vol. 239, 545-551 **[0147]**
- **HUGHES-JONES, N.C. ; BYE, J.M. ; BEALE, D. ; COADWELL, J.** *Biochem.J.,* 1990, vol. 268, 135-140 **[0147] [0148]**
- **KAMETANI, F. ; YOSHIMURA, K. ; TONOIKE, H. ; HOSHI, A. ; SHINODA, T. ; ISOBE, T.** *Biochem.Biophys.Res.Commun.,* 1985, vol. 126, 848-852 **[0147]**
- **KIEFER, C.R. ; MCGUIRE, B.S., JR. ; OSSERMAN, E.F. ; GARVER, F.A.** *J.Immunol.,* 1983, vol. 131, 1871-1875 **[0147]**
- **KIEFER, C.R. ; PATTON, H.M., JR. ; MCQUIRE, B.S., JR. ; GARVER, F.A.** *J.Immunol.,* 1984, vol. 124, 301-306 **[0147]**
- **KISHIMOTO, T. ; OKAJIMA, H. ; OKUMOTO, T. ; TANIGUCHI, M.** *Nud.Acids Res.,* 1989, vol. 17, 4385 **[0147]**
- **KLAFKI, H.-W. ; KRATZIN, H.D. ; PICK, A.I. ; ECKART, K. ; HILSCHMANN, N.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0147]**
- **KOHLER, H. ; RUDOFSKY, S. ; KLUSKENS, L.** *J.Irmnunology,* 1975, vol. 114, 415-421 **[0147]**
- **KOJIMA, M. ; ODANI, S. ; IKENAKA, T.** *Mol.Immunol.,* 1980, vol. 17, 1407-1414 **[0147]**
- **KOMORI, S. ; YAMASAKI, N. ; SHIGETA, M. ; ISOJIMA, S. ; WATANABE, T.** *Clin.Exp.immunol.,* 1988, vol. 71, 508-516 **[0147]**
- **KRATZIN, H.D. ; PALM, W. ; STANGEL, M. ; SCHMIDT, W.E. ; FRIEDRICH, J. ; HILSCHMANN, N.** *Biol.Cheni.Hoppe-Seyler,* 1989, vol. 370, 263-272 **[0147]**
- **KRATZIN, H.D. ; PICK, A.I. ; STANGEL, M. ; HILSCHMANN, N.** Amyloid And Amyloidosis. Kluwer Academic Publishers, 1990, 181 **[0147]**
- **LANGER, B. ; STEINMETZ-KAYNE, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1968, vol. 349, 945-951 **[0147]**
- **LARRICK, J.W. ; DANIELSSON, L. ; BRENNER, C.A. ; WALLACE, E.F. ; ABRAHAMSON, M. ; FRY, K.E. ; BORREBAECK, C.A.K.** *Bio/Tech.,* 1989, vol. 7, 934-938 **[0147]**
- **LEVY, S. ; MENDEL, E. ; KON, S. ; AVNUR, Z. ; LEVY, R.** *J.Exp.Med.,* 1988, vol. 168, 475-489 **[0147]**
- **LEWIS, A.P. ; LEMON, S.M. ; BARBER, K.A. ; MURPHY, P. ; PARRY, N.R. ; PEAKMAN, T.C. ; SIMS, M.J. ; WORDEN, J. ; CROWE, J.S.** *J.Immunol.,* 1993, vol. 151, 2829-2838 **[0147] [0148]**
- **LIU. V.Y.S. ; LOW, T.L.K. ; INFANTE, A. ; PUTNAM, F.W.** *Science,* 1976, vol. 193, 1017-1020 **[0147]**
- **INFANTE. A. ; PUTNAM, F.W.** *J.Biol.Chem.,* 1979, vol. 254, 9006-9016 **[0147]**
- **LOPEZ DE CASTRO, J.A. ; CHIU, Y.Y.H. ; POLJAK, R.J.** *Biochemistry,* 1978, vol. 17, 1718-1723 **[0147]**
- **MANTOVANI, L. ; WILDER, R.L. ; CASALI, P.** *J.Immunol.,* 1993, vol. 151, 473-488 **[0147] [0148]**
- **MARKS, J.D. ; HOOGENBOOM, H.R. ; BONNERT, T.P. ; MCCAFFERTY, J. ; GRIFFITHS, A.D. ; WINTER, G.** *J.Mol.Biol.,* 1991, vol. 222, 581-597 **[0147] [0148]**
- **MIHAESCO, E. ; ROY, J.-P. ; CONGY, N. ; PERAN-RIVAT, L. ; MIHAESCO, C.** *Eur.J.Biochem.,* 1985, vol. 150, 349-357 **[0147]**
- **MILSTEIN, C. ; CLEGG, J.B. ; JARVIS, J.M.** *Biochem.J.,* 1968, vol. 110, 631-652 **[0147]**
- **MORAN. M.J. ; ANDRIS, J.S. ; MATSUMATO, Y.-I. ; CAPRA, J.D. ; HERSH, E.M.** *Mol.Immunol.,* 1993, vol. 30, 1543-1551 **[0147]**
- **NABESHIMA, Y. ; IKENAKA, T.** *Mol.Immunol.,* 1979, vol. 16, 439-444 **[0147]**
- **PAUL, E. ; ILIEV, A.A. ; LIVNEH, A. ; DIAMOND, B.** *J.Immunol.,* 1992, vol. 149, 3588-3595 **[0147]**
- **PICK, A.I. ; KRATZIN, H.D. ; BAMIKOL-WATANABE, S. ; HILSCHMANN, N.** Amyloid And Amyloidosis. Kluwer Academic, 1990 **[0147]**
- **PONSTINGL, H. ; HILSCHMANN, N.** *Z.Physioi.Chem.,* 1969, vol. 350, 1148-1152 **[0147]**
- *Z.Physiol.Chem.,* 1971, vol. 352, 859-877 **[0147]**
- **PONSTINGL, H. ; HESS, M. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1968, vol. 349, 867-871 **[0147]**
- *Z.Physiol.Chem.,* 1971, vol. 352, 247-266 **[0147]**
- **RANDEN, L ; PASCUAL, V. ; VICTOR, K. ; THOMPSON, K.M. ; FORRE, O. ; CAPRA, J.D. ; NATVIG. J.B.** *Eur.J.Immunol.,* 1993, vol. 23, 1220-1225 **[0147]**
- **SCHOLZ, R. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1975, vol. 356, 1333-1335 **[0147]**
- **SETTMACHER, U. ; JAHN, S. ; SIEGEL, P. ; VON BAEHR, R. ; HANSEN, A.** *Mol.Immunol.,* 1993, vol. 30, 953-954 **[0147]**
- **SHINODA, T. ; TITANI, K. ; PUTNAM, F.W.** *J.Biol.Chem.,* 1970, vol. 245, 4475-4487 **[0147]**
- **SLETTEN, K. ; HUSBY, G. ; NATVIG, J.B.** *Scand.J.Immunol.,* 1974, vol. 3, 833-836 **[0147]**
- **SLETTEN, K. ; NATVIG, J.B. ; HUSBY, G. ; JUUL, J.** *Biochem.J.,* 1981, vol. 195, 561-572 **[0147]**
- **SOLOMON, A. ; FRANGIONE, B. ; FRANKLIN, E.C.** *J.Clin.invest.,* 1982, vol. 70, 453-460 **[0147]**

- **FRANGIONE, B. ; MOLOSHOK, T. ; SOLOMON, A.** *J.Immunol.,* 1983, vol. 131, 2490-2493 **[0147]**
- **TAKAHASHI, N. ; TAKAYASU, T. ; ISOBE, T. ; SHINODA, T. ; OKUYAMA, T. ; SHIMIZU, A.** *J.Biochem.,* 1979, vol. 86, 1523-1535 **[0147]**
- **TAKAHASHI, N. ; TAKAYASU, T. ; SHINODA, T. ; ITO, S. ; OKUYAMA, T. ; SHIMIZU, A.** *Biomed.Res.,* 1980, vol. 1, 321-333 **[0147]**
- **TAKAHASHI, Y. ; TAKAHASHI, N. ; TETAERT, D ; PUTNAM, F.W.** *Proc.Nat.Acad.Sci.USA,* 1983, vol. 80, 3686-3690 **[0147]**
- **TAKAYASU, T. ; TAKAHASHI, N. ; SHINODA, T. ; OKUYAMA, T. ; TOMIOKA, H.** *J.Biochem.,* 1980, vol. 89, 421-436 **[0147]**
- **TITANI. K. ; WIKLER, M. ; SHINODA, T. ; PUTNAM, F.W.** *J. Biol.Chem.,* 1970, vol. 245, 2171-2176 **[0147]**
- **TOFT, K.G. ; SLETTEN, K. ; HUSBY, G.** *Biol.Cliem.Hoppe-Seyler.,* 1985, vol. 366, 617-625 **[0147]**
- **TONOIKE, H. ; KAMETANI, F. ; HOSHI, A. ; SHINODA, T. ; ISOBE, T.** *Biochem.Biophys.Res.Commun.,* 1985, vol. 126, 1228-1234 **[0147]**
- **TONOIKE, H. ; KAMETANI, F. ; HOSHI, A. ; SHINODA, T. ; ISOBE, T.** *Febs Letters,* 1985, vol. 185, 139-141 **[0147]**
- **TSUJIMOTO, Y. ; CROCE, C.M.** *Nucl.Acids Res.,* 1984, vol. 12, 8407-8414 **[0147]**
- **TSUNETSUGU-YOKOTA, Y. ; MINEKAWA, T. ; SHIGEMOTO, K. ; SHIRASAWA, T. ; TAKEMORI, T.** *Mol.Immunol.,* 1992, vol. 29, 723-728 **[0147]**
- **TVETERAAS, T. ; SLETTEN, K. ; WESTERMARK, P.** *Biochem.J.,* 1985, vol. 232, 183-190 **[0147]**
- **VASICEK, T.J. ; LEDER, P.** *J.Exp.Med.,* 1990, vol. 172, 609-620 **[0147]**
- **WIKLER, M. ; PUTNAM, F.W.** *J.Biol.Chem.,* 1970, vol. 245, 4488-4507 **[0147]**
- **YAGO, K. ; ZENITA, K. ; OHWAKI, I. ; HARADA, Y. ; NOZAWA, S. ; TSUKAZAKI, K. ; IWAMORI, M. ; ENDO, N. ; YASUDA, N. ; OKUMA, M.** *MoLImmunol.,* 1993, vol. 30, 1481-1489 **[0147]**
- **YAMASAKI, N. ; KOMORI, S. ; WATANABE, T.** *Mol.Immunol.,* 1987, vol. 24, 981-985 **[0147]**
- **ZHU, D. ; KIM, H.S. ; DEUTSCH, H.F.** *Mol.Immunol.,* 1983, vol. 20, 1107-1116 **[0147]**
- **ZHU. D. ; ZHANG, H. ; ZHU, N. ; LUO, X.** *Scientia Sinica,* 1986, vol. 29, 746-755 **[0147]**
- **ADDERSON, E.E. ; AZMI, F.H. ; WILSON, P.M. ; SHACKELFORD, P.G. ; CARROLL, W.L.** *J.Immunol.,* 1993, vol. 151, 800-809 **[0148] [0151]**
- **ADDERSON, E.E. ; SHACKELFORD, P.G. ; QUINN, A. ; CARROLL, W.L.** *J.Immunol.,* 1991, vol. 147, 1667-1674 **[0148]**
- **AKAHORI, Y. ; KUROSAWA, Y. ; KAMACHI, Y. ; TORII, S. ; MATSUOKA, H.** *J.Clin.Invest.,* 1990, vol. 85, 1722-1727 **[0148]**
- **ANDRIS, J.S. ; EHRLICH, P.H. ; OSTBERG. L. ; CAPRA, J.D.** *J.Immunol.,* 1992, vol. 149, 4053-4059 **[0148]**
- **ANDRIS, J.S. ; JOHNSON, S. ; ZOLLA-PAZNER. S. ; CAPRA, J.D.** *Proc.Natl.Acad.Sci.USA,* 1991, vol. 88, 7783-7787 **[0148]**
- **ANKER, R. ; CONLEY, M.E. ; POLLOK, B.A.** *J.Exp.Med.,* 1989, vol. 169, 2109-2119 **[0148]**
- **BAKKUS, M.H.C. ; HEIRMAN, C. ; VAN RIET, I. ; VAN CAMP, B. ; THIELEMANS, K.** *Blood,* 1992, vol. 80, 2326-2335 **[0148]**
- **BARBAS, C.F., III ; COLLET, T.A. ; AMBERG, W. ; ROBEN, P. ; BINLEY, J.M. ; HOEKSTRA, D. ; CABABA, D. ; JONES, T.M. ; WILLIAMSON, R.A. ; PILKINGTON, G.R.** *J.Mol.Biol.,* 1993, vol. 230, 812-823 **[0148]**
- **BERMAN, J.E. ; HUMPHRIES, C.G. ; BARTH, J. ; ALT, F.W. ; TUCKER, P.W.** *J.Exp.Med.,* 1991, vol. 173, 1529-1535 **[0148]**
- **BERMAN, J.E. ; MELLIS, S.J. ; POLLOCK. R. ; SMITH, C.L. ; SUH, H. ; HEINKE, B. ; KOWAL, C. ; SURTI, U. ; CHESS, L. ; CANTOR, C.R.** *Embo J.,* 1988, vol. 7, 727-738 **[0148]**
- **BHAT, N.M. ; BIEBER, M.M. ; CHAPMAN, C.J. ; STEVENSON, F.K. ; TENG. N.N.H.** *J.Immunol.,* 1993, vol. 151, 5011-5021 **[0148]**
- **BIRD, J. ; GALILI, N. ; LINK, M. ; STITES, D. ; SKLAR, J.** *J.Exp.Med.,* 1988, vol. 168, 229-245 **[0148]**
- **CAI, J. ; HUMPHRIES, C. ; RICHARDSON, A. ; TUCKER, P.W.** *J.Exp.Med.,* 1992, vol. 176, 1073-1081 **[0148]**
- **CAPRA, J.D. ; HOPPER, J.E.** *Immunochemistry,* 1976, vol. 13, 995-999 **[0148]**
- **HOPPER, J.E. ; NOYES, C. ; HEINRIKSON, R. ; KESSEL, J.W.** *J.Immunol.,* 1976, vol. 116, 743-746 **[0148]**
- **CAPRA, J.D. ; KEHOE, J.M.** *Proc.Nat.Acad.Sci.USA,* 1974, vol. 71, 845-848 **[0148]**
- **CHEN, P.P. ; LIU, M.-F. ; GLASS, C.A. ; SINHA, S. ; KIPPS, T.J. ; CARSON, D.A.** *Arthritis & Rheumatism,* 1989, vol. 32, 72-76 **[0148] [0151]**
- **KIPPS, T.J. ; TOMHAVE, E. ; PRATT, L.F. ; DUFFY, S. ; CHEN, P.P. ; CARSON. D.A.** *Proc.Natl.Acad.Sci.USA,* 1989, vol. 86, 5913-5917 **[0148]**
- **CHIU, Y.Y.H. ; LOPEZ DE CASTRO, J.A. ; POLJAK, R.J.** *Biochemistry,* 1979, vol. 18, 553-560 **[0148]**
- **CLEARY, M.L. ; MEEKER, T.C. ; LEVY, S. ; LEE, E. ; TRELA, M. ; SKLAR, J. ; LEVY, R.** *Cell,* 1986, vol. 44, 97-106 **[0148]**
- **CUISINIER, A.-M. ; GAUTHIER, L. ; BOUBLI, L. ; FOUGEREAU, M. ; TONNELLE, C.** *Eur.J.Immunol.,* 1993, vol. 23, 110-118 **[0148]**
- **CUNNINGHAM. B.A. ; GOTTLIEB.P.D. ; PFLUMM, M.N. ; EDELMAN, G.M.** Progress In Immunology. Academic Press, 1971, 3-24 **[0148]**

- **CUNNINGHAM, B.A. ; RUTISHAUSER, U. ; GALL, W.E. ; GOTTLIEB, P.D. ; WAXDAL, M.J. ; EDELMAN, G.M.** *Biochemistry,* 1970, vol. 9, 3161-3170 **[0148]**
- **DEANE, M. ; NORTON, J.D.** *Eur.J.Immunol.,* 1990, vol. 20, 2209-2217 **[0148]**
- **DERSIMONIAN, H. ; SCHWARTZ, R.S. ; BARRETT, K.J. ; STOLLAR, B.D.** *J.Immunol.,* 1987, vol. 139, 2496-2501 **[0148] [0148]**
- **CHEN, P.P. ; LIU, M.-F. ; SINHA, S. ; CARSON, D.A.** *Arth.Rheum.,* 1988, vol. 31, 1429-1431 **[0148] [0148] [0151]**
- **DESAI, R. ; SPATZ, L. ; MATSUDA, T. ; ILYAS, A.A. ; BERMAN, J.E. ; ALT, F.W. ; KABAT, E.A. ; LATOV, N.** *J.Neuroimmunol.,* 1990, vol. 26, 35-41 **[0148]**
- **EZAKI, I. ; KANDA, H. ; SAKAI, K. ; FUKUI, N. ; SHINGU, M. ; NOBUNAGA, M. ; WATANABE, T.** *Arthritis And Rheumatism.,* 1991, vol. 34, 343-350 **[0148]**
- **FLORENT, G. ; LEHMAN, D. ; PUTNAM, F.W.** *Biochemistry,* 1974, vol. 13, 2482-2498 **[0148] [0148]**
- **FRIEDLANDER, R.M. ; NUSSENZWEIG, M.C. ; LEDER, P.** *Nucl.Acids Res.,* 1990, vol. 18, 4278 **[0148]**
- **GORMAN, S.D. ; CLARK, M.R. ; ROUTLEDGE, E.G. ; COBBOLD, S.P. ; WALDMANN, H.** *Proc.Natl.Acad.Sci.,* 1991, vol. 88, 4181-4185 **[0148]**
- **GRILLOT-COURVALIN, C. ; BROUET, J.-C. ; PILLER, F. ; RASSENTI, L.Z. ; LABAUME, S. ; SILVERMAN, G.J. ; SILBERSTEIN, L. ; KIPPS, T.J.** *Eur.J.Immunol.,* 1992, vol. 22, 1781-1788 **[0148]**
- **GUILLAUME, T. ; RUBINSTEIN, D.B. ; YOUNG, F. ; TUCKER, L. ; LOGTENBERG, T. ; SCHWARTZ, R.S. ; BARRETT, K.L.** *J.Immunol.,* 1990, vol. 145, 1934-1945 **[0148]**
- **HILLSON, J.L. ; OPPLIGER, I.R. ; SASSO, E.H. ; MILNER, E.C.B. ; WENER, M.H.** *J.Immunol.,* 1992, vol. 149, 3741-3752 **[0148]**
- **HOCH, S. ; SCHWABER, J.** *J.Immunol.,* 1987, vol. 139, 1689-1693 **[0148]**
- **HUANG, C. ; STEWART, A.K. ; SCHWARTZ, R.S. ; STOLLAR, B.D.** *J.Clin.Invest.,* 1992, vol. 89, 1331-1343 **[0148]**
- **IKEMATSU, H. ; HARINDRANATH, N. ; UEKI, Y. ; NOTKINS, A.L. ; CASALI, P.** *J.Immunol.,* 1993, vol. 150, 1325-1337 **[0148]**
- **IKEMATSU, H. ; KASAIAN, M.T. ; SCHETTINO, E.W. ; CASALI, P.** *J.Immunol.,* 1993, vol. 151, 3604-3616 **[0148]**
- **KELLY, P.J. ; PASCUAL, V. ; CAPRA, J.D. ; LIPSKY, P.E.** *J.Immunol.,* 1992, vol. 148, 1294-1301 **[0148]**
- **KIPPS, T.J. ; DUFFY, S.F.** *J.Clin.Invest.,* 1991, vol. 87, 2087-2096 **[0148]**
- **KIPPS, T.J. ; TOMHAVE, E. ; PRATT, L.F. ; DUFFY, S. ; CHEN, P.P. ; CARSON, D.A.** *Proc.Natl.Acad.Sci.USA,* 1989, vol. 86, 5913-5917 **[0148]**
- **KISHIMOTO, T. ; OKAJIMA, H. ; OKUMOTO, T. ; TANIGUCHI, M.** *Nucl.Acids Res.,* 1989, vol. 17, 4385 **[0148]**
- **KNIGHT, G.B. ; AGNELLO, V. ; BONAGURA, V. ; BARNES, J.L. ; PANKA, D.J. ; ZHANG, Q.-X.** *J.Exp.Med.,* 1993, vol. 178, 1903-1911 **[0148]**
- **KOHLER, H. ; SHIMIZU, A. ; PAUL, C. ; MOORE, V. ; PUTNAM, F.W.** *Nature,* 1970, vol. 227, 1318-1320 **[0148]**
- **KOMORI, S. ; YAMASAKI, N. ; SHIGETA, M. ; ISOJIMA, S. ; WATANABE, T.** *Clin.Exp.Immunol.,* 1988, vol. 71, 508-516 **[0148]**
- **KON. S. ; LEVY, S. ; LEVY, R.** *Proc.Natl.Acad.Sci.USA,* 1987, vol. 84, 5053-5057 **[0148]**
- **KRATZIN, H. ; ALTEVOGT, P. ; RUBAN, E. ; KORTT, A. ; STAROSCIK, K. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1975, vol. 356, 1337-1342 **[0148]**
- **KRATZIN, H. ; ALTEVOGT, P. ; KORTT, A. ; RUBAN, E. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1978, vol. 359, 1717-1745 **[0148]**
- **KUDO, A. ; ISHIHARA, T. ; NISHIMURA, Y. ; WATANABE, T.** *Gene,* 1985, vol. 33, 181-189 **[0148]**
- **LEHMAN, D.W. ; PUTNAM, F.W.** *Proc.Nat.Acad.Sci.USA,* 1980, vol. 77, 3239-3243 **[0148]**
- **LIU, V.Y.S. ; LOW, T.L.K. ; INFANTE, A. ; PUTNAM, F.W.** *Science,* 1976, vol. 193, 1017-1020 **[0148]**
- **LOGTENBERG, T. ; YOUNG, F.M. ; VAN ES, J. ; GMELIG-MEYLING, F.H.J. ; BERMAN, J.E. ; ALT, F.W.** *J. Autoimmunity,* 1989, vol. 2, 203-213 **[0148]**
- **LOGTENBERG, T. ; YOUNG, F.M. ; VAN ES, J.H. ; GNIELIG-MEYLING, F.H.J. ; ALT, F.W.** *J.Exp.Med.,* 1989, vol. 170, 1347-1355 **[0148]**
- **MANFTEIMER-LORY, A. ; KATZ. J.B. ; PILLINGER, M. ; GHOSSEIN, C. ; SMITH, A. ; DIAMOND, B.** *J.Exp.Med.,* 1991, vol. 174, 1639-1652 **[0148]**
- **MEEKER, T.C. ; GRIMALDI, J. ; O'ROURKE, R. ; LOEB, J. ; JULIUSSON, G. ; EINHORN, S.** *J.Immol.,* 1988, vol. 141, 3994-3998 **[0148]**
- **MILILI, M. ; FOUGEREAU, M. ; GUGLIELMI, P. ; SCHIFF, C.** *Mol.Immunol.,* 1991, vol. 28, 753-761 **[0148]**
- **MORTARI, F. ; WANG, J.-Y. ; SCHROEDER, JR., H.W.** *J.Inununol.,* 1993, vol. 150, 1348-1357 **[0148]**
- **NICKERSON, K.G. ; BERMAN, J. ; GLICKMAN, E. ; CHESS, L. ; ALT, F.W.** *J.Exp.Med.,* 1989, vol. 169, 1391-1403 **[0148]**
- **PASCUAL, V. ; RANDEN, I. ; THOMPSON, K. ; SIOUD, M. ; FORRE, O. ; NATVIG, J. ; CAPRA, J.D.** *J.Clin.Invest.,* 1990, vol. 86, 1320-1328 **[0148] [0148]**

- **PASCUAL, V. ; VICTOR, K. ; RANDEN, L ; THOMPSON, K. ; STEINITZ, M. ; FORRE, O. ; FU, S.-M. ; NATVIG, J.B. ; CAPRA, J.D.** *Scand.J.Immunol.,* 1992, vol. 36, 349-362 **[0148]**
- **PASCUAL, V. ; VICTOR, K. ; SPELLERBERG, M. ; HAMBLIN, T.J. ; STEVENSON, F.K. ; CAPRA, J.D.** *J.Immunol.,* 1992, vol. 149, 2337-2344 **[0148]**
- **PONSTINGL, H. ; SCHWARZ, J. ; REICHEL, W. ; HILSCHMANN, N.** *Z. Physiol. Chem.,* 1970, vol. 351, 1591-1594 **[0148]**
- **PONSTINGL, H. ; HILSCHMANN, N. ; 1571-1604.** *Z.Physiol.Chem.,* 1976, vol. 357 **[0148]**
- **PORTOLANO, S. ; MCLACHLAN, S.M. ; RAPOPORT, B.** *J.Iminunol.,* 1993, vol. 151, 2839-2851 **[0148]**
- **PRATT, L.F. ; SZUBIN, R. ; CARSON, D.A. ; KIPPS, T.J.** *J.Iminunol.,* 1991, vol. 147, 2041-2046 **[0148]**
- **PRESS, E.M. ; HOGG, N.M.** *Biochem.J.,* 1970, vol. 117, 641-660 **[0148]**
- **PUTNAM, F.W. ; SHIMIZU, A. ; PAUL., C. ; SHINODA, T. ; KOHLER, H.** *Ann.N.Y.Acad.Sci.,* 1971, vol. 190, 83-103 **[0148]**
- **PUTNAM, F.W. ; TAKAHASHI, N. ; TETAERT, D. ; DEBUIRE, B. ; LIN, L.C.** *Proc.Nat.Acad.Sci.USA,* 1981, vol. 78, 6168-6172 **[0148]**
- **TAKAHASHI, N. ; TETAERT, D. ; DEBUIRE, B. ; LIN, L. ; PUTNAM, F.W.** *Proc.Nat.Acad.Sci.USA,* 1982, vol. 79, 2850-2854 **[0148]**
- **RAAPHORST, F.M. ; TIMMERS, E. ; KENTER, M.J.H. ; VAN TOL, M.J.D. ; VOSSEN, J.M. ; SCHUURMAN, R.K.B.** *Eur.J.Immunol.,* 1992, vol. 22, 247-251 **[0148]**
- **RABBITTS. T.H. ; BENTLEY, D.L. ; DUNNICK, W. ; FORSTER, A ; MATTHYSSENS, G. ; MILSTEIN, C.** Symp.Quanti.Biol. Cold Spring Harb, 1980, vol. 45, 867-878 **[0148]**
- **MATTHYSSENS, G. ; RABBITTS, T.H.** *Proc.Nat.Acad.Sci.USA,* 1980, vol. 77, 6561-6565 **[0148]**
- **RANDEN, I. ; PASCUAL, V. ; VICTOR, K. ; THOMPSON, K.M. ; FORRE, O. ; CAPRA, J.D. ; NATVIG, J.B.** *Eur.J.Immunol.,* 1993, vol. 23, 1220-1225 **[0148]**
- **RASSENTI, L.Z. ; KIPPS, T.J.** *J.Exp.Med.,* 1993, vol. 177, 1039-1046 **[0148]**
- **REIDL, L.S. ; FRIEDMAN, D.F. ; GOLDMAN, J. ; HARDY, R.R. ; JEFFERIES, L.C. ; SILBERSTEIN, L.E.** *J.Immunol.,* 1991, vol. 147, 3623-3631 **[0148]**
- **ROUDIER, J. ; SILVERMAN, G.J. ; CHEN, P.P. ; CARSON, D.A. ; KIPPS, T.J.** *J.Immunol.,* 1990, vol. 144, 1526-1530 **[0148]**
- **SANZ, I. ; CASALI, P. ; THOMAS, J.W. ; NOTKINS, A.L. ; CAPRA, J.D.** *J.Immunol.,* 1989, vol. 142, 4054-4061 **[0148]**
- **SANZ, I. ; DANG, H. ; TAKEI, M. ; TALAL, N. ; CAPRA, J.D.** *J.Immunol.,* 1989, vol. 142, 883-887 **[0148]**
- **SCHMIDT, W.E. ; JUNG, H-.D. ; PALM, W. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1983, vol. 364, 713-747 **[0148]**
- **SCHROEDER. H.W., JR. ; WANG, J.Y.** *Proc.Natl.Acad.Sci.USA,* 1990, vol. 87, 6146-6150 **[0148]**
- **SCHROEDER, H.W., JR. ; HILLSON, J.L. ; PERLMUTTER, R.M.** *Science,* 1987, vol. 238, 791-793 **[0148] [0148] [0148]**
- **CHEN. P.P. ; LIU, M.-F. ; GLASS. C.A. ; SINHA, S. ; KIPPS, T.J. ; CARSON, D.A.** *Arthritis & Rheumatism,* 1989, vol. 32, 72-76 **[0148]**
- **SCHUTTE, M.E. ; EBELING, S.B. ; AKKERMANS, K.E. ; GMELIG-MEYLING, F.H. ; LOGTENBERG, T.** *Eur.J.Immunol.,* 1991, vol. 21, 1115-1121 **[0148]**
- **SCHUTTE, M.E. ; EBELING, S.B. ; AKKERMANS, K.E. ; GMELIG-MEYLING, F.H.J. ; LOGTENBERG, T.** *Eur.J.Immunol.,* 1991, vol. 21, 1115-1121 **[0148]**
- **SETTMACHER, U. ; JALIN, S. ; SIEGEL, P. ; VON BAEHR, R. ; HANSEN, A.** *Mol.Immunol.,* 1993, vol. 30, 953-954 **[0148]**
- **SHEN, A. ; HUMPHRIES, C. ; TUCKER, P. ; BLATTNER, F.** *Proc.Natl.Acad.Sci.USA,* 1987, vol. 84, 8563-8567 **[0148]**
- **SHIMIZU, A. ; NUSSENZWEIG, M.C. ; MIZUTA, T.-R. ; LEDER, P. ; HONJO, T.** *Proc. Natl.Acad. Sci. USA,* 1989, vol. 86, 8020-8023 **[0148]**
- **SHIN, E.K. ; MATSUDA, F. ; FUJIKURA, J. ; AKAMIZU, T. ; SUGAWA, H. ; MORI, T. ; HONJO, T.** *Eur.J.Immunol.,* 1993, vol. 23, 2365-2367 **[0148]**
- **SILBERSTEIN, L.E. ; LITWIN, S. ; CARMACK, C.E.** *J.Exp.Med.,* 1989, vol. 169, 1631-1643 **[0148]**
- **SINGAL, D.P. ; FRAME. B. ; JOSEPH. S. ; BLAJCHMAN, M.A. ; LEBER, B.F.** *Immunogenet.,* 1993, vol. 38, 242 **[0148]**
- **STEINER, L.A. ; GARCIA-PARDO, A. ; MARGOLIES, M.N.** *Biochemistry,* 1979, vol. 18, 4068-4080 **[0148]**
- **STEWART, A.K. ; HUANG, C. ; STOLLAR, B.D. ; SCHWARTZ, R.S.** *J.Exp.Med.,* 1993, vol. 177, 409-418 **[0148]**
- **THOMAS, J.W.** *J.Immunol.,* 1993, vol. 150, 1375-1382 **[0148]**
- **TORANO, A. ; PUTNAM, F.W.** *Proc.Nat.Acad.Sci.USA,* 1978, vol. 75, 966-969 **[0148]**
- **VAN DER HEIJDEN, R.W.J. ; BUNSCHOTEN, H. ; PASCUAL, V. ; UYTDEHAAG, F.G.C.M. ; OSTERHAUS, A.D.M.E. ; CAPRA, J.D.** *J.Immunol.,* 1990, vol. 144, 2835-2839 **[0148]**
- **VAN DER STOEP, N. ; VAN DER LINDEN, J. ; LOGTENBERG, T.** *J.Exp.Med.,* 1993, vol. 177, 99-107 **[0148]**
- **VAN ES, J.H. ; GMELIG-MEYLING, F.H.J. ; LOGTENBERG, T.** *Eur.J.Immunol.,* 1992, vol. 22, 2761-2764 **[0148]**

- **VARADE, W.S. ; MARIN, E. ; KITTELBERGER, A.M. ; INSEL, R.A.** *J.Immunol.,* 1993, vol. 150, 4985-4995 **[0148]**
- **WATANABE, S. ; BARNIKOL, H.U. ; HORN, J. ; BERTRAM, J. ; HILSCHMANN, N.** *Z.Physiol.Chem.,* 1973, vol. 354, 1505-1509 **[0148]**
- **WENG, N.-P. ; YU-LEE, L.-Y. ; SANZ, L ; PATTEN, B.M. ; MARCUS, D.M.** *J.Intmunol.,* 1992, vol. 149, 2518-2529 **[0148]**
- **WHITE, M.B. ; WORD, C.J. ; HUMPHRIES, C.G. ; BLATTNER, F.R. ; TUCKER, P.W.** *Mol.Cell.Biol.,* 1990, vol. 10, 3690-3699 **[0148]**
- **YAGO, K. ; ZENITA, K. ; OHWAKI, I. ; HARADA, Y. ; NOZAWA, S. ; TSUKAZAKI, K. ; IWAMORI, M. ; ENDO, N. ; YASUDA, N. ; OKUMA, M.** *Mol.Immunol.,* 1993, vol. 30, 1481-1489 **[0148]**
- **ZELENETZ, A.D. ; CHEN, T.T. ; LEVY, R.** *J.Exp.Med.,* 1992, vol. 176, 1137-1148 **[0148]**
- **COX, J.P.L. ; TOMLINSON, I.M. ; WINTER, G.** *Eur.J.Immunol.,* 1994, vol. 24, 827-836 **[0149]**
- **HUBER, C. et al.** *Eur.J.Immunol.,* 1993, vol. 23, 2868 **[0149]**
- **KLOBECK, H.G. ; BORNKAMMM, G.W. ; COMBRIATO, G. ; MOCIKAT, R. ; POHLENZ, H.D. ; ZACHAU, H.G.** *Nucl.Acids Res.,* 1985, vol. 13, 6515-6529 **[0149]**
- **LAUTNER-RIESKE. A. ; HUBER, C. ; MEINDL, A. ; PARGENT, W. ; SCHÄBLE, K.F. ; THIEBE, R. ; ZOCHER, I. ; ZACHAU, H.G.** *Eur.J.Immunol.,* 1992, vol. 22, 1023 **[0149]**
- **LORENZ, W. ; SCHABLE, K.F. ; THIEBE, R. ; STAVNEZER, J. ; ZACHAU, H.G.** *Mol.Immunol.,* 1988, vol. 25, 479 **[0149]**
- **PARGENT, W. ; MEINDL, A. ; THIEBE, R. ; MITZEL, S. ; ZACHAU, H.G.** *Eur.J.Immunol.,* 1991, vol. 21, 1821-1827 **[0149]**
- **SCOTT, M.G. ; CRIMMINS, D.L. ; MCCOURT, D.W. ; CHUNG, G. ; SCHABLE, K.F. ; THIEBE, R. ; QUENZEL, E.-M. ; ZACHAU, H.G. ; NAHM, M.H.** *J.Immunol.,* 1991, vol. 147, 4007-4013 **[0149]**
- **STRAUBINGER, B. ; HUBER, E. ; LORENZ, W. ; OSTERHOLZER, E. ; PARGENT, W. ; PECH, M. ; POHFENZ, H.-D. ; ZIMMER, F.-J. ; ZACHAU, H.G.** *J.Mol.Biol.,* 1988, vol. 199, 23-34 **[0149]**
- **STRAUBINGER, B. ; THIEBE, R. ; HUBER, C. ; OSTERHOLZER, E. ; ZACHAU, H.G.** *Biol. Chem. Hoppe-Seyer,* 1988, vol. 369, 601-607 **[0149]**
- **WILLIAMS, S.C. ; WINTER, G.** *Eur.J.Immunol.,* 1993, vol. 23, 1456-1461 **[0150]**
- **SIMINOVITCH, K.A. ; MISENER, V. ; KWONG, P.C. ; SONG, Q.-L. ; CHEN, P.P.** *J.Clin.Invest.,* 1989, vol. 84, 1675-1678 **[0150]**
- **BROCKLY, F. ; ALEXANDRE, D. ; CHUCHANA, P. ; HUCK, S. ; LEFRANC, G. ; LEFRANC, M.-P.** *Nuc.Acids.Res.,* 1989, vol. 17, 3976 **[0150]**
- **DALEY, M.D. ; PENG, H.-Q. ; MISENER, V. ; LIU, X.-Y. ; CHEN, P.P. ; SIMINOVITCH, K.A.** *Mol.Immunol.,* 1992, vol. 29, 1515-1518 **[0150]**
- **DEFTOS, M. ; SOTO-GIL, R. ; QUAN, M. ; OLEE, T. ; CHEN, P.P.** *Scand. J. Immunol.,* 1994, vol. 39, 95 **[0150]**
- **STIERNHOLM, N.B.J. ; KUZNIAR, B. ; BERINSTEIN, N.L.** *J. Immunol.,* 1994, vol. 152, 4969-4975 **[0150]**
- **ANDERSON, M.L.M. ; SZAJNERT, M.F. ; KAPLAN, J.C. ; MCCOLL, L. ; YOUNG, B.D.** *Nuc.Acids Res.,* 1984, vol. 12, 6647-6661 **[0150]**
- **ANDRIS. J.S. ; BRODEUR, B.R. ; CAPRA, J.D.** *Mol.Immunol.,* 1993, vol. 30, 1601-1616 **[0151]**
- **BERMAN, J.E. ; MELLIS, S.J. ; POLLOCK, R. ; SMITH, C.L. ; SUH, H. ; HEINKE, B. ; KOWAL, C. ; SURTI, U. ; CHESS, L. ; CANTOR, C.R.** *Embo J.,* 1988, vol. 7, 727-738 **[0151]**
- **BULUWELA, L. ; RABBITTS, T.H.** *Eur.J.Immunol.,* 1988, vol. 18, 1843-1845 **[0151]**
- **BULUWELA, L. ; ALBERTSON, D.G. ; SHERRITTGTON, P. ; RABBITTS, P.H. ; SPURR, N. ; RABBITTS, T.H.** *Embo J.,* 1988, vol. 7, 2003-2010 **[0151]**
- **COOK, G.P. et al.** *Nature Genetics,* 1994, vol. 7, 162-168 **[0151]**
- **HAINO, M. et al.** *J. Biol. Chem.,* 1994, vol. 269, 2619-2626 **[0151]**
- **HUMPHRIES, C.G. ; SHEN, A. ; KUZIEL, W.A. ; CAPRA, J.D. ; BLATTNER, F.R. ; TUCKER, P.W.** *Nature,* 1988, vol. 331, 446-449 **[0151]**
- **KODAIRA, M. ; KINASHI, T. ; UMEMURA, I. ; MATSUDA, F. ; NOMA, T. ; ONE. Y. ; HONJO, T.** *J.Mol.Biol.,* 1986, vol. 190, 529-541 **[0151]**
- **LEE, K.H. ; MATSUDA, F. ; KINASHI, T. ; KODAIRA, M. ; HONJO, T.** *J.Mol.Biol.,* 1987, vol. 195, 761-768 **[0151]**
- **MATSUDA, F. ; LEE, K.H. ; NAKAI, S. ; SATO, T. ; KODAIRA, M. ; ZONG, S.Q. ; OHNO, H. ; FUKUHARA, S. ; HONJO, T.** *Embo J.,* 1988, vol. 7, 1047-1051 **[0151]**
- **MATSUDA, F. ; SHIN, E.K. ; HIRABAYASHI, Y. ; NAGAOKA, H. ; YOSHIDA, M.C. ; ZONG, S.Q. ; HONJO, T.** *Embo J.,* 1990, vol. 9, 2501-2506 **[0151]**
- **MATSUDA, F. ; SHIN, E.K. ; NAGAOKA, H. ; MATSUMURA, R. ; HAINO, M. ; FUKITA, Y. ; TAKA-ISHI, S. ; IMAI, T. ; RILEY, J.H. ; ANAND, R.** *Nature Genet.,* 1993, vol. 3, 88-94 **[0151]**
- **NAGAOKA, H. ; OZAWA, K. ; MATSUDA, F. ; HAYASHIDA, H. ; MATSUMURA, R. ; HAINO, M. ; SHIN, E.K. ; FUKITA, Y. ; IMAI, T. ; ANAND, R.** *Temporal,* 1993 **[0151]**
- **RECHAVI, G. ; BIENZ, B. ; RAM, D. ; BEN-NERIAH, Y. ; COHEN, J.B. ; ZAKUT, R. ; GIVOL, D.** *Proc.Nat.Acad.Sci.USA,* 1982, vol. 79, 4405-4409 **[0151]**
- **SANZ, I. ; KELLY, P. ; WILLIAMS, C. ; SCHOLL, S. ; TUCKER, P. ; CAPRA, J.D.** *Embo J.,* 1989, vol. 8, 3741-3748 **[0151]**
- **SHIN. E.K. ; MATSUDA. F. ; FUJIKURA, J. ; AKAMIZU, T. ; SUGAWA, H. ; MORI, T. ; HONJO, T.** *Eur.J.Immunol.,* 1993, vol. 23, 2365-2367 **[0151]**

- **TOMLINSON, IM. ; WALTER, G. ; MARKS. JD. ; LLEWELYN, MB. ; WINTER. G.** *J.MoLBiol.,* 1992, vol. 227, 776-798 **[0151]**
- **VAN DER MAAREL, S. ; VAN DIJK, K.W. ; ALEXANDER, C.M. ; SASSO, E.H. ; BULL, A. ; MILNER, E.C.B.** *J.Immunol.,* 1993, vol. 150, 2858-2868 **[0151]**
- **VAN DIJK, K.W. ; MORTARI, F. ; KIRKHAM, P.M. ; SCHROEDER, JR., H.W. ; MILNER, E.C.B.** *Eur.J.Immunol.,* 1993, vol. 23, 832-839 **[0151]**
- **VAN ES, J.H. ; AANSTOOT, H. ; GMELIG-MEYLING, F.H.J. ; DERKSEN, R.H.W.M. ; LOGTENBERG, T.** *J.Immunol.,* 1992, vol. 149, 223 4-2240 **[0151]**
- **WENG, N.-P. ; SNYDER, J.G. ; YU-LEE, L.-Y. ; MARCUS, D.M.** *Eur.J.Immunot.,* 1992, vol. 22, 1075-1082 **[0151]**
- **OLEE, T. ; YANG, P.M. ; SIMINOVITCH, K.A. ; OLSEN, N.J. ; HILLSON, J.L. ; WU, J. ; KOZIN, F. ; CARSON, D.A. ; CHEN, P.P.** *J. Clin. Invest.,* 1991, vol. 88, 193-203 **[0151]**
- **CHEN, P.P. ; YANG, P.M.** *Scand. J. Immunol.,* 1990, vol. 31, 593-599 **[0151]**