(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
2 August 2001 (02.08.2001)

PCT

(10) International Publication Number
WO 01/55182 A1

(51) International Patent Classification[7]: C07K 14/22, C12N 15/31

(21) International Application Number: PCT/AU01/00069

(22) International Filing Date: 25 January 2001 (25.01.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/177,917 25 January 2000 (25.01.2000) US

(71) Applicant: THE UNIVERSITY OF QUEENSLAND [AU/AU]; St Lucia, Brisbane, Queensland 4072 (AU).

(72) Inventors: PEAK, Ian, Richard, Anselm; Unit 10, 81 Armadale Street, St Lucia, Brisbane, Queensland 4067 (AU). JENNINGS, Michael, Paul; 20 Picasso Street, Carina, Brisbane, Queensland 4152 (AU).

(74) Agent: FISHER, Adams, Kelly; Level 13, Amp Place, 10 Eacle Street, Brisbane, Queensland 4000 (AU).

(81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

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

WO 01/55182 A1

(54) Title: PROTEINS COMPRISING CONSERVED REGIONS OF NEISSERIA MENINGITIDIS SURFACE ANTIGEN NhhA

(57) Abstract: Novel proteins that constitute modified forms of a *Neisseria meningitidis* surface antigen and encoding nucleic acids are provided. The modified surface proteins are characterized by having deletions of non-conserved amino acids, and thereby being capable of eliciting cross-protective immune responses against *Neisseria meningitidis*. The invention extends to the use of the modified surface antigens in diagnostics, in therapeutic and prophylactic vaccines and in the design and/or screening of medicaments. The modified surface antigens are particularly usefulin vaccines which effectively immunize against a broader spectrum of *N. meningitidis* strains than would be expected from a corresponding wild-type surface antigen.

Proteins comprising conserved regions of Neisseria meningitidis surface antigen NhhA

FIELD OF THE INVENTION

THIS INVENTION relates to novel proteins that constitute modified forms of a *Neisseria meningitidis* surface antigen, to nucleic acids encoding such novel peptides and polypeptides, and to the use of these in diagnostics, in therapeutic and prophylactic vaccines and in the design and/or screening of medicaments. More particularly, by having deletions of non-conserved amino acids, the modified surface antigens of the invention may be useful in vaccines which effectively immunize against a broader spectrum of *N. meningitidis* strains than would be expected from a corresponding wild-type surface antigen.

BACKGROUND OF THE INVENTION

Neisseria meningitidis is a Gram-negative bacterium and the causative agent of meningococcal meningitis and septicemia. Its only known host is the human, and it may be carried asymptomatically by approximately 10% of the population (Caugant *et al*, 1994, Journal of Clinical Microbiology **32** 323).

N. meningitidis may express a polysaccharide capsule, and this allows classification of the bacteria according to the nature of the capsule expressed. There are at least twelve serogroups of *N. meningitidis*: A, B, C, 29-E, H, I, K, L, W135, X, Y and Z, of which serogroups A, B, and C cause 90% of meningococcal disease (Poolman *et al*, 1995, Infectious Agents and Disease **4** 13). Vaccines directed against serogroups A and C are available, but the serogroup B capsular polysaccharide is poorly immunogenic and does not induce protection in humans.

Other membrane and extracellular components are therefore being examined for their suitability for inclusion in vaccines. Examples include the outer membrane proteins of classes 1, 2 and 3 (porin; encoded by *por* genes), and classes 4 (Rmp) and 5 (Opacity proteins; encoded by *opa* and *opc* genes).

However, to date, none of these candidates is able to induce complete protection, particularly in children (Romero *et al.*, 1994, Clinical Microbiology Review, 7 559; Poolman *et al.*, 1995, *supra*).

To create an effective vaccine, it is necessary to identify components of *N. meningitidis* which are present in a majority of strains, and which are capable of inducing a protective immune response (for example, bactericidal antibodies).

In this regard, reference is made to International Publications WO 99/24578, WO99/36544, WO99/58683 and WO99/57280, each of which is incorporated herein by reference and describe a number of candidate proteins that may be useful in vaccines to immunize against *Neisseria meningitidis*.

In this regard, particular reference is made to International Publication WO99/31132 and Peak *et al.* 2000, FEMS Immunol. Med. Microbiol. **28** 329, each of which is incorporated herein by reference and describe a novel surface antigen isolated from a number of different strains of *N. meningitidis,* which surface antigen, and allelic variants thereof, for the purposes of this specification will be referred to as NhhA.

SUMMARY OF THE INVENTION

The present inventors have discovered that the NhhA surface antigen has polypeptide regions which are variable between *N. meningitidis* strains, and other regions which are conserved between strains. The variable regions may be immunogenic and tend to elicit strain-specific immune responses, such that vaccines incorporating an NhhA antigen derived from a particular strain of *N. meningitidis* tend to preferentially immunize against that particular strain. As a result, the present inventors have sought to produce a modified NhhA polypeptide which elicts an immune response which is not as strain-specific as that elicited by wild-type NhhA. This modified NhhA antigen will be useful for the production of therapeutic and/or prophylactic vaccines against *N. meningitidis* as will be described hereinafter. By directing the immune response primarily against conserved epitopes, such vaccines should effectively immunize against a broader spectrum of *N. meningitidis* strains than would be expected following immunization with wild-type NhhA.

The present invention is therefore broadly directed to isolated proteins having conserved amino acids of NhhA polypeptides.

Proteins of the invention may therefore have one or more deletions of non-conserved amino acids compared to a corresponding wild-type NhhA polypeptide.

In a first aspect, the invention provides an isolated protein comprising twelve or more contiguous conserved amino acids sequences of an NhhA polypeptide, said isolated protein excluding wild-type NhhA polypeptides.

Suitably, the protein of the invention is capable of eliciting an immune response.

Preferably, the immune response is less strain-specific than that elicited by said corresponding wild-type NhhA polypeptide.

More preferably, said immune response provides protection against one or more strains of *N. meningitidis*, or even more preferably a plurality of strains of *N. meningitidis*

Wild-type NhhA polypeptide sequences are exemplified in FIG.1 (SEQ ID NOS: 1 to 10).

A consensus amino acid sequence is also set forth in FIG.1 (SEQ ID NO:11).

The isolated protein of the invention preferably comprises one or more constant regions of an NhhA polypeptide, herein designated C1, C2, C3, C4 and C5 regions in FIG. 1.

It will be appreciated that according to this aspect, suitably one or more non-conserved amino acids of a variable region of an NhhA polypeptide, designated as V1, V2, V3 or V4 regions in FIG. 1, are deleted with respect to a wild-type NhhA polypeptide.

Preferably, a V1 region, or at least a substantial portion thereof, is deleted.

In particular embodiments, the isolated protein has an amino acid sequence as set forth in any one of FIGS. 5 to 9 (SEQ ID NOS: 23 to 27) which are examples of "*modified NhhA polypeptides of the invention*". In FIG. 14 (SEQ ID NOS: 33 to 39) further examples are provided of "mature" polypeptides predicted to result of removal of N-terminal signal sequences.

According to a second aspect, the invention provides an isolated nucleic acid encoding a polypeptide according to the first aspect.

Wild-type *nhhA* nucleic acid sequences are exemplified in FIG.2 (SEQ ID NOS: 12 to 21).

A consensus nucleic acid sequence is also set forth in FIG.2 (SEQ ID NO:22).

Preferably, the C1, C2, C3, C4 and C5 regions are encoded by respective nucleotide sequences as set forth in FIG. 2.

Preferably, the V1, V2, V3 and V4 regions are encoded by respective nucleotide sequences as set forth in FIG. 2.

In a particular embodiment, the isolated nucleic acid of the invention has a nucleotide sequence as set forth in any one of FIGS. 5 to 9 (SEQ ID NOS: 28 to 32), which are particular examples of "*modified nhhA nucleic acids of the invention*".

The invention according to the first and second aspects extends to homologs, fragments, variants and derivatives of the isolated proteins and nucleic acids of the invention.

Specifically excluded from the scope of the invention are wild-type NhhA polypeptides and *nhhA* nucleic acids.

In a third aspect, the invention resides in an expression construct comprising an expression vector and a nucleic acid according to the second aspect, wherein said sequence is operably linked to one or more regulatory nucleic acids in said expression vector.

In a fourth aspect, the invention provides a host cell containing an expression construct according to the third aspect.

In a fifth aspect of the invention, there is provided a method of producing a recombinant isolated protein according to the first aspect, said method comprising the steps of:

(i) culturing a host cell containing an expression vector according to the third aspect such that said polypeptide is expressed in said host cell; and

(ii) isolating said recombinant polypeptide.

In a sixth aspect, the invention provides an antibody or antibody fragment that binds to a protein of the invention, fragment, variant or derivative thereof.

In a seventh aspect, the invention provides a method of detecting *N. meningitidis* in a biological sample suspected of containing same, said method comprising the steps of:-

(i) isolating the biological sample from an individual;

(ii) combining the above-mentioned antibody or antibody fragment with the biological sample; and

(iii) detecting specifically bound antibody or antibody fragment which indicates the presence of *N. meningitidis*.

In an eighth aspect, there is provided a method of detecting *N. meningitidis* bacteria in a biological sample suspected of containing said bacteria, said method comprising the steps of:-

(i) isolating the biological sample from a patient;

(ii) detecting a nucleic acid sequence according to the second-mentioned aspect in said sample which indicates the presence of said bacteria.

In a ninth aspect, the invention provides a method for diagnosing infection of an individual by *N. meningitidis*, said method comprising the steps of:-

(i) contacting a biological sample from an individual with a polypeptide, fragment, variant or derivative of the invention; and

(ii) determining the presence or absence of a complex between said polypeptide, fragment, variant or derivative and *N. meningitidis*-specific antibodies in said sample, wherein the presence of said complex is indicative of said infection.

Preferably, the individual is a mammal.

More preferably, the individual is a human.

In a tenth aspect, the invention also extends to the use of an isolated protein according to the first-mentioned aspect, the use of isolated nucleic acids according to the second aspect or the use of the antibody or antibody fragment mentioned above in a kit for detecting *N. meningitidis* bacteria in a biological sample.

In an eleventh aspect of the invention, there is provided a pharmaceutical composition comprising an isolated protein according to the first mentioned aspect.

Preferably, said pharmaceutical composition is a vaccine.

In a twelfth aspect, the invention provides a method of preventing infection of a patient by *N. meningitidis*, comprising the step of administrating a pharmaceutically effective amount of the above-mentioned vaccine.

In a thirteenth aspect, the invention provides a method of identifying an immunogenic fragment of an isolated protein, variant or derivative according to the first mentioned aspect, comprising the steps of:-

(i) producing a fragment of said polypeptide, variant or derivative;

(ii) administering said fragment to an individual; and

(iii) detecting an immune response in said individual, which response includes production of elements which specifically bind *N. meningitidis* and/or said polypeptide, variant or derivative, and/or a protective effect against *N. meningitidis* infection.

Preferably, the individual is a mammal.

More preferably, the individual is a human.

<u>BRIEF DESCRIPTION OF THE FIGURES AND TABLES</u>

Table 1: Identification of amino acids of the conserved regions (C1, C2, C3, C4 and C5) and variable regions (V1, V2, V3 and V4) of an NhhA polypeptide from each of ten (10) indicated strains of *N. meningitidis*. Relevant SEQ ID NOS are also indicated. Column 1 = strain designation. SEQ ID NOS:1-9 were previously described in copending application WO99/31132; the sequences of NhhA and *nhhA* of strain Z2491 were obtained from *http://www.sanger.ac.uk/Projects/N_meningitidis/*; column 2 = amino acid

numbering of C1 region; column 3 = amino acid numbering of V1 region; column 4 = amino acid numbering of C2 region; column 5 = amino acid numbering of V2 region; column 6 = amino acid numbering of C3 region, column 7 = amino acid numbering of V2 region; column 8 = amino acid numbering of C4 region; column 9 = amino acid numbering of V4 region; column 10 = amino acid numbering of C5 region. Note that the amino acid numbering of the consensus sequence (SEQ ID NO:11) is also indicated.

Table 2: Table of amino acid substitutions.

FIG. 1: Amino acid sequence alignments of NhhA polypeptide amino acid sequences from ten (10) *N. meningitidis* strains (SEQ ID NOS: 1-10) together with consensus sequence (SEQ ID NO:11). Strain names and polypeptide sequences used in this alignment correspond to the strain names and SEQ ID NOS in column 1 of Table 1. Amino acids are indicated by standard single letter abbreviations. Consensus amino acids are shown only where residues are completely conserved. Conserved regions (double underlined, labeled C1, C2, C3, C4, C5) and variable regions (single underlined, labeled V1, V2, V3, V4) are indicated under the consensus sequence.

FIG. 2: Nucleotide sequence alignment of *nhhA* nucleic acids from ten (10) *N. meningitidis* strains, which sequences encode the amino acid sequences of FIG. 1. Regions C1, C2, C3, C4, C5 and V1, V2, V3, V4 are as described in FIG. 1 and Table 1.

FIG. 3: Plasmid map corresponding to pCO14K with a PCR amplification product encoding wild-type PMC21 NhhA operably linked to the *porA* promoter. (Not drawn to scale) 3A: Solid arrows indicate the arrangement of the *porA* and *kanR* genes in pCO14K.Oligonucleotide primers HOMP5' and HOMP3'AN used to amplify the *nhhA* gene of strain PMC21 are shown. The *nhhA* gene is shown by dotted arrow, the *porA* promoter by a black box, and *Eag*I and *Nco*I restriction sites used to replace *porA* with *nhhA* in as described in Example 2 are shown. 3B Arrangement of genes in pIP52(PMC21), as described in Example2. the *Bgl*II site used to construct a mutant as described in Example 4 is shown.

FIG. 4: Schematic representation of Splice Overlap Extension PCR strategy for deletion of specific regions of NhhA polypeptides. A schematic of the wild-type *nhhA* gene is shown at the top of Figures 4A-C, and the recombinant *nhhA* is shown at the bottom of these figures, with variable regions shown as black and constant regions by unfilled boxes. Arrows indicate approximate location of oligonucleotide primers. Vertical hatched lines indicate amplification products. Where oligonucleotide sequence is from discontinuous regions of an *nhhA* nucleic acid, this is shown by a dotted line between such discontinuous regions. Approximate scale indicated. Double vertical lines indicate that only a portion of the C5 region is shown. A: shows the strategy as described in Example 6. B: shows the strategy as described in Example 7. C: shows the strategy as described in Example 8.

FIG. 5: (A) Amino acid sequence of PMC 21 NhhA deletion mutant polypeptide (SEQ ID NO: 23) produced in Example 4; and (B) encoding nucleotide sequence (SEQ ID NO: 28).

FIG. 6: (A) Amino acid sequence of H41 NhhA deletion mutant polypeptide (SEQ ID NO: 24) produced in Example 5; and (B) encoding nucleotide sequence (SEQ ID NO: 29).

FIG. 7: (A) Amino acid sequence of PMC21 NhhA deletion mutant polypeptide (SEQ ID NO: 25) produced by splice overlap PCR in Example 6; and (B) encoding nucleotide sequence (SEQ ID NO: 30).

FIG. 8: (A) Amino acid sequence of PMC21 NhhA deletion mutant polypeptide (SEQ ID NO: 26) produced by splice overlap PCR in Example 7; and (B) encoding nucleotide sequence (SEQ ID NO: 31).

FIG. 9: (A) Amino acid sequence of PMC21 NhhA deletion mutant polypeptide (SEQ ID NO: 27) produced by splice overlap PCR in Example 8; and (B) encoding nucleotide sequence (SEQ ID NO: 32).

FIG.10: Amino acid sequence alignments of wild type and NhhA deletion mutant polypeptide sequences. These polypeptides were produced as described in Example 2, Example 3, Example 4 and Example 5. Amino acids are indicated by the one letter abbreviation. Conserved regions labelled C1, C2, C3, C4 and C5

corresponding to those defined in Table 1 and FIG. 1 are indicated by double underlining of full length sequences from H41 and PMC21, and variable regions labelled V1, V2, V3, V4 corresponding to those defined in Table 1 and FIG. 1 are indicated by single underlining of full length sequences from H41 and PMC21.

FIG. 11: Western immunoblot showing over expressed NhhA. 45 μg total cell protein was separated on 4-20% gradient SDS-PAGE before transfer to a nitrocellulose filter and western immunoblot as described in Example 9. Lane 1: Parental strain showing wild-type level of NhhA expression. Lane 2: Strain P6 (overexpresses PMC 21 NhhA as described in Example 2). Lane 3: Strain PΔ6 (overexpresses the truncated PMC 21NhhA described in Example 4). Lane 4: Strain H14 (overexpresses H41 NhhA described in Example 3). Lane 5: Strain HΔ8 (overexpresses the truncated H41 NhhA described in Example 5). Lane 6: Strain 2A (NhhA expression abolished by mutation of *nhhA* gene as described in International Publication WO99/31132). Migration of standards is indicated: 185 kDa, 119 kDa, 85 kDa, 62 kDa, 51,2 kDa, 38.2 kDa, 22.4 kDa. Wild-type NhhA polypeptide is present as a high molecular weight immunoreactive band present in lane 1 but absent from lane 6.

FIG. 12: Isolated NhhA deletion mutant polypeptides. NhhA polypeptides were isolated as described in Example 9 before separation on 4-20% SD-PAGE. The polyacrylamide gel was Coomassie stained. Lane 1: OMC preparation of Strain overexpressing the truncated PMC21 NhhA polypeptide described in Example 6. Lane 2: Purified truncated PMC21 NhhA polypeptide. Lane 3: OMC preparation of Strain over-expressing the truncated PMC21 NhhA polypeptide described in Example 4. Lane 4: Purified truncated PMC21 NhhA polypeptide. Lane 5: OMC preparation of a strain overexpressing PMC21 NhhA polypeptide described in Example 2. Lane 6: Purified PMC21 NhhA polypeptide. Lane 7: Molecular weight standards of 173 kDa, 111 kDa, 80 kDa, 61 kDa, 49 kDa, 36 kDa. Note that the reactive high molecular weight species in all lanes except 6 probably represents multimers of NhhA polypeptides. Other bands are probably less stable forms of NhhA or breakdown products. Note these are absent from lane 6.

FIG. 13: Western Immunoblot using anti-NhhA protein mouse sera. In all panels, lanes 1, 3, 5, 7, contain OMC of Strain over expressing PMC21 NhhA polypeptide, and lanes 2, 4, 6, and 8 contain OMC of strain 2A which does not express NhhA. PanelA: Lanes 1 and 2: mouse A inoculated with wild-type PMC21 NhhA at a 1:1000 dilution. Lanes 3 and 4: mouse A inoculated with wild-type PMC21 NhhA at a 1:10.000 dilution. Lanes 5 and 6, mouse B inoculated with wild-type PMC21 NhhA at a 1:1000 dilution. Lanes 7 and 8: mouse B inoculated with wild-type PMC21 NhhA at a 1:10.000 dilution. Panel B: Lanes 1 & 2: mouse C inoculated with truncated PMC21 NhhA polypeptide (Example 4) at a 1:1000 dilution. Lanes 3 & 4: mouse C inoculated with truncated PMC21 NhhA polypeptide (Example 4) at a 1:10,000 dilution. Lanes 5 & 6: mouse D inoculated with truncated PMC21 NhhA (Example 4) at a1:1000 dilution. Lanes 7 and 8: mouse D inoculated with truncated PMC21 NhhA (Example 4)1:1000 dilution. Panel C: Lanes 1 & 2: mouse E inoculated with truncated PMC21 NhhA (Example 6) at a 1:1000 dilution. Lanes 3 and 4: mouse E inoculated with truncated PMC21 NhhA (Example 6) at a 1:10,000 dilution. Lanes 5 & 6: mouse F inoculated with truncated PMC21 NhhA (Example 6) at a 1:1000 dilution. Lanes 7 & 8: mouse F inoculated with truncated PMC21 NhhA (Example 6) at a 1:1000 dilution.

FIG. 14: Predicted mature NhhA polypeptide deletion mutants. A: predicted mature protein described in Example 2 (SEQ ID NO:33); B: predicted mature protein described in Example 3 (SEQ ID NO:34); C: predicted mature protein described in Example 4 (SEQ ID NO:35); D: predicted mature protein described in Example 5 (SEQ ID NO:36); E: predicted mature protein described in Example 6 (SEQ ID NO:37): F: predicted mature protein described in Example 7 (SEQ ID NO:38); and G: predicted mature protein described in Example 8 (SEQ ID NO:39).

DETAILED DESCRIPTION OF THE INVENTION

Throughout this specification, unless the context requires otherwise, the words "comprise", "comprises" and "comprising" will be understood to imply the inclusion of a stated integer or group of integers but not the exclusion of any other integer or group of integers.

With regard to nomenclature, NhhA is used herein when reference is made to proteins of the invention, while *nhhA* is used herein when reference is made to nucleic acids of the invention. It will also be understood that NhhA/*nhhA* proteins and nucleic acids include the HiaNm/*hianm* proteins and nucleic acids referred to in WO99/31132, for example, without limitation thereto.

The present invention is predicated, at least in part, by the elucidation of conserved and less-conserved regions in the NhhA polypeptide in ten (10) strains of *N. meningitidis*. Corresponding regions are predicted to be conserved in other allelic variants of the exemplified NhhA polypeptides.

It will be appreciated that central to the present invention is the realization that by deleting non-conserved amino acids in a wild-type NhhA polypeptide to form a modified NhhA polypeptide of the invention, an immune response may be elicited upon immunization by said polypeptide of the invention which, by directing the immune response against conserved epitopes, will provide protection against one or more heterologous strains of *N. meningitidis*.

As used herein, "*non-conserved*" amino acids are amino acid residues present in a wild-type NhhA polypeptide from a first *N. meningitidis* strain, but which are not present in a wild-type NhhA polypeptide from one or more other strains.

Suitably, the polypeptides of the first aspect have at least a portion of one of the V1, V2, V3 or V4 regions deleted with respect to the corresponding wild-type sequence, and accordingly, may be collectively referred to as examples of "*deletion mutants*".

It will be appreciated that the present inventors have identified the V1, V2, V3 and V4 regions as being regions of wild-type NhhA polypeptides having relatively high frequencies of non-conserved amino acids compared to the relatively conserved C1-5 regions.

Of the V regions, the V1 (hypervariable) and V2 regions have the highest frequency of non-conserved amino acids, while V3 and V4 have relatively lower frequencies. However, the V1 region constitutes a more significant proportion of wild-type NhhA polypeptides than does the V2 region (in terms of total amino acids). Therefore, it is preferred that the isolated proteins according to the first-mentioned aspect have at least a substantial portion of the V1 region deleted.

It will also be realized by the skilled person that in constructing said deletion mutants, "shuffling"of regions between NhhA polypeptides of different *N. meningitidis* strains is possible. For example, an NhhA polypeptide of the invention may comprise a H41 C1 region together with a PMC21 C5 region.

Such "shuffling" is particularly well-suited to recombinant DNA methods.

For the purposes of this invention, by *"isolated"* is meant material that has been removed from its natural state or otherwise been subjected to human manipulation. Isolated material may be substantially or essentially free from components that normally accompany it in its natural state, or may be manipulated so as to be in an artificial state together with components that normally accompany it in its natural state. Isolated material may be in native or recombinant form.

By "*protein*"is meant an amino acid polymer. The amino acids may be natural or non-natural amino acids as are well understood in the art.

A "*peptide*" is a protein having no more than fifty (50) amino acids.

A polypeptide is a protein having fifty (50) or more amino acids.

As used herein, the phrase "*elicits an immune response*" refers to the ability of an isolated polypeptide of the invention to produce an immune response in a mammal to which it is administered, wherein the response is directed to *N. meningitidis* and/or said polypeptide. Preferably, the immune response includes production of bactericidal antibodies. More preferably, the immune response is protective against *N. meningitidis* infection.

"*Strain-specific*" is used herein in the context of an immune response which is directed to, or at least predominantly directed to, an autologous *N. meningitidis* strain.

As used herein, "*cross-reactive*" means an ability of a polypeptide of the invention to elicit an immune response directed to one or more heterologous

N. meningitidis strains.

As used herein, "*cross-protective*" means an ability of a polypeptide of the invention to elicit an immune response and thereby provide protection against infection by one or more heterologous *N. meningitidis* strains.

Therefore, in light of the foregoing, said polypeptide of the invention may be referred to herein as an "*immunogen*", or as being "*immunogenic*".

Although for the purposes of the present invention, said modified NhhA proteins have been exemplified by the amino acid sequences set forth in FIGS 5 to 9 (SEQ ID NOS: 23-27) and FIG. 14, the present invention also contemplates fragments, derivatives and variants (such as allelic variants) of the exemplified proteins.

For example, amino acids can be deleted from any of the C1-5 sequences set forth in FIG. 1, while not all non-conserved amino acids in the V1-4 regions need be deleted in order to reduce strain-specific immunogenicity.

Therefore, isolated proteins of the invention may include fragments of the C1-5 and V1-4 regions.

Indeed, as will be described hereinafter in the Examples, it may be advantageous for the purposes of recombinant DNA-based production of polypeptides of the invention, to delete one or a few amino acids of a C1, C2, C3, C4 and/or C5 region or a V1, V2, V3 and/or V4 region in the interests of utilizing convenient restriction endonuclease sites and achieving high level expression of stable, immunogenic protein.

In one embodiment, a "*fragment*" includes an amino acid sequence that constitutes less than 100%, but at least 20%, preferably at least 50%, more preferably at least 80% or even more preferably at least 90% of said C1, C2, C3, C4 or C5 regions.

Fragments, for example, may be peptides comprising as few as twelve amino acids such as the C2 region (SEQ ID NO:11) or sequences of at least twenty contiguous amino acids, or more than one hundred contiguous amino acids corresponding to some or all of the C1, C2, C3, C4 and/or C5 regions described herein.

Other fragments exemplified herein are modified NhhA polypeptides of the invention which have undergone post-translational processing to form a mature polypeptide, such as shown in FIG. 14.

In another embodiment, a *"fragment"* is a small peptide, for example of at least 6, preferably at least 10 and more preferably at least 20 amino acids in length, which comprises one or more antigenic determinants or epitopes derived from modified NhhA proteins of the invention. Larger fragments comprising more than one peptide are also contemplated, and may be obtained through the application of standard recombinant nucleic acid techniques or synthesized using conventional liquid or solid phase synthesis techniques. For example, reference may be made to solution synthesis or solid phase synthesis as described, for example, in Chapter 9 entitled *"Peptide Synthesis"* by Atherton and Shephard which is included in a publication entitled *"Synthetic Vaccines"* edited by Nicholson and published by Blackwell Scientific Publications. Alternatively, peptides can be produced by digestion of a polypeptide of the invention with proteinases such as endoLys-C, endoArg-C, endoGlu-C and staphylococcins V8-protease. The digested fragments can be purified by, for example, high performance liquid chromatographic (HPLC) techniques.

As used herein, *"variant"* polypeptides are polypeptides of the invention in which one or more amino acids have been replaced by different amino acids. It is well understood in the art that some amino acids may be changed to others with broadly similar properties without changing the nature of the activity of the polypeptide (conservative substitutions). Exemplary conservative substitutions in the polypeptide may be made according to Table 2.

Substantial changes in function are made by selecting substitutions that are less conservative than those shown in Table 2. Other replacements would be non-conservative substitutions and relatively fewer of these may be tolerated. Generally, the substitutions which are likely to produce the greatest changes in a polypeptide's properties are those in which (a) a hydrophilic residue (*e.g.*, Ser or Thr) is substituted for, or by, a hydrophobic residue (*e.g.*, Ala, Leu, Ile, Phe or

Val); (b) a cysteine or proline is substituted for, or by, any other residue; (c) a residue having an electropositive side chain (*e.g.*, Arg, His or Lys) is substituted for, or by, an electronegative residue (*e.g.*, Glu or Asp) or (d) a residue having a bulky side chain (*e.g.*, Phe or Trp) is substituted for, or by, one having a smaller side chain (*e.g.*, Ala, Ser)or no side chain (*e.g.*, Gly).

The term *"variant"* also includes NhhA polypeptides of the invention produced from allelic variants of the sequences exemplified in this specification.

NhhA polypeptide variants may fall within the scope of the term "*polypeptide homologs*".

Polypeptide homologs share at least 70%, preferably at least 80% and more preferably at least 90% sequence identity with the amino acid sequences of modified NhhA polypeptides of the invention as hereinbefore described.

As generally used herein, a "*homolog*" shares a definable nucleotide or amino acid sequence relationship with a nucleic acid or polypeptide of the invention as the case may be.

For example, such homologs are contemplated as having amino acid sequences that differ from those exemplified herein, but which are immunogenic and provide cross-protective immunity.

Specifically excluded from the scope of the term "*homologs*" are wild-type NhhA polypeptides and *nhhA* nucleic acids.

Included within the scope of homologs are "*orthologs*", which are functionally-related polypeptides and their encoding nucleic acids, isolated from bacterial species other than *N. meningitidis*.

Terms used herein to describe sequence relationships between respective nucleic acids and polypeptides include "comparison window", "sequence identity", "percentage of sequence identity" and "substantial identity". Because respective nucleic acids/polypeptides may each comprise (1) only one or more portions of a complete nucleic acid/polypeptide sequence that are shared by the nucleic acids/polypeptides, and (2) one or more portions which are divergent between the nucleic acids/polypeptides, sequence comparisons are typically performed by comparing sequences over a "comparison window" to identify and compare local regions of sequence similarity. A "*comparison window*" refers to a conceptual segment of typically 12 contiguous residues that is compared to a reference sequence. The comparison window may comprise additions or deletions (*i.e.*, gaps) of about 20% or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the respective sequences. Optimal alignment of sequences for aligning a comparison window may be conducted by computerised implementations of algorithms (Geneworks program by Intelligenetics; GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package Release 7.0, Genetics Computer Group, 575 Science Drive Madison, WI, USA, incorporated herein by reference) or by inspection and the best alignment (i.e., resulting in the highest percentage homology over the comparison window) generated by any of the various methods selected. Reference also may be made to the BLAST family of programs as for example disclosed by Altschul *et al.*, 1997, Nucl. Acids Res. **25** 3389, which is incorporated herein by reference.

A detailed discussion of sequence analysis can be found in Unit 19.3 of CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel *et al.* (John Wiley & Sons Inc NY, 1995-1999).

The term "*sequence identity*" is used herein in its broadest sense to include the number of exact nucleotide or amino acid matches having regard to an appropriate alignment using a standard algorithm, having regard to the extent that sequences are identical over a window of comparison. Thus, a "*percentage of sequence identity*" is calculated by comparing two optimally aligned sequences over the window of comparison, determining the number of positions at which the identical nucleic acid base (e.g., A, T, C, G, I) occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison (i.e., the window size), and multiplying the result by 100 to yield the percentage of sequence identity. For example, "*sequence identity*" may be understood to mean the "match percentage" calculated by the DNASIS computer program (Version 2.5 for windows; available from Hitachi Software engineering Co., Ltd., South San Francisco, California, USA).

Thus, it is well within the capabilities of the skilled person to prepare polypeptide homologs of the invention, such as variants as hereinbefore defined, by recombinant DNA technology. For example, nucleic acids of the invention can be mutated using either random mutagenesis for example using transposon mutagenesis, or site-directed mutagenesis. The resultant DNA fragments are then cloned into suitable expression hosts such as *E. coli* using conventional technology and clones that retain the desired activity are detected. Where the clones have been derived using random mutagenesis techniques, positive clones would have to be sequenced in order to detect the mutation.

As used herein, *"derivative"* polypeptides are polypeptides of the invention which have been altered, for example by conjugation or complexing with other chemical moieties or by post-translational modification techniques as would be understood in the art. Such derivatives include amino acid deletions and/or additions to NhhA polypeptides of the invention, or variants thereof, wherein said derivatives elicit an immune response.

"*Additions*" of amino acids may include fusion of the polypeptides or variants thereof with other polypeptides or proteins. In this regard, it will be appreciated that the polypeptides or variants of the invention may be incorporated into larger polypeptides, and such larger polypeptides may also be expected to be immunogenic. The polypeptides as described above may be fused to a further protein, for example, which is not derived from *N. meningitidis*. The other protein may, by way of example, assist in the purification of the protein. For instance a polyhistidine tag, or a maltose binding protein may be used. Alternatively, it may produce an immune response which is effective against *N. meningitidis* or it may produce an immune response against another pathogen. Other possible fusion proteins are those which produce an immunomodulatory response. Particular examples of such proteins include Protein A or glutathione S-transferase (GST). In addition, the polypeptide may be fused to an oligosaccharide based vaccine component where it acts as a carrier protein.

Other derivatives contemplated by the invention include, but are not limited to, modification to side chains, incorporation of unnatural amino acids and/or their derivatives during peptide, polypeptide or protein synthesis and the use of crosslinkers and other methods which impose conformational constraints on the polypeptides, fragments and variants of the invention. Examples of side chain modifications contemplated by the present invention include modifications of amino groups such as by acylation with acetic anhydride; acylation of amino groups with succinic anhydride and tetrahydrophthalic anhydride; amidination with methylacetimidate; carbamoylation of amino groups with cyanate; pyridoxylation of lysine with pyridoxal-5-phosphate followed by reduction with $NaBH_4$; reductive alkylation by reaction with an aldehyde followed by reduction with $NaBH_4$; and trinitrobenzylation of amino groups with 2, 4, 6-trinitrobenzene sulphonic acid (TNBS).

The carboxyl group may be modified by carbodiimide activation via O-acylisourea formation followed by subsequent derivitization, by way of example, to a corresponding amide.

The guanidine group of arginine residues may be modified by formation of heterocyclic condensation products with reagents such as 2,3-butanedione, phenylglyoxal and glyoxal.

Sulphydryl groups may be modified by methods such as performic acid oxidation to cysteic acid; formation of mercurial derivatives using 4-chloromercuriphenylsulphonic acid, 4-chloromercuribenzoate; 2-chloromercuri-4-nitrophenol, phenylmercury chloride, and other mercurials; formation of a mixed disulphides with other thiol compounds; reaction with maleimide, maleic anhydride or other substituted maleimide; carboxymethylation with iodoacetic acid or iodoacetamide; and carbamoylation with cyanate at alkaline pH.

Tryptophan residues may be modified, for example, by alkylation of the indole ring with 2-hydroxy-5-nitrobenzyl bromide or sulphonyl halides or by oxidation with N-bromosuccinimide.

Tyrosine residues may be modified by nitration with tetranitromethane to form a 3-nitrotyrosine derivative.

The imidazole ring of a histidine residue may be modified by N-carbethoxylation with diethylpyrocarbonate or by alkylation with iodoacetic acid derivatives.

Examples of incorporating unnatural amino acids and derivatives during peptide synthesis include but are not limited to, use of 4-amino butyric acid, 6-aminohexanoic acid, 4-amino-3-hydroxy-5-phenylpentanoic acid, 4-amino-3-hydroxy-6-methylheptanoic acid, t-butylglycine, norleucine, norvaline, phenylglycine, ornithine, sarcosine, 2-thienyl alanine and/or D-isomers of amino acids.

The invention also contemplates covalently modifying a polypeptide, fragment or variant of the invention with dinitrophenol, in order to render it immunogenic in humans

Isolated proteins of the invention (inclusive of fragments, variants, derivatives and homologs) may be prepared by any suitable procedure known to those of skill in the art.

For example, the protein may be prepared as a recombinant polypeptide by a procedure including the steps of:

(i) preparing an expression construct which comprises a modified *nhhA* nucleic acid of the invention, operably linked to one or more regulatory nucleotide sequences;

(ii) transfecting or transforming a suitable host cell with the expression construct; and

(iii) expressing the recombinant polypeptide in said host cell.

A number of Examples will be provided hereinafter which describe production of modified *nhhA* nucleic acids of the invention by PCR.

In one particular embodiment, PCR is splice overlap PCR, as will be described hereinafter, which method is based on that described in Ho *et al.*, 1989, Gene **77** 51, and by Horton *et al.*, 1989, Gene **77** 61, which are both incorporated herein by reference.

For the purposes of host cell expression, the recombinant nucleic acid is operably linked to one or more regulatory sequences in an expression vector.

An *"expression vector"* may be either a self-replicating extra-chromosomal vector such as a plasmid, or a vector that integrates into a host genome.

By *"operably linked"* is meant that said regulatory nucleotide sequence(s) is/are positioned relative to the recombinant nucleic acid of the invention to initiate, regulate or otherwise control transcription.

Regulatory nucleotide sequences will generally be appropriate for the host cell used for expression. Numerous types of appropriate expression vectors and suitable regulatory sequences are known in the art for a variety of host cells.

Typically, said one or more regulatory nucleotide sequences may include, but are not limited to, promoter sequences, leader or signal sequences, ribosomal binding sites, transcriptional start and termination sequences, translational start and termination sequences, and enhancer or activator sequences.

Constitutive or inducible promoters as known in the art are contemplated by the invention. The promoters may be either naturally occurring promoters, or hybrid promoters that combine elements of more than one promoter.

In a preferred embodiment, the expression vector contains a selectable marker gene to allow the selection of transformed host cells. Selectable marker genes are well known in the art and will vary with the host cell used.

In an embodiment, the expression vector is pCO14K, which has a *porA* promoter and kanamycin selection gene, as will be described in detail hereinafter. According to this embodiment, the host cell is a bacterium selected from the group consisting of *E. coli* and *N. meningitidis*.

The expression vector may also include a fusion partner (typically provided by the expression vector) so that the recombinant polypeptide of the invention is expressed as a fusion polypeptide with said fusion partner. The main advantage of fusion partners is that they assist identification and/or purification of said fusion polypeptide.

In order to express said fusion polypeptide, it is necessary to ligate a nucleotide sequence according to the invention into the expression vector so that the translational reading frames of the fusion partner and the nucleotide sequence of the invention coincide.

Well known examples of fusion partners include, but are not limited to, glutathione-S-transferase (GST), Fc potion of human IgG, maltose binding protein (MBP) and hexahistidine (HIS_6), which are particularly useful for isolation of the fusion polypeptide by affinity chromatography. For the purposes of fusion polypeptide purification by affinity chromatography, relevant matrices for affinity chromatography are glutathione-, amylose-, and nickel- or cobalt-conjugated resins respectively. Many such matrices are available in "kit" form, such as the QIAexpress™ system (Qiagen) useful with (HIS_6) fusion partners and the Pharmacia GST purification system.

A preferred fusion partner is MBP, which is described hereinafter in Example 11.

Another fusion partner well known in the art is green fluorescent protein (GFP). This fusion partner serves as a fluorescent "tag" which allows the fusion polypeptide of the invention to be identified by fluorescence microscopy or by flow cytometry. The GFP tag is useful when assessing subcellular localization of the fusion polypeptide of the invention, or for isolating cells which express the fusion polypeptide of the invention. Flow cytometric methods such as fluorescence activated cell sorting (FACS) are particularly useful in this latter application.

Preferably, the fusion partners also have protease cleavage sites, such as for Factor X_a or Thrombin, which allow the relevant protease to partially digest the fusion polypeptide of the invention and thereby liberate the recombinant polypeptide of the invention therefrom. The liberated polypeptide can then be isolated from the fusion partner by subsequent chromatographic separation.

Fusion partners according to the invention also include within their scope "epitope tags", which are usually short peptide sequences for which a specific antibody is available. Well known examples of epitope tags for which specific monoclonal antibodies are readily available include c-myc, influenza virus haemagglutinin and FLAG tags.

As hereinbefore, polypeptides of the invention may be produced by culturing a host cell transformed with said expression construct comprising a nucleic acid encoding a polypeptide, or polypeptide homolog, of the invention. The conditions appropriate for protein expression will vary with the choice of expression vector and the host cell. This is easily ascertained by one skilled in the art through routine experimentation.

Suitable host cells for expression may be prokaryotic or eukaryotic. One preferred host cell for expression of a polypeptide according to the invention is a bacterium. The bacterium used may be *Escherichia coli* or *N. meningitidis*.

In a preferred embodiment, the host cell is *N. meningitidis* which has been modified so as to not express PorA, Opa, Opc or capsular polysaccharide and expresses a desired lipopolysaccharide phenotype.

Alternatively, the host cell may be an insect cell such as, for example, *SF9* cells that may be utilized with a baculovirus expression system.

The recombinant protein may be conveniently prepared by a person skilled in the art using standard protocols as for example described in Sambrook, *et al.*, MOLECULAR CLONING. A Laboratory Manual (Cold Spring Harbor Press, 1989), incorporated herein by reference, in particular Sections 16 and 17; CURRENT PROTOCOLS IN MOLECULAR BIOLOGY Eds. Ausubel *et al.*, (John Wiley & Sons, Inc. 1995-1999), incorporated herein by reference, in particular Chapters 10 and 16; and CURRENT PROTOCOLS IN PROTEIN SCIENCE Eds. Coligan *et al.*, (John Wiley & Sons, Inc. 1995-1999) which is incorporated by reference herein, in particular Chapters 1, 5 and 6.

Preferred methods of expression of recombinant modified NhhA proteins of the invention, and methods for detection of expressed protein, are provided hereinafter in the Examples.

Nucleotide sequences

The invention provides an isolated nucleic acid that encodes a modified NhhA protein of the invention

Preferably, said isolated nucleic acid has a nucleotide sequence that encodes one or more NhhA polypeptide constant (C) regions as described in FIGS.

1 and 2. The isolated nucleic acid may further encode one or more non-conserved (V region) amino acids such as also identified in FIGS 1 and 2.

Particular embodiments of such isolated nucleic acids are provided in SEQ ID NOS: 28-32 and FIGS. 5-9.

The term *"nucleic acid"* as used herein designates single-or double-stranded mRNA, RNA, cRNA and DNA, said DNA inclusive of cDNA and genomic DNA.

A *"polynucleotide"* is a nucleic acid having eighty (80) or more contiguous nucleotides, while an *"oligonucleotide"* has less than eighty (80) contiguous nucleotides.

A *"probe"* may be a single or double-stranded oligonucleotide or polynucleotide, suitably labeled for the purpose of detecting complementary sequences in Northern or Southern blotting, for example.

A *"primer"* is usually a single-stranded oligonucleotide, preferably having 15-50 contiguous nucleotides, which is capable of annealing to a complementary nucleic acid "template" and being extended in a template-dependent fashion by the action of a DNA polymerase such as *Taq* polymerase, RNA-dependent DNA polymerase or Sequenase™.

The present invention also contemplates homologs of nucleic acids of the invention as hereinbefore defined.

Such nucleic acid homologs exclude nucleic acids encoding full-length wild-type NhhA polypeptides.

For example, nucleic acid homologs encode peptides and polypeptides, structurally related to NhhA V and C regions of the invention, that may be useful for the purposes of providing cross-protective immunity to *N. meningitidis* by immunization.

In one embodiment, nucleic acid homologs encode polypeptide homologs of the invention, inclusive of variants, fragments and derivatives thereof.

In another embodiment, nucleic acid homologs share at least 60%, preferably at least 70%, more preferably at least 80%, and even more preferably at least 90% sequence identity with the nucleic acids of the invention.

In yet another embodiment, nucleic acid homologs hybridize to nucleic acids of the invention under at least low stringency conditions, preferably under at least medium stringency conditions and more preferably under high stringency conditions.

"*Hybridize and Hybridization*" is used herein to denote the pairing of at least partly complementary nucleotide sequences to produce a DNA-DNA, RNA-RNA or DNA-RNA hybrid. Hybrid sequences comprising complementary nucleotide sequences occur through base-pairing between complementary purines and pyrimidines as are well known in the art.

In this regard, it will be appreciated that modified purines (for example, inosine, methylinosine and methyladenosine) and modified pyrimidines (thiouridine and methylcytosine) may also engage in base pairing.

"Stringency" as used herein, refers to temperature and ionic strength conditions, and presence or absence of certain organic solvents and/or detergents during hybridisation. The higher the stringency, the higher will be the required level of complementarity between hybridizing nucleotide sequences.

"Stringent conditions" designates those conditions under which only nucleic acid having a high frequency of complementary bases will hybridize.

Reference herein to low stringency conditions includes and encompasses:-

(i) from at least about 1% v/v to at least about 15% v/v formamide and from at least about 1 M to at least about 2 M salt for hybridisation at 42°C, and at least about 1 M to at least about 2 M salt for washing at 42°C; and

(ii) 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (i) 2xSSC, 0.1% SDS; or (ii) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at room temperature.

Medium stringency conditions include and encompass:-

(i) from at least about 16% v/v to at least about 30% v/v formamide and from at least about 0.5 M to at least about 0.9 M salt for hybridisation at 42°C, and at least about 0.5 M to at least about 0.9 M salt for washing at 42°C; and

(ii) 1% Bovine Serum Albumin (BSA), 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C and (a) 2 x SSC, 0.1% SDS; or (b) 0.5% BSA, 1 mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 5% SDS for washing at 42°C.

High stringency conditions include and encompass:-

(i) from at least about 31% v/v to at least about 50% v/v formamide and from at least about 0.01 M to at least about 0.15 M salt for hybridisation at 42°C, and at least about 0.01 M to at least about 0.15 M salt for washing at 42°C;

(ii) 1% BSA, 1 mM EDTA, 0.5 M $NaHPO_4$ (pH 7.2), 7% SDS for hybridization at 65°C, and (a) 0.1 x SSC, 0.1% SDS; or (b) 0.5% BSA, 1mM EDTA, 40 mM $NaHPO_4$ (pH 7.2), 1% SDS for washing at a temperature in excess of 65°C for about one hour; and

(iii) 0.2 x SSC, 0.1% SDS for washing at or above 68°C for about 20 minutes.

In general, washing is carried out at T_m = 69.3 + 0.41 (G + C) % - 12°C. In general, the T_m of a duplex DNA decreases by about 1°C with every increase of 1% in the number of mismatched bases.

Notwithstanding the above, stringent conditions are well known in the art, such as described in Chapters 2.9 and 2.10 of. Ausubel *et al., supra*, which are herein incorporated be reference. A skilled addressee will also recognize that various factors can be manipulated to optimize the specificity of the hybridization. Optimization of the stringency of the final washes can serve to ensure a high degree of hybridization.

Typically, complementary nucleotide sequences are identified by blotting techniques that include a step whereby nucleotides are immobilized on a matrix (preferably a synthetic membrane such as nitrocellulose), a hybridization step, and a detection step. Southern blotting is used to identify a complementary DNA sequence; northern blotting is used to identify a complementary RNA sequence. Dot blotting and slot blotting can be used to identify complementary DNA/DNA, DNA/RNA or RNA/RNA polynucleotide sequences. Such techniques are well known by those skilled in the art, and have been described in Ausubel *et al.*, *supra*, at pages 2.9.1 through 2.9.20. According to such methods, Southern blotting involves separating DNA molecules according to size by gel electrophoresis, transferring the size-separated DNA to a synthetic membrane, and hybridizing the membrane bound DNA to a complementary nucleotide sequence.

In dot blotting and slot blotting, DNA samples are directly applied to a synthetic membrane prior to hybridization as above.

An alternative blotting step is used when identifying complementary nucleic acids in a cDNA or genomic DNA library, such as through the process of plaque or colony hybridization. Other typical examples of this procedure is described in Chapters 8-12 of Sambrook *et al.*, *supra* which are herein incorpoated by reference.

Typically, the following general procedure can be used to determine hybridization conditions. Nucleic acids are blotted/transferred to a synthetic membrane, as described above. A wild type nucleotide sequence of the invention is labeled as described above, and the ability of this labeled nucleic acid to hybridize with an immobilized nucleotide sequence analyzed.

A skilled addressee will recognize that a number of factors influence hybridization. The specific activity of radioactively labeled polynucleotide sequence should typically be greater than or equal to about 10^8 dpm/μg to provide a detectable signal. A radiolabeled nucleotide sequence of specific activity 10^8 to 10^9 dpm/μg can detect approximately 0.5 pg of DNA. It is well known in the art that sufficient DNA must be immobilized on the membrane to permit detection. It is desirable to have excess immobilized DNA, usually 1-10 μg. Adding an inert polymer such as 10% (w/v) dextran sulfate (MW 500,000) or polyethylene glycol 6000 during hybridization can also increase the sensitivity of hybridization (see

Ausubel *et al.*, *supra* at 2.10.10).

To achieve meaningful results from hybridization between a nucleic acid immobilized on a membrane and a labeled nucleic acid, a sufficient amount of the labeled nucleic acid must be hybridized to the immobilized nucleic acid following washing. Washing ensures that the labeled nucleic acid is hybridized only to the immobilized nucleic acid with a desired degree of complementarity to the labeled nucleic acid.

Methods for detecting labeled nucleic acids hybridized to an immobilized nucleic acid are well known to practitioners in the art. Such methods include autoradiography, chemiluminescent, fluorescent and colorimetric detection.

In another embodiment, nucleic acid homologs of the invention may be prepared according to the following procedure:

(i) obtaining a nucleic acid extract from a suitable host;

(ii) creating primers which are optionally degenerate wherein each comprises a portion of a nucleotide sequence of the invention; and

(iii) using said primers to amplify, via nucleic acid amplification techniques, one or more amplification products from said nucleic acid extract.

Suitably, the host is a bacterium.

Preferably, the host is of the genus *Neisseria*.

More preferably, the host is *N. meningitidis* or *N. lactamica*.

Primers useful according to nucleic acid sequence amplification methods include SEQ ID NOS:40-51 as described in detail hereinafter.

Suitable nucleic acid amplification techniques are well known to the skilled addressee, and include polymerase chain reaction (PCR) as for example described in Chapter 15 of Ausubel *et al. supra*, which is incorporated herein by reference; strand displacement amplification (SDA) as for example described in U.S. Patent No 5,422,252 which is incorporated herein by reference; rolling circle replication (RCR) as for example described in Liu *et al.*, 1996, J. Am. Chem. Soc. **118** 1587 and International application WO 92/01813 and Lizardi *et al.*, (International Application WO 97/19193) which are incorporated herein by reference; nucleic acid sequence-based amplification (NASBA) as for example described by Sooknanan *et al.*,1994, Biotechniques **17** 1077) which is incorporated herein by reference; and Q-β replicase amplification as for example described by Tyagi *et al.*, 1996, Proc. Natl. Acad. Sci. USA **93** 5395 which is incorporated herein by reference.

As used herein, an *"amplification product"* refers to a nucleic acid product generated by nucleic acid amplification techniques.

Antibodies

The invention also contemplates antibodies against the isolated proteins fragments, variants and derivatives of the invention. Antibodies of the invention may be polyclonal or monoclonal. Well-known protocols applicable to antibody production, purification and use may be found, for example, in Chapter 2 of Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY (John Wiley & Sons NY, 1991-1994) and Harlow, E. & Lane, D. *Antibodies: A Laboratory Manual*, Cold Spring Harbor, Cold Spring Harbor Laboratory, 1988, which are both herein incorporated by reference.

Generally, antibodies of the invention bind to or conjugate with a polypeptide, fragment, variant or derivative of the invention. For example, the antibodies may comprise polyclonal antibodies. Such antibodies may be prepared for example by injecting a polypeptide, fragment, variant or derivative of the invention into a production species, which may include mice or rabbits, to obtain polyclonal antisera. Methods of producing polyclonal antibodies are well known to those skilled in the art. Exemplary protocols which may be used are described for example in Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, *supra*, and in Harlow & Lane, 1988, *supra*.

In lieu of the polyclonal antisera obtained in the production species, monoclonal antibodies may be produced using the standard method as for example, described in an article by Köhler & Milstein, 1975, Nature **256**, 495, which is herein incorporated by reference, or by more recent modifications thereof as for example, described in Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, *supra* by immortalizing spleen or other antibody producing cells derived from a production species which has been inoculated with one or more of the polypeptides, fragments, variants or derivatives of the invention.

The invention also includes within its scope antibodies which comprise Fc or Fab fragments of the polyclonal or monoclonal antibodies referred to above. Alternatively, the antibodies may comprise single chain Fv antibodies (scFvs) against the peptides of the invention. Such scFvs may be prepared, for example, in accordance with the methods described respectively in United States Patent No 5,091,513, European Patent No 239,400 or the article by Winter & Milstein, 1991, Nature **349** 293, which are incorporated herein by reference.

The antibodies of the invention may be used for affinity chromatography in isolating natural or recombinant *N. meningitidis* polypeptides. For example reference may be made to immunoaffinity chromatographic procedures described in Chapter 9.5 of Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, *supra.*

The antibodies may be used to:

(i) screen expression libraries to identify variant polypeptides of the invention;

(ii) identify immunoreactive fragments or immunoreactive epitopes; and/or

(iii) detect *N. meningitidis* infection;

as will be described hereinafter but without limitation to these particular uses.

Detection of N. meningitidis

The presence or absence of *N. meningitidis* in an individual may determined by isolating a biological sample from said individual, mixing an antibody or antibody fragment described above with the biological sample, and detecting specifically bound antibody or antibody fragment which indicates the presence of *N. meningitidis* in the sample.

The term *"biological sample"* as used herein refers to a sample that may be extracted, untreated, treated, diluted or concentrated from an individual, such as a patient. Suitably, the biological sample is selected from the group consisting of whole blood, serum, plasma, saliva, urine, sweat, ascitic fluid, peritoneal fluid, synovial fluid, amniotic fluid, cerebrospinal fluid, skin biopsy, and the like.

Any suitable technique for determining formation of the complex may be used. For example, an antibody or antibody fragment according to the invention having a label associated therewith may be utilized in immunoassays. Such immunoassays may include, but are not limited to, radioimmunoassays (RIAs), enzyme-linked immunosorbent assays (ELISAs) and immunochromatographic techniques (ICTs) which are well known those of skill in the art.

For example, reference may be made to Chapter 7 of Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, *supra* which discloses a variety of immunoassays that may be used in accordance with the present invention. Immunoassays may include competitive assays as understood in the art.

The label associated with the antibody or antibody fragment may include the following:

(A) direct attachment of the label to the antibody or antibody fragment;

(B) indirect attachment of the label to the antibody or antibody fragment; i.e., attachment of the label to another assay reagent which subsequently binds to the antibody or antibody fragment; and

(C) attachment to a subsequent reaction product of the antibody or antibody fragment.

The label may be selected from a group including a chromogen, a catalyst, an enzyme, a fluorophore, a chemiluminescent molecule, a lanthanide ion such as Europium (Eu^{34}), a radioisotope and a direct visual label. In the case of a direct visual label, use may be made of a colloidal metallic or non-metallic particle, a dye particle, an enzyme or a substrate, an organic polymer, a latex particle, a liposome, or other vesicle containing a signal producing substance and the like.

A large number of enzymes useful as labels is disclosed in United States Patent Specifications U.S. 4,366,241, U.S. 4,843,000, and U.S. 4,849,338, all of which are herein incorporated by reference. Enzyme labels useful in the present invention include alkaline phosphatase, horseradish peroxidase, luciferase, β-galactosidase, glucose oxidase, lysozyme, malate dehydrogenase and the like.

The enzyme label may be used alone or in combination with a second enzyme in solution.

Suitably, the fluorophore is selected from a group including fluorescein isothiocyanate (FITC), tetramethylrhodamine isothiocyanate (TRITL) or R-Phycoerythrin (RPE).

The invention also extends to a method for detecting infection of patients by *N. meningitidis*, said method comprising the steps of contacting a biological sample from a patient with a polypeptide, fragment, variant or derivative of the invention, and determining the presence or absence of a complex between said polypeptide, fragment, variant or derivative and *N. meningitidis*-specific antibodies in said serum, wherein the presence of said complex is indicative of said infection.

In a preferred embodiment, detection of the above complex is effected by detectably modifying said polypeptide, fragment, variant or derivative with a suitable label as is well known in the art and using such modified compound in an immunoassay as for example described above.

In another aspect, the invention provides a method of detecting *N. meningitidis* bacteria in a biological sample suspected of containing said bacteria, said method comprising the steps of isolating the biological sample from a patient, detecting a nucleic acid sequence according to the invention in said sample which indicates the presence of said bacteria. Detection of the said nucleic acid sequence may be determined using any suitable technique. For example, a labeled nucleic acid according to the invention may be used as a probe in a Southern blot of a nucleic acid extract obtained from a patient as is well known in the art.

Alternatively, a labeled nucleic acid according to the invention may be utilized as a probe in a Northern blot of a RNA extract from the patient.

Preferably, a nucleic acid extract from the patient is utilized in concert with oligonucleotide primers corresponding to sense and antisense sequences of a nucleic acid sequence according to the invention, or flanking sequences thereof, in a nucleic acid amplification reaction such as PCR, or the ligase chain reaction (LCR) as for example described in International Application WO89/09385 which is incorporated by reference herein.

A variety of automated solid-phase detection techniques are also appropriate. For example, very large scale immobilized primer arrays (VLSIPS™) are used for the detection of nucleic acids as for example described by Fodor *et al.*,1991, Science **251** 767 and Kazal *et al.*, 1996, Nature Medicine **2** 753. The above generic techniques are well known to persons skilled in the art.

Pharmaceutical compositions

A further feature of the invention is the use of the polypeptide, fragment, variant or derivative of the invention ("*immunogenic agents*") as actives in a pharmaceutical composition for protecting patients against infection by *N. meningitidis*.

Suitably, the pharmaceutical composition comprises a pharmaceutically-acceptable carrier, diluent or excipient.

By *"pharmaceutically-acceptable carrier, diluent or excipient"* is meant a solid or liquid filler, diluent or encapsulating substance that may be safely used in systemic administration. Depending upon the particular route of administration, a variety of carriers, well known in the art may be used. These carriers may be selected from a group including sugars, starches, cellulose and its derivatives, malt, gelatine, talc, calcium sulfate, vegetable oils, synthetic oils, polyols, alginic acid, phosphate buffered solutions, emulsifiers, isotonic saline and salts such as mineral acid salts including hydrochlorides, bromides and sulfates, organic acids such as acetates, propionates and malonates and pyrogen-free water.

A useful reference describing pharmaceutically acceptable carriers, diluents and excipients is Remington's Pharmaceutical Sciences (Mack Publishing Co. N.J. USA, 1991) which is incorporated herein by reference.

Any safe route of administration may be employed for providing a patient with the composition of the invention. For example, oral, rectal, parenteral, sublingual, buccal, intravenous, intra-articular, intra-muscular, intra-dermal, subcutaneous, inhalational, intraocular, intraperitoneal, intracerebroventricular, transdermal and the like may be employed. Intra-muscular and subcutaneous injection is appropriate, for example, for administration of immunogenic compositions, vaccines and DNA vaccines.

Dosage forms include tablets, dispersions, suspensions, injections, solutions, syrups, troches, capsules, suppositories, aerosols, transdermal patches and the like. These dosage forms may also include injecting or implanting controlled releasing devices designed specifically for this purpose or other forms of implants modified to act additionally in this fashion. Controlled release of the therapeutic agent may be effected by coating the same, for example, with hydrophobic polymers including acrylic resins, waxes, higher aliphatic alcohols, polylactic and polyglycolic acids and certain cellulose derivatives such as hydroxypropylmethyl cellulose. In addition, the controlled release may be effected by using other polymer matrices, liposomes and/or microspheres.

Pharmaceutical compositions of the present invention suitable for oral or parenteral administration may be presented as discrete units such as capsules, sachets or tablets each containing a pre-determined amount of one or more therapeutic agents of the invention, as a powder or granules or as a solution or a suspension in an aqueous liquid, a non-aqueous liquid, an oil-in-water emulsion or a water-in-oil liquid emulsion. Such compositions may be prepared by any of the methods of pharmacy but all methods include the step of bringing into association one or more immunogenic agents as described above with the carrier which constitutes one or more necessary ingredients. In general, the compositions are prepared by uniformly and intimately admixing the immunogenic agents of the invention with liquid carriers or finely divided solid carriers or both, and then, if necessary, shaping the product into the desired presentation.

The above compositions may be administered in a manner compatible with the dosage formulation, and in such amount as is immunogenically-effective to protect patients from *N. meningitidis* infection. The dose administered to a patient, in the context of the present invention, should be sufficient to effect a beneficial response in a patient over time such as a reduction in the level of *N. meningitidis*, or to inhibit infection by *N. meningitidis*. The quantity of the immunogenic agent(s) to be administered may depend on the subject to be treated inclusive of the age, sex, weight and general health condition thereof. In this regard, precise amounts of the immunogenic agent(s) required to be administered will depend on the judgement of the practitioner.

In determining the effective amount of the immunogenic agent to be administered in the treatment or prophylaxis against *N. meningitidis*, the physician may evaluate circulating plasma levels, progression of disease, and the production of anti-*N. meningitidis* antibodies. In any event, suitable dosages of the immunogenic agents of the invention may be readily determined by those of skill in the art. Such dosages may be in the order of nanograms to milligrams of the immunogenic agents of the invention.

The above compositions may be used as therapeutic or prophylactic vaccines. Accordingly, the invention extends to the production of vaccines containing as actives one or more of the immunogenic agents of the invention. A variety of applicable procedures are contemplated for producing such vaccines. Exemplary procedures include, for example, those described in NEW GENERATION VACCINES (1997, Levine *et al.*, Marcel Dekker, Inc. New York, Basel Hong Kong) which is incorporated herein by reference.

An immunogenic agent according to the invention can be mixed, conjugated or fused with other antigens, including B or T cell epitopes of other antigens. In addition, it can be conjugated to a carrier as described below.

When an haptenic peptide of the invention is used (*i.e.*, a peptide which reacts with cognate antibodies, but cannot itself elicit an immune response), it can be conjugated with an immunogenic carrier. Useful carriers are well known in the art and include for example: thyroglobulin; albumins such as human serum albumin; toxins, toxoids or any mutant crossreactive material (CRM) of the toxin from tetanus, diptheria, pertussis, *Pseudomonas*, *E. coli*, *Staphylococcus*, and *Streptococcus*; polyamino acids such as poly(lysine:glutamic acid); influenza; Rotavirus VP6, Parvovirus VP1 and VP2; hepatitis B virus core protein; hepatitis B virus recombinant vaccine and the like. Alternatively, a fragment or epitope of a carrier protein or other immnogenic protein may be used. For example, a haptenic peptide of the invention can be coupled to a T cell epitope of a bacterial toxin, toxoid or CRM. In this regard, reference may be made to U.S. Patent No 5,785,973 which is incorporated herein by reference.

In addition, a polypeptide, fragment, variant or derivative of the invention may act as a carrier protein in vaccine compositions directed against *Neisseria*, or against other bacteria or viruses.

The immunogenic agents of the invention may be administered as multivalent subunit vaccines in combination with antigens of *N. meningitidis*, or antigens of other organisms inclusive of the pathogenic bacteria *H. influenzae*, *M. catarrhalis*, *N. gonorrhoeae*, *E. coli*, *S. pneumoniae* etc. Alternatively or additionally, they may be administered in concert with oligosaccharide or polysaccharide components of *N. meningitidis*.

The vaccines can also contain a pharmaceutically-acceptable carrier, diluent or excipient as hereinbefore defined..

The vaccines and immunogenic compositions may include an adjuvant as is well known in the art. Adjuvants contemplated by the present invention include, but are not limited to: surface active substances such as hexadecylamine, octadecylamine, octadecyl amino acid esters, lysolecithin, dimethyldioctadecylammonium bromide, N, N-dicoctadecyl-N', N'bis(2-hydroxyethyl-propanediamine), methoxyhexadecylglycerol, and pluronic polyols; polyamines such as pyran, dextransulfate, poly IC carbopol; peptides such as muramyl dipeptide and derivatives, dimethylglycine, tuftsin; oil emulsions; and mineral gels such as aluminum phosphate, aluminum hydroxide or alum; lymphokines, QuilA and immune stimulating complexes (ISCOMS).

With regard to examples of adjuvants, reference is also made to International Publication WO99/36544 incorporated herein by reference.

Vaccination by DNA delivery

Expression constructs comprising modified NhhA proteins of the invention may be administered to humans to prophylactically and/or therapeutically treat the host. In this regard, expression constructs may encode one or more modified NhhA peptides, polypeptides, fragments or derivatives of these, collectively referred to as "*immunogenic agents*".

Expression constructs also include gene therapy constructs, which employ specialized gene therapy vectors such as vaccinia, and viral vectors useful in gene therapy. The latter include adenovirus and adenovirus-associated viruses

(AAV) such as described in Franceschi *et al.*, 2000, J. Cell Biochem. **78** 476, Braun-Falco *et al.*,1999, Gene Ther. **6** 432, retroviral and lentiviral vectors such as described in Buchshacher *et al.*, 2000, Blood **95** 2499 and vectors derived from herpes simplex virus and cytomegalovirus. A general review of gene therapy vectors and delivery methods may be found in Robbins *et al.*, 1998, Trends in Biotech. **16** 35. An exemplary reference which describes a number of vectors potentially suitable for gene therapy using *Neisseria* proteins, and methods of delivery, is International Publication WO99/36544 incorporated herein by reference.

The immunogenic agents of the invention may be expressed by attenuated viral hosts. By "*attenuated viral hosts*" is meant viral vectors that are either naturally, or have been rendered, substantially avirulent. A virus may be rendered substantially avirulent by any suitable physical (*e.g.*, heat treatment) or chemical means (*e.g.*, formaldehyde treatment). By "*substantially avirulent*" is meant a virus whose infectivity has been destroyed. Ideally, the infectivity of the virus is destroyed without affecting the proteins that carry the immunogenicity of the virus. From the foregoing, it will be appreciated that attenuated viral hosts may comprise live viruses or inactivated viruses.

Attenuated viral hosts which may be useful in a vaccine according to the invention may comprise viral vectors inclusive of adenovirus, cytomegalovirus and preferably pox viruses such as vaccinia (see for example Paoletti and Panicali, U.S. Patent No. 4,603,112 which is incorporated herein by reference) and attenuated *Salmonella* strains (see for example Stocker, U.S. Patent No. 4,550,081 which is herein incorporated by reference). Live vaccines are particularly advantageous because they lead to a prolonged stimulus that can confer substantially long-lasting immunity. Another reference which describes a variety of viral vectors potentially suitable for immunization using *Neisseria* proteins, and methods of delivery, is International Publication WO99/36544 incorporated herein by reference.

Multivalent vaccines can be prepared from one or more microorganisms that express different epitopes of *N. meningitidis* (e.g., other surface proteins or epitopes of *N. meningitidis*). In addition, epitopes of other pathogenic microorganisms can be incorporated into the vaccine.

In a preferred embodiment, this will involve the construction of a recombinant vaccinia virus to express a nucleic acid sequence according to the invention. Upon introduction into a host, the recombinant vaccinia virus expresses the immunogenic agent, and thereby elicits a host CTL response. For example, reference may be made to U.S. Patent No 4,722,848, incorporated herein by reference, which describes vaccinia vectors and methods useful in immunization protocols.

A wide variety of other vectors useful for therapeutic administration or immunization with the immunogenic agents of the invention will be apparent to those skilled in the art from the present disclosure.

In a further embodiment, the nucleotide sequence may be used as a vaccine in the form of a "naked DNA" vaccine as is known in the art. For example, an expression vector of the invention may be introduced into a mammal, where it causes production of a polypeptide *in vivo*, against which the host mounts an immune response as for example described in Barry, M. *et al.*, (1995, *Nature*, **377**:632-635) which is hereby incorporated herein by reference.

Detection kits

The present invention also provides kits for the detection of *N. meningitidis* in a biological sample. These will contain one or more particular agents described above depending upon the nature of the test method employed. In this regard, the kits may include one or more of a polypeptide, fragment, variant, derivative, antibody, antibody fragment or nucleic acid according to the invention. The kits may also optionally include appropriate reagents for detection of labels, positive and negative controls, washing solutions, dilution buffers and the like. For example, a nucleic acid-based detection kit may include (i) a nucleic acid according to the invention (which may be used as a positive control), (ii) an oligonucleotide primer according to the invention, and optionally a DNA polymerase, DNA ligase etc depending on the nucleic acid amplification technique employed.

Preparation of immunoreactive fragments

The invention also extends to a method of identifying an immunoreactive fragment of a polypeptide, variant or derivatives according to the invention. This method essentially comprises generating a fragment of the polypeptide, variant or derivative, administering the fragment to a mammal; and detecting an immune response in the mammal. Such response will include production of elements which specifically bind *N. meningitidis* and/or said polypeptide, variant or derivative, and/or a protective effect against *N. meningitidis* infection.

Prior to testing a particular fragment for immunoreactivity in the above method, a variety of predictive methods may be used to deduce whether a particular fragment can be used to obtain an antibody that cross-reacts with the native antigen. These predictive methods may be based on amino-terminal or carboxy-terminal sequence as for example described in Chapter 11.14 of Ausubel *et al.*, *supra*. Alternatively, these predictive methods may be based on predictions of hydrophilicity as for example described by Kyte & Doolittle 1982, J. Mol. Biol. **157** 105 and Hopp & Woods, 1983, Mol. Immunol. **20** 483) which are incorporated by reference herein, or predictions of secondary structure as for example described by Choo & Fasman,1978, Ann. Rev. Biochem. **47** 251), which is incorporated herein by reference.

In addition, "*epitope mapping*" uses monoclonal antibodies of the invention to identify cross-reactive epitopes by first testing their ability to provide cross-protection, followed by identifying the epitope recognized by said antibodies. An exemplary method is provided in Coligan *et al.*, CURRENT PROTOCOLS IN IMMUNOLOGY, *supra*.

Generally, peptide fragments consisting of 10 to 15 residues provide optimal results. Peptides as small as 6 or as large as 20 residues have worked successfully. Such peptide fragments may then be chemically coupled to a carrier molecule such as keyhole limpet hemocyanin (KLH) or bovine serum albumin (BSA) as for example described in Sections 11.14 and 11.15 of Ausubel *et al.*, *supra*).

It will also be appreciated that peptides may be synthetically circularized, as for example described in Hoogerhout *et al.*, 1995, Infect. Immun. **63** 3473, which is herein incorporated by reference.

The peptides may be used to immunize an animal as for example discussed above. Antibody titers against the native or parent polypeptide from which the peptide was selected may then be determined by, for example, radioimmunoassay or ELISA as for instance described in Sections 11.16 and 114 of Ausubel *et al.*, *supra*.

Antibodies may then be purified from a relevant biological fluid of the animal by ammonium sulfate fractionation or by chromatography as is well known in the art. Exemplary protocols for antibody purification are given in Sections 10.11 and 11.13 of Ausubel *et al.*, *supra*, which are herein incorporated by reference.

Immunoreactivity of the antibody against the native or parent polypeptide may be determined by any relevant procedure such as, for example, Western blot.

Functional blockers

The wild-type NhhA/HiaNm polypeptides disclosed in WO99/31132 are believed to have adhesin properties. They in fact have some similarity to adhesins of *Haemophilus influenzae* which are surface antigens. Specifically they are approximately 67% homologous to the Hia protein of *H. influenzae* (Barenkamp & St. Geme III, 1996, Molecular Microbiology **19** 1215), and 74% homologous to the Hsf protein of *H. influenzae* (St. Geme III, J. *et al*, 1996, Journal of Bacteriology **178** 6281; and U.S. Patent No 5,646,259). For these comparisons, a gap weight of 3, and length weight of 0.01 was used using the GAP program (Deveraux, 1984, *supra*). Thus, interruption of the function of these polypeptides would be of significant therapeutic benefit since they would prevent *N. meningitidis* bacteria from adhering to and invading cells. Interruption of the function may be effected in several ways.

For example, moieties such as chemical reagents or polypeptides which block receptors on the cell surface which interact with a polypeptides of the invention may be administered. These compete with the infective organism for receptor sites. Such moieties may comprise for example polypeptides of the invention, in particular fragments, or functional equivalents of these as well as mimetics.

The term *"mimetics"* is used herein to refer to chemicals that are designed to resemble particular functional regions of the proteins or peptides. Anti-idiotypic antibodies raised against the above-described antibodies which block the binding of the bacteria to a cell surface may also be used. Alternatively, moieties which interact with the receptor binding sites in the polypeptides of the invention may effectively prevent infection of a cell by *N. meningitidis*. Such moieties may comprise blocking antibodies, peptides or other chemical reagents.

All such moieties, pharmaceutical compositions in which they are combined with pharmaceutically acceptable carriers and methods of treating patients suffering from *N. meningitidis* infection by administration of such moieties or compositions form a further aspect of the invention.

The polypeptides of the invention may be used in the screening of compounds for their use in the above methods. For example, polypeptides of the invention may be combined with a label and exposed to a cell culture in the presence of a reagent under test. The ability of reagent to inhibit the binding of the labeled polypeptide to the cell surface can then be observed. In such a screen, the labeled polypeptides may be used directly on an organism such as *E. coli*. Alternatively, *N. meningitidis* itself may be engineered to express a modified and detectable form of the polypeptide. The use of engineered *N. meningitidis* strains in this method is preferred as it is more likely that the tertiary structure of the protein will resemble more closely that expressed in wild-type bacteria.

In order that the invention may be readily understood and put into practical effect, particular preferred embodiments will now be described by way of the following non-limiting examples.

EXAMPLE 1

Identification of constant and variable regions of NhhA polypeptides

The present inventors have elucidated NhhA amino acid sequences which are conserved and/or non-conserved between ten (10) strains of *N. meningitidis*. The non-conserved regions are subdivided into four variable regions (V1, V2, V3 and V4) and the conserved regions are subdivided into C1, C2, C3, C4 and C5 (as shown in FIG. 1 and Table 1; SEQ ID NOS: 1-11). The corresponding nucleotide sequence comparison is shown in FIG 2 (SEQ ID NOS: 12-22).

EXAMPLE 2

PMC 21 NhhA polypeptide over-expression

The NhhA protein encoded by the *nhhA* gene of *N. meningitidis* strain PMC21was over expressed by making an expression construct wherein the *nhhA* gene is operably linked to a promoter.

The following oligonucleotide primers were used to amplify an *N. meningitidis* PMC21 strain *nhhA* nucleic acid open reading frame by PCR:-

HOMP5': 5'-CAA TTA A<u>CG</u> <u>GCC</u> <u>G</u>AA TAA AAG GAA GCC GAT **ATG AAC AAA ATA TAC CGC ATC**-3' (SEQ ID NO 40); which contains an *Eag*I restriction site (underlined) and the sequence encoding the first 7 (seven) amino acids of NhhA (bold type)

HOMP3'AN 5'-TGG AAT <u>CCA TGG</u> **AAT CGC CAC CCT TCC CTT C**-3' (SEQ ID NO 41); which contains an *Nco*I restriction site (underlined) and the reverse complement of sequence 48-61 nucleotides past the end of the *nhhA* open reading frame of strain ¢3 (bold type)

The amplification product contained restriction sites which were subsequently digested with *Eag*I and *Nco*I restriction endonucleases.

The plasmid used for subcloning was pCO14K, which plasmid contains a *porA* promoter upstream of the gene encoding the strongly expressed Class 1 outer membrane protein of *N. meningitidis* together with flanking sequence of *N. meningitidis* strain 2996 and a selectable kanamycin resistance gene as described by Rouppe van der Voort, *et al.*, Infect Immun 1996 **64** 2745.

The digested amplification product was then ligated into *Eag*I and *Nco*I restriction endonuclease-digested pCO14K. This ligation resulted in the replacement of the majority of the *porA* open reading frame with the *nhhA* amplification product (FIG 3). This created a recombinant nucleic acid expression construct (open reading frame shown in SEQ ID NO 12) which encodes a polypeptide of 591 amino acids as shown in SEQ ID NO 1.

This places expression of the *nhhA* nucleic acid of the invention under the control of the strong *porA* promoter. Translation begins at the ATG codon beginning at position 31 of HOMP5'. In order to prevent formation of a fusion between the *porA* and *nhhA*, the HOMP5' sequence contains a TAA stop codon prior to the initiating ATG described above.

The resulting plasmid, pIP52(PMC21), was linearized by restriction digestion and used to transform *N. meningitidis* strain 7G2 using the method described by Janik *et al*, 1976, Journal of Clinical Microbiology **4** 71. Transformants were selected by overnight incubation at 37 °C in 5% CO_2 on solid media containing 100 μg/ml kanamycin. Kanamycin resistant colonies were selected, subcultured overnight and screened for over-expression of NhhA polypeptide by separating total cell proteins electrophoretically on 10% SDS-PAGE followed by transfer to nitrocellulose membrane using a Semi-Dry Blotter (BioRad). The membrane was then incubated sequentially with rabbit anti-NhhA sera (as described in International Publication WO99/31132) and alkaline-phosphatase conjugated anti-Rabbit IgG (Sigma) before colorimetric detection with NBT/BCIP (Sigma). One clone was isolated which expressed NhhA polypeptide at a higher level compared with the parental strain (FIG 11).

Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:33).

The plasmid construct pIP52(PMC21) may be transformed into any transformation-competent strain of *N. meningitidis*

EXAMPLE 3

H41 NhhA polypeptide over-expression

The NhhA protein encoded by the *nhhA* gene of *N. meningitidis* strain H41 was over expressed using the same methods as described in Example 2. This created a recombinant nucleic acid expression construct (open reading frame shown in SEQ ID NO:13) which encodes a polypeptide of 591 amino acids as shown in SEQ ID NO: 2. In this example the resulting plasmid pIP52(H41) was linearized, and transformed into *N. meningitidis* strain 7G2. Kanamycin resistant colonies were analysed and one was chosen which when examined by Western immunoblot, demonstrated overexpression of NhhA. (FIG 11). Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:34).

This strategy may be employed to create expression constructs containing the wild-type *nhhA* sequence of other *N. meningitidis* strains.

EXAMPLE 4

NhhA deletion mutant construction using convenient restriction site

For ease of reference, the amino acid sequence of the NhhA polypeptide encoded by the *nhhA* nucleic acid of strain PMC21 is shown in SEQ ID NO 1. The present inventors created a deletion mutant version of wild-type PMC21 *nhhA*, in which the most variable region between strains was deleted. An amplification product encoding amino acids 1-54 of the wild-type PMC21 NhhA polypeptide was generated by PCR amplification from *nhhA* nucleic acid template using the following primers:

HOMP5': 5'-CAA TTA ACG GCC GAA TAA AAG GAA GCC GAT ATG AAC AAA ATA TAC CGC ATC-3' (SEQ ID NO 40); which is the same oligonucleotide used to create the overexpression construct pIP52.

NH3'BG: 5'-GGT CAG ATC TGT **TTC ATT GTT AGC ACT TGC**-3' (SEQ ID NO 42); which contains a *Bgl*II restriction site (underlined) and the reverse complement of sequence encoding amino acids 134, (double underlined) and 49-54 of wild-type PMC21 NhhA (bold type).

The resulting amplification product included an *Eag*I and *Bgl*II restriction endonuclease sites. pIP52(PMC21) includes a single *Eag*I site 20 bp upstream of the start of the *nhhA* open reading frame (ORF) and a single *Bgl*II site located within the ORF (see Figure 3B). Therefore, pIP52(PMC21) and the amplification product were subjected to restriction endonuclease digestion with *Eag*I and *Bgl*II, ligated and used to transform competent DH5α strain *E. coli* bacteria; this replaces the *Eag*I/*Bgl*II fragment of pIP52(PMC21) with the PCR product. This created a recombinant nucleic acid expression construct (open reading frame shown in FIG. 5; SEQ ID NO 28) which encodes a polypeptide of 512 amino acids as shown in FIG. 5 (SEQ ID NO 23). This amino acid sequence includes amino acids 1-54 and 134-592 of the wild-type sequence, and thereby deletes the majority of the V1 region, all of the V2 and C2 regions, and part of the C3 region of the wild-type PMC21 NhhA polypeptide.

This plasmid was linearised by restriction digestion and transformed in to *N. meningitidis* strain 7G2. Using methods as described in Example 1, one clone was isolated which overexpresses the truncated PMC21 NhhA (FIG 11).

Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:35). To confirm the presence of a cleavable signal sequence and to confirm the identity of the over expressed protein, outer membrane proteins were semi-purified by isolating the fraction that is insoluble in the detergent sarkosyl.

The isolated membrane proteins were separated electrophoretically before transfer to Nylon membrane. The position of the over expressed protein was revealed by Coomassie stain. This region of the membrane was excised and the protein was N-terminal sequenced. The first eleven amino acids of this protein were XXETDLTSVGT which corresponds to amino acid residues 52 to 62 (inclusive) of the amino acid sequence predicted to be expressed by the over expression construct as defined in this example.

This is an example of a deletion using existing restriction sites within the polynucleotide sequence. This construct may be transformed into any transformation competent *N. meningitidis*.

EXAMPLE 5

NhhA deletion mutant construction using convenient restriction site

An expression construct containing the wild-type *nhhA* sequence of H41 was made as described in Example 2. The resulting expression construct was named pIP52(H41). A deletion mutant was made, using the strategy outlined in this example. In this instance the oligonucleotide primers used were:

HOMP5': 5'-CAA TTA ACG GCC GAA TAA AAG GAA GCC GAT ATG AAC AAA ATA TAC CGC ATC-3' (SEQ ID NO:40); which is the same oligonucleotide used to create the overexpression construct pIP52

NH3'STU: 5'-GAT CAG GCC TGT **ATC TTC ATC GGT AGC ATT** -3' (SEQ ID NO 43); which contains a *Stu*I restriction site (underlined) and the reverse complement of sequence encoding amino acids 134, (double underlined) and 49-54 of wild-type H41 NhhA (bold type).

The resulting amplification product contains single *Eag*I and *Stu*I restriction endonuclease sites. The expression construct pIP52(H41) contains these restriction sites. Therefore, pIP52(H41) and the amplification product were subjected to restriction endonuclease digestion with *Eag*I and *Stu*I, ligated and used to transform competent DH5α strain *E. coli* bacteria; this ligation replaces the *Eag*I/*Stu*I fragment of pIP52(H41) with the PCR product. This created a recombinant nucleic acid expression construct (open reading frame shown in FIG. 6 and SEQ ID NO 29) which encodes a polypeptide of 513 amino acids as shown in FIG. 6 and SEQ ID NO 24. This amino acid sequence includes amino acids 1-54 and 134-593 of the wild-type sequence, and thereby deletes the majority of the V1 region, all of the V2 and C2 regions, and part of the C3 region of the wild-type H41 NhhA polypeptide.

This plasmid was linearised by restriction digestion and transformed in to *N. meningitidis* strain 7G2. Using methods as described in Example 1, one clone was isolated which overexpresses the truncated H41 NhhA (FIG 11).

Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:36).

This construct may be transformed into any competent *N. meningitidis*.

EXAMPLE 6

NhhA deletion mutant construction using splice-overlap PCR

In addition to using convenient restriction sites to delete variable regions from nucleotides encoding NhhA, mutants may also be constructed by use of "Splice Overlap Extension" PCR, as described by Ho *et al.*, 1989, *supra* and by Horton, R.M., *et al.*, 1989, *supra*. In this way, polynucleotides can be generated which encode constant regions, but have variable regions deleted (see Figure 5A, 5B, 5C).

In this example, a construct was made containing the C1 and C5 regions, and all other regions deleted (see Figure 5A).

The following oligonucleotide primers were used in PCR reactions to amplify DNA corresponding to region C1 (see FIG.1) from chromosomal DNA of strain PMC21:

HOMP5': 5'-CAA TTA ACG GCC GAA TAA AAG GAA GCC GAT ATG AAC AAA ATA TAC CGC ATC-3' (SEQ ID NO:40); which is the same oligonucleotide used to create the over-expression construct pIP52(PMC21)

SO-C: 5'-GAC GAA ATC AAC GTT **CTT AGC ACT TGC CTG AAC CGT TGC**-3' (SEQ ID NO 44); which sequence is the reverse complement of sequence encoding amino acids 237-241 at the start of the C5 region (underlined) and amino acids 45-52 at the end of the C1 region (bold type) of wild-type NhhA of strain PMC21.

The amplification product of this reaction is HOMP5'/SO-C.

The following oligonucleotide primers were used in PCR reactions to amplify C5 from chromosomal DNA of strain PMC21:

SO- D: 5'-AAC GTT GAT TTC GTC CGC ACT TAC-3' (SEQ ID NO 45); which encodes amino acids 237-244 at the start of C5 (underlined indicates reverse complement of Primer SO-C),

HO3'AN: 5'-TGG AAT CCA TGG AAT CGC CAC CCT TCC CTT C-3' (SEQ ID NO 41); which is the same primer used in the construction of pIP52.

The amplification product of this reaction is SO-D/HO3'AN.

The amplification products HOMP5'/SO-C and SO-D/HO3'AN were purified from an agarose gel following separation by electrophoresis, were mixed, and subjected to further amplification using primers HOMP5' and

HO3'AN. The resulting amplification product encodes amino acids 1-52 and 337-591 of wild-type NhhA of PMC21. This amplification product was subjected to restriction digestion with *Eag*I and *Nco*I, and cloned into pCO14K, as described in Example 1. This recombinant molecule contains regions C1 and C5, thus deleting regions V1 to 4 and C2 to 4. The nucleotide sequence of the open reading frame is shown in FIG. 7 and SEQ ID NO 30, and the predicted polypeptide sequence derived from this nucleotide sequence is shown in FIG. 7 and SEQ ID NO 25.

This plasmid was linearized by restriction digestion and transformed in to *N. meningitidis* strain 7G2. Using methods as described in Example 2, one clone was isolated which overexpresses the truncated PMC21 NhhA.

Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:37).

This plasmid may be transformed into any transformation competent strain of *N. meningitidis*.

EXAMPLE 7

NhhA deletion mutant construction using splice-overlap PCR

It will be appreciated that a similar strategy can be used to create recombinant polynucleotides encoding various regions of NhhA. A construct can be made comprising regions C1, C4, V4 and C5 using the following strategy (see Figure 5B):

The C1 region is amplified using oligonucleotide primers:

HOMP5': 5'-CAA TTA ACG GCC GAA TAA AAG GAA GCC GAT ATG AAC AAA ATA TAC CGC ATC-3' (SEQ ID NO:40);

SO-E: 5'-AAC GCT TGC CGC ACG **CTT AGC ACT TGC CTG CAA CGT TGC**-3' (SEQ ID NO 46); which encodes the reverse complement of amino acids 211-215 at the start of the C4 region (underlined) and at the end of the C1 region (bold type) of strain PMC21.

The amplification product of this reaction is HOMP5'/SO-E.

The following oligonucleotide primers are used in PCR reactions to amplify the region C4-V4-C5 from chromosomal DNA of strain PMC21:

SO-F: 5'-<u>CGT GCG GCA AGC GTT</u> AAA GAC GTA-3' (SEQ ID NO 47); which encodes amino acids 211-218 at the start of C4 (underlined indicates reverse complement of Primer SO-E),

HO3'AN: 5'-TGG AAT CCA TGG AAT CGC CAC CCT TCC CTT C-3 (SEQ ID NO: 41).

The amplification product of this reaction is SO-F/HOMP3'

The amplification products HOMP5'/SO-E and SO-F/HO3'AN will be purified from agarose gel following separation by electrophoresis, and will be mixed, and subjected to further amplification using primers HOMP5' and HO3'AN. The resulting product encodes amino acids 1-52 and 211-591 of wild-type NhhA of PMC21. This amplification product will be subjected to restriction digestion with *Eag*I and *Nco*I, and cloned into pCO14K. This recombinant molecule contains regions C1, C4, V4 and C5 thus deleting regions V1-3 and C2-3. The nucleotide sequence of the open reading frame is shown in FIG. 8 and SEQ ID NO 31, and the predicted polypeptide sequence derived from this nucleotide sequence is shown in FIG. 8 and SEQ ID NO 26. Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 51 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:38).

This construct can be transformed into any transformation competent *N. meningitidis*.

EXAMPLE 8

NhhA deletion mutant construction using splice-overlap PCR

It will be appreciated that a similar strategy can be used to create recombinant polynucleotides encoding various regions of NhhA. A construct can be made comprising regions C1, C2, C3, C4, and C5 using the following strategy (see Figure 5C):

C1 and C2 will be amplified using oligonucleotide primers:

HOMP5': 5'-CAA TTA ACG GCC GAA TAA AAG GAA GCC GAT ATG AAC AAA ATA TAC CGC ATC-3' (SEQ ID NO 40);

SO-G: 5'- <u>CAG CGA GTA GGT GAA</u> <u>**TTG TTT GAT TTT CAG GTT GTC GCC GGC TTT GAG GGT**</u> **GTT AGC ACT TGC CTG AAC CGT**-3' (SEQ ID NO 48); which encodes the reverse complement of amino acids 125-129 at the start of the C3 region (underlined), all of the C2 region (amino acids 109-120, bold and double underlined) and the end of the C1 region (amino acids 46-52, bold type) of strain PMC21.

The amplification product of this reaction is HOMP5'/SO-G.

The C3 and part of C4 regions will be amplified using the following oligonucleotide primers:

SO-H: 5'-<u>TTC ACC TAC TCG CTG</u> AAA AAA GAC-3' (SEQ ID NO 49); which encodes amino acids 125-132 at the start of C3 (underlined indicates reverse complement of Primer SO-G)

SO-I: 5'- **GCC AGC GTT TAA TAC GTC TTT AAC GCT TGC CGC ACG** <u>ATC GGT CAA AGT CGA ACC AAT</u> -3' (SEQ ID NO 50); which encodes the reverse complement of amino acids 182-88 at the end of C3 (underlined) and amino acids 211-222 of C4 (bold type).

The amplification product of this reaction is SO-H/SO-I.

The amplification products HOMP5'/SO-G and SO-H/SO-I are purified from agarose gel following separation by electrophoresis, mixed and subjected to further amplification using primers HOMP5' and SO-I to yield a product encoding amino acids 1-52, 103-114, 125-188, and 211-222, *i.e.* regions C1, C2, C3 and part of C4. The amplification product of this reaction is HOMP5'/SO-I.

The C5 and part of C4 regions are amplified using the following oligonucleotide primers:

SO-J: 5' **GTA TTA AAC GCT GGC** TGG AAC ATT AAA GGC GTT AAA **AAC GTT GAT TTC GTC CGC ACT**-3' (SEQ ID NO 51); which encodes amino acids 218-229 of C4 (underlined), and amino acids 237-243 of C5 (bold type) of wild-type NhhA of strain PMC21. (Bold underlined type indicates reverse complement of SO-I)

HO3'AN: 5'-TGG AAT CCA TGG AAT CGC CAC CCT TCC CTT C-3' (SEQ ID NO:41).

The amplification product of this reaction is SO-J/HO3'AN.

The amplification products HOMP5'/SO-I and SO-J/HO3'AN will be purified from agarose gel following separation by electrophoresis, and will be mixed, and subjected to further amplification using primers HOMP5' and HO3'AN. The resulting product encodes amino acids 1-52, 103-114, 125-188, 211-229, and 237-591 of wild-type NhhA of strain PMC21. The resulting product will be subjected to restriction digestion with *Eag*I and *Nco*I, and cloned into pCO14K. This recombinant molecule contains regions C1, C2, C3, C4 and C5, thus deleting regions V1, V2, V3, and V4. The nucleotide sequence of the open reading frame is shown in FIG. 9 and SEQ ID NO 32, and the predicted polypeptide sequence derived from this nucleotide sequence is shown in FIG. 9 and SEQ ID NO 27. Analysis of the predicted amino acid sequence using the computer program SIGCLEAVE (part of the eGCG suite of programs hosted at www.angis.org.au) indicates that the first 49 amino acids will be cleaved to produce the mature polypeptide (FIG. 14; SEQ ID NO:39).

This construct can be transformed into any transformation competent strain of *N. meningitidis*.

EXAMPLE 9

Purification of over expressed NhhA polypeptides

Recombinant NhhA polypeptide as described in the previous Examples may be isolated by the following procedure. Bacteria are grown overnight (12-14 hours) at 37° C in an atmosphere of 5% CO_2. (In this example, media was BHI agar supplemented with Leventhal's base. Other growth media are well known to those skilled in the art). Bacteria from ten 25 mL agar plates were collected and suspended in 25 mL 10mM Tris adjusted to ph 8.0 with HCl. An equal volume of 10mM Tris (pH 8.0) containing 2% sarkosyl was added and the mixture mixed gently for 1 hour at 4° C. This was centrifuged at 100,000 × g for seventy minutes at 20° C and the supernatant discarded. The pellet was resuspended in 25 mL 10 mM Tris (pH 8.0) containing 1% sarkosyl by passing through a 25 gauge needle. This was centrifuged at 100,000 × g for seventy minutes at 20° C and the supernatant discarded. The pellet was resuspended in 10mL 10mM Tris (pH 8.0) by passing through a 25 gauge needle. This fraction contains the sarkosyl insoluble components of the cell, and is enriched for outer membrane proteins. (An additional step may be incorporated to remove residual sarkosyl detergent, whereby the protein solution is dialysed for four cycles of 4-8 hours against 100-1000 volumes of, for example, 10 mM Tris.Cl pH 8.0 or PBS (phosphate buffered saline) at 4 °C

Having determined the concentration of protein in the suspension by absorbance at wavelength of 280 nm, or by using a BCA kit (Pierce), approximately 1 mL of solution containing 10 mg of protein in a solution containing 1% SDS (sodium lauryl sulphate), 2% β-mercaptoethanol was separated on 1.5 mm thick 6% SDS-PAGE in the BioRad mini-protean II apparatus. The high molecular weight NhhA was eluted from the gel using the BioRad "mini Whole gel Eluter". Approximately 10% of each eluted fraction was checked by SDS-PAGE separation followed by Coomassie staining. Fractions containing NhhA essentially free of other proteins were pooled. This procedure was carried out to isolate over expressed mature NhhA as described in Example 2 (SEQ ID NO: 1), over expressed *Bgl*II deletion mature NhhA as described in Example 4 (SEQ ID NO: 23) and over expressed NhhA deletion mutant as described in Example 6 (SEQ ID NO: 25). Isolated protein is shown in FIG. 12.

EXAMPLE 10

Immunogenicity of purified NhhA deletion mutant polypeptides.

Mice were inoculated with purified wild-type NhhA polypeptides and deletion mutants as described in the previous Examples.. In one group, each Balb/C mouse was inoculated subcutaneously with approximately 130 μg PMC21 NhhA with MPL + TDM ™ adjuvant (obtained fromSigma-Aldrich) on day 0, 115 μg on day 14. In a second group, each mouse was inoculated with approximately 120 μg protein with MPL + TDM ™ adjuvant (obtained fromSigma-Aldrich) at day 0 and 190 μg at day 14. In a third group, each mouse was inoculated with approximately 260 μg protein with MPL + TDM ™ adjuvant (obtained fromSigma-Aldrich) at day 0 and 1240 μg at day 14. Blood samples were taken at day 21 and serum was extracted. These sera were tested for the presence of antibodies recognising full length PMC 21 NhhA by Western immunoblot (FIG. 13). OMC preparations (5 mg) of P6 (overexpresses PMC21 NhhA) and Strain 2A (NhhA expression abolished) were separated by 6% SDS-PAGE using the BioRad Mini Protean II electrophoresis apparatus. The proteins were transferred to nitrocellulose electrophoretically, and the filter was cut into 3 mm strips then blocked in 5% skim milk in PBS. Mouse sera was diluted to 1:1000 and 1:10000 in 5% skim milk powder and icnubated with the nitrocellulose strips. Antibody binding was detected using alkaline-phosphatase conjugated anti-mouse IgG (Sigma) before colorimetric detection with NBT/BCIP (Sigma). As can be seen from FIG. 13, it is possible to elicit an immune response against the full length mature PMC21 NhhA polypeptide by inoculation with NhhA deletion mutants or with full length mature NhhA polypeptides.

EXAMPLE 11

Expression of deletion mutant polypeptide in E. coli

In addition to expression of the mutant polypeptides of the invention in *N. meningitidis*, they may also be expressed in *E. coli* bacteria. Any of the recombinant *nhhA* deletion mutants of Examples 4-8 may be used as template for PCR amplifcation. Oligonucleotide primers used may be as described in International Publication WO99/31132 (such as SEQ ID NO 24 and SEQ ID NO 25 of that document). The amplification product may be restriction digested with *BamHI*/*Hind*III enzymes and ligated with *BamHI*/*Hind*III restriction digested plasmid pMALC2 (New England BioLabs), and the resultant plasmid transformed into competent *E. coli* strain DH5α. The resulting strain can be induced to express high levels of recombinant protein using conditions recommended by the manufacturer of pMALC2. The resulting recombinant protein is a fusion of maltose binding protein and the deletion mutant NhhA polypeptide of the invention.. This may be semi-purified by separation on SDS-PAGE followed by electroelution using the Mini-Gel Electro-eluter (BioRad) according to manufacturers instructions. The semi-purified fusion protein may then be dialysed against PBS, before digestion with the protease enzyme Factor Xa. to cleave the maltose binding protein moiety from the recombinant NhhA protein. The recombinant NhhA protein may be purifed by standard methods , as for example described by R. K. Scopes, Protein Purification (Springer-Verlag, New York, NY USA, 1993).

Throughout the specification the aim has been to describe the preferred embodiments of the invention without limiting the invention to any one embodiment or specific collection of features. It will therefore be appreciated by those of skill in the art that, in light of the instant disclosure, various modifications and changes can be made in the particular embodiments exemplified without departing from the scope of the present invention.

All computer programs, algorithms, patent and scientific literature referred to herein is incorporated herein by reference.

TABLE 1

	C1	V1	C2	V2	C3	V3	C4	V4	C5
Consensus SEQ ID NO: 11	1-50	51-108	109-120	121-134	135-198	199-220	221-239	240-248	249-604
PMC21 SEQ ID NO: 1	1-50	51-108	109-120	121-124	125-188	189-210	211-229	230-236	237-591
H41 SEQ ID NO: 2	1-50	51-102	103-114	115-124	125-188	189-210	211-229	230-236	237-591
P20 SEQ ID NO: 3	1-50	51-105	106-117	118-121	122-185	186-205	206-224	225-234	235-589
EG327 SEQ ID NO: 4	1-50	51-104	105-116	117-126	127-190	191-212	213-231	232-238	239-594
EG329 SEQ ID NO: 5	1-50	51-108	109-120	121-124	125-188	189-210	211-229	230-236	237-591
H38 SEQ ID NO: 6	1-50	51-105	106-117	118-131	132-195	196-217	218-236	237-243	244-599
H15 SEQ ID NO: 7	1-50	51-104	105-116	117-130	131-194	195-216	217-235	236-242	243-598
BZ10 SEQ ID NO: 8	1-50	51-104	105-116	117-130	131-194	195-216	217-235	236-242	243-598
BZ198 SEQ ID NO: 9	1-50	51-104	105-116	117-126	127-190	191-212	213-231	232-238	239-594
Z2491 SEQ ID NO: 10	1-50	51-102	103-114	115-124	125-188	189-208	209-227	228-236	237-592

TABLE 2

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

CLAIMS

1. An isolated protein comprising twelve or more contiguous conserved amino acids of an NhhA polypeptide, wherein said isolated protein is not a wild-type NhhA polypeptide.

2. The isolated protein of Claim 1 which is capable of eliciting an immune response.

3. The isolated protein of Claim 2, wherein the immune response is less strain-specific than that elicited by a corresponding said NhhA polypeptide.

4. The isolated protein of Claim 3, wherein said immune response provides protection against one or more strains of *N. meningitidis*

5. The isolated protein of Claim 3, wherein said immune response provides protection against a plurality of strains of *N. meningitidis*

6. The isolated protein of Claim 1 comprising twenty or more contiguous conserved amino acids.

7. The isolated protein of Claim 6 comprising fifty or more contiguous conserved amino acids.

8. The isolated protein of Claim 7 comprising one hundred or more contiguous conserved amino acids.

9. The isolated protein of Claim 1, wherein the NhhA polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO: 1; SEQ ID NO: 2; SEQ ID NO: 3; SEQ ID NO: 4; SEQ ID NO: 5; SEQ ID NO: 6; SEQ ID NO: 7; SEQ ID NO: 8; SEQ ID NO: 9; and SEQ ID NO: 10.

10. An isolated protein comprising an amino acid sequence selected from the group consisting of:

(i) residues 1 to 50 of SEQ ID NO:11;

(ii) residues 109 to120 of SEQ ID NO:11;

(iii) residues 135 to198 of SEQ ID NO:11;

(iv) residues 221 to 239 of SEQ ID NO:11; and

(v) residues 249 to 604 of SEQ ID NO:11.

wherein said isolated protein is not a wild type NhhA polypeptide.

11. The isolated protein of Claim 10, wherein the isolated protein has have an amino acid sequence selected from the group consisting of:

(i) residues 1 to 50 of SEQ ID NO:1;
(ii) residues 1 to 50 of SEQ ID NO:2;
(iii) residues 1 to 50 of SEQ ID NO:3;
(vii) residues 1 to 50 of SEQ ID NO:4;
(viii) residues 1 to 50 of SEQ ID NO:5;
(ix) residues 1 to 50 of SEQ ID NO:6;
(x) residues 1 to 50 of SEQ ID NO:7;
(xi) residues 1 to 50 of SEQ ID NO:8;
(xii) residues 1 to 50 of SEQ ID NO:9;
(xiii) residues 1 to 50 of SEQ ID NO:10;
(xiv) residues 125 to 188 of SEQ ID NO:1;
(xv) residues 125 to 188 of SEQ ID NO:2;
(xvi) residues 122 to 185 of SEQ ID NO:3;
(xvii) residues 127 to 190 of SEQ ID NO: 4;
(xviii) residues 125 to 188 of SEQ ID NO:5;
(xix) residues 132 to 195 of SEQ ID NO:6;
(xx) residues 131 to 194 of SEQ ID NO:7;
(xxi) residues 131 to 194 of SEQ ID NO: 8;
(xxii) residues 127 to 190 of SEQ ID NO:9;
(xxiii) residues 125 to 188 of SEQ ID NO:10;
(xxiv) residues 211 to 229 of SEQ ID NO:1;
(xxv) residues 206 to 224 of SEQ ID NO:3;
(xxvi) residues 237 to 591 of SEQ ID NO:1;
(xxvii) residues 237 to 592 of SEQ ID NO:2;
(xxviii)residues 235 to 589 of SEQ ID NO:3;
(xxix) residues 239 to 594 of SEQ ID NO:4;
(xxx) residues 237 to 591 of SEQ ID NO:5;
(xxxi) residues 244 to 599 of SEQ ID NO:6;
(xxxii) residues 243 to 598 of SEQ ID NO:7;
(xxxiii)residues 243 to 598 of SEQ ID NO:8.
(xxxiv)residues 239 to 594 of SEQ ID NO:9; and
(xxxv) residues 237 to 592 of SEQ ID NO:10.

11. The isolated protein of Claim 10 further comprising one or more variable (V) region amino acids of an NhhA polypeptide.

12. The isolated protein of Claim 11 having an amino acid sequence selected from the group consisting of: SEQ ID NO:23; SEQ ID NO:24, SEQ ID NO:25; SEQ ID NO:26; SEQ ID NO:27; SEQ ID NO: 33; SEQ ID NO: 34 SEQ ID NO: 35; SEQ ID NO: 36; SEQ ID NO: 37; SEQ ID NO: 38; and SEQ ID NO: 39.

13. An allelic variant of the isolated protein of Claim 10.

14. A fragment or derivative of the isolated protein of Claim 10.

15. The fragment of Claim 13 which is immunogenic.

16. A pharmaceutical composition comprising one or more isolated proteins according to any one of Claims 1-15.

17. The pharmaceutical composition of Claim 16 which is a vaccine.

18. An isolated nucleic acid encoding the isolated protein of any one of Claims 1-15.

19. The isolated nucleic acid of Claim 18 which has a nucleotide sequence selected from the group consisting of:

(i) residues 1 to 150 of SEQ ID NO:22;

(ii) residues 325 to 361 of SEQ ID NO:22;

(iii) residues 403 to595 of SEQ ID NO:22;

(iv) residues 661 to 717 of SEQ ID NO:22; and

(v) residues 745 to 1815 of SEQ ID NO:22.

20. The isolated nucleic acid of Claim 19 which has a nucleotide sequence selected from the group consisting of SEQ ID NO:28; SEQ ID NO:29: SEQ ID NO:30; SEQ ID NO:31 and SEQ ID NO:32.

21. An expression vector which includes the isolated nucleic acid of Claim 19 or Claim 20.

22. A host cell transformed with the expression vector of Claim 21.

23. The host cell of Claim 22 which is a bacterium.

24. The host cell of Claim 23 which is *Neisseria meningitidis*.

	1				50
EG327	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAS
BZ198	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
BZ10	MNKISRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
H15	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLFATVQAN
EG329	MNEILRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAS
PMC21	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAS
H38	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
P20	MNKIYRIIWN	SALNAWVVVS	ELTRNHTKRA	SATVATAVLA	TLLSATVQAN
Z2491	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
H41	MNKIYRIIWN	SALNAWVAVS	ELTRNHTKRA	SATVKTAVLA	TLLFATVQAN
Consensus	MN-I-RIIWN	SALNAWV-VS	ELTRNHTKRA	SATV-TAVLA	TLL-ATVQA-
			C1		

	51				100
EG327	TTDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.VTED	SNWGVYFDKK
BZ198	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
BZ10	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
H15	ATDDD...DL	YLEPVQRTAV	VLSFRSDKEG	TGEKE.GTED	SNWAVYFDEK
EG329	ANNEEQEEDL	YLDPVLRTVA	VLIVNSDKEG	TGEKEKVEEN	SDWAVYFNEK
PMC21	ANNEEQEEDL	YLDPVQRTVA	VLIVNSDKEG	TGEKEKVEEN	SDWAVYFNEK
H38	ATDED..EEE	ELEPVVRSAL	VLQFMIDKEG	NGENE.STGN	IGWSIYYDNH
P20	ATDTD..EDE	ELESVARSAL	VLQFMIDKEG	NGEIESTGDI	GWSIYYDDHN
Z2491	ATDED..EEE	ELESVQR.SV	VGSIQASMEG	SGELET...I	SLSMTNDSKE
H41	ATDED..EEE	ELESVQR.SV	VGSIQASMEG	SVELET...I	SLSMTNDSKE
Consensus	----------	-L--V-R---	V-------EG	--E-E-----	----------
			V1		

	101				150
EG327	GVLTAGTITL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTDLTSVG
BZ198	RVLKAGAITL	KAGDNLKIKQ	NTNENTNDSS	....FTYSLK	KDLTDLTSVE
BZ10	RVLKAGAITL	KAGDNLKIKQ	NTNENTNENT	NDSSFTYSLK	KDLTDLTSVE
H15	RVLKAGAITL	KAGDNLKIKQ	NTNENTNENT	NDSSFTYSLK	KDLTDLTSVE
EG329	GVLTAREITL	KAGDNLKIKQ	NG...TN...	....FTYSLK	KDLTDLTSVG
PMC21	GVLTAREITL	KAGDNLKIKQ	NG...TN...	....FTYSLK	KDLTDLTSVG
H38	NTLHGATVTL	KAGDNLKIKQ	NTNKNTNENT	NDSSFTYSLK	KDLTDLTSVE
P20	TLHG.ATVTL	KAGDNLKIKQ	SGKD......	....FTYSLK	KELKDLTSVE
Z2491	FVDPYIVVTL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTGLINVE
H41	FVDPYIVVTL	KAGDNLKIKQ	NTNENTNASS	....FTYSLK	KDLTGLINVE
Consensus	--------TL	KAGDNLKIKQ	----------	----FTYSLK	K-L--L--V-
	V1	C2	V2		C3

	151				200
EG327	TEKLSFSANS	NKVNITSDTK	GLNFAKKTAE	TNGDTTVHLN	GIGSTLTDTL
BZ198	TEKLSFGANG	NKVNITSDTK	GLNFAKETAG	TNGDPTVHLN	GIGSTLTDTL
BZ10	TEKLSFGANG	NKVNITSDTK	GLNFAKETAG	TNGDPTVHLN	GIGSTLTDTL
H15	TEKLSFGANG	NKVNITSDTK	GLNFAKETAG	TNGDPTVHLN	GIGSTLTDTL
EG329	TEKLSFSANG	NKVNITSDTK	GLNFAKETAG	TNGDTTVHLN	GIGSTLTDTL
PMC21	TEKLSFSANG	NKVNITSDTK	GLNFAKETAG	TNGDTTVHLN	GIGSTLTDTL
H38	TEKLSFGANG	NKVNITSDTK	GLNFAKETAG	TNGDTTVHLN	GIGSTLTDTL
P20	TEKLSFGANG	NKVNITSDTK	GLNFAKETAG	TNGDPTVHLN	GIGSTLTDTL
Z2491	TEKLSFGANG	KKVNIISDTK	GLNFAKETAG	TNGDTTVHLN	GIGSTLTDTL
H41	TEKLSFGANG	KKVNIISDTK	GLNFAKETAG	TNGDTTVHLN	GIGSTLTDML
Consensus	TEKLSF-AN-	-KVNI-SDTK	GLNFAK-TA-	TNGD-TVHLN	GIGSTLTD-L
			C3		

FIG. 1

```
           201                                                 250
EG327      LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
BZ198      LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
BZ10       LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
H15        LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
EG329      LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
PMC21      LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
H38        LNTGATTNVT NDNVTDDKKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
P20        AGSSASHVDA GNQST..HYT RAASIKDVLN AGWNIKGVKT GSTTGQSENV
Z2491      AGSSASHVDA GNQST..HYT RAASIKDVLN AGWNIKGVKT GSTTGQSENV
H41        LNTGATTNVT NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTAS..DNV
Consensus  ----A----- ----T----- RAAS-KDVLN AGWNIKGVK- G-T-----NV
                 V3                        C4                V4

           251                                                 300
EG327      DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
BZ198      DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
BZ10       DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
H15        DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
EG329      DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
PMC21      DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
H38        DFVHTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
P20        DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
Z2491      DFVRTYDTVE FLSADTKTTT VNVESKDNGK RTEVKIGAKT SVIKEKDGKL
H41        DFVRTYDTVE FLSADTKTTT VNVESKDNGK KTEVKIGAKT SVIKEKDGKL
Consensus  DFV-TYDTVE FLSADTKTTT VNVESKDNGK -TEVKIGAKT SVIKEKDGKL
                                           C5

           301                                                 350
EG327      VTGKDKGEND SSTDKGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
BZ198      VTGKGKDENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
BZ10       VTGKGKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
H15        VTGKGKDENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
EG329      VTGKDKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
PMC21      VTGKDKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
H38        VTGKGKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
P20        VTGKGKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
Z2491      VTGKGKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
H41        VTGKGKGENG SSTDEGEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
Consensus  VTGK-K-EN- SSTD-GEGLV TAKEVIDAVN KAGWRMKTTT ANGQTGQADK
                                           C5

           351                                                 400
EG327      FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS
BZ198      FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
BZ10       FETVTSGTKV TFASGNGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
H15        FETVTSGTKV TFASGNGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
EG329      FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS
PMC21      FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS
H38        FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
P20        FETVTSGTKV TFASGNGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
Z2491      FETVTSGTNV TFASGKGTTA TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS
H41        FETVTSGTKV TFASGNGTTA TVSKDDQGNI TVKYDVNVGD ALNVNQLQNS
Consensus  FETVTSGT-V TFASG-GTTA TVSKDDQGNI TV-YDVNVGD ALNVNQLQNS
                                           C5
```

FIG. 1 cont.

```
          401                                                  450
EG327     GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
BZ198     GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
BZ10      GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
H15       GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
EG329     GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
PMC21     GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
H38       GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
P20       GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
Z2491     GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EISRNGKNID
H41       GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EITRNGKNID
Consensus GWNLDSKAVA GSSGKVISGN VSPSKGKMDE TVNINAGNNI EI-RNGKNID
                                        C5

          451                                                  500
EG327     IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
BZ198     IATSMAPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDTNK PVRITNVAPG
BZ10      IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
H15       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
EG329     IATSMTPQFS SVSLGAGADA PTLSVDG.DA LNVGSKKDNK PVRITNVAPG
PMC21     IATSMTPQFS SVSLGAGADA PTLSVDG.DA LNVGSKKDNK PVRITNVAPG
H38       IATSMTPQFS SVSLGAGADA PTLSVDDKGA LNVGSKDANK PVRITNVAPG
P20       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
Z2491     IATSMAPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
H41       IATSMTPQFS SVSLGAGADA PTLSVDDEGA LNVGSKDANK PVRITNVAPG
Consensus IATSM-PQFS SVSLGAGADA PTLSVD---A LNVGSK--NK PVRITNVAPG
                                        C5

          501                                                  550
EG327     VKEGDVTNVA QLKGVAQNLN NHIDNVDGNA RAGIAQAIAT AGLVQAYLPG
BZ198     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
BZ10      VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLAQAYLPG
H15       VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLAQAYLPG
EG329     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
PMC21     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
H38       VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
P20       VKEGDVTNVA QLKGVAQNLN NRIDNVNGNA RAGIAQAIAT AGLAQAYLPG
Z2491     VKEGDVTNVA QLKGVAQNLN NRIDNVDGNA RAGIAQAIAT AGLVQAYLPG
H41       VKEGDVTNVA QLKGVAQNLN NRIDNVNGNA RAGIAQAIAT AGLVQAYLPG
Consensus VKEGDVTNVA QLKGVAQNLN N-IDNV-GNA RAGIAQAIAT AGL-QAYLPG
                                        C5

          551                                                  600
EG327     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
BZ198     KSMMAIGGDT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
BZ10      KSMMAIGGGT YRGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGTSASV
H15       KSMMAIGGGT YRGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGASASV
EG329     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
PMC21     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
H38       KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
P20       KSMMAIGGGT YLGEAGYAIG YSSISDTGNW VIKGTASGNS RGHFGTSASV
Z2491     KSMMAIGGGT YRGEAGYAIG YSSISDGGNW IIKGTASGNS RGHFGASASV
H41       KSMMAIGGGT YLGEAGYAIG YSSISAGGNW IIKGTASGNS RGHFGASASV
Consensus KSMMAIGG-T Y-GEAGYAIG YSSIS--GNW -IKGTASGNS RGHFG-SASV
                                        C5
```

FIG. 1 cont.

	601
EG327	GYQW.
BZ198	GYQW.
BZ10	GYQW.
H15	GYQW.
EG329	GYQW.
PMC21	GYQW.
H38	GYQW.
P20	GYQW.
Z2491	GYQW.
H41	GYQW.
Consensus	GYQW.
	C5

FIG. 1 cont.

```
           1                                                                          70
H15        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
BZ10       ATGAACAAAA TATCCCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
BZ198      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
P20        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGTCGTATCC GAGCTCACAC
H38        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT AGTCGTATCC GAGCTCACAC
Z2491      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
H41        ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
EG329      ATGAACGAAA TATTGCGCAT CATTTGGAAT AGCGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
PMC21      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT CGTTGTATCC GAGCTCACAC
EG327      ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT CGCCGTATCC GAGCTCACAC
Consensus  ATGAAC-AAA TAT--CGCAT CATTTGGAAT AG-GCCCTCA ATGC-TGGGT -G--GTATCC GAGCTCACAC
                                        C1

           71                                                                        140
H15        GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
BZ10       GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
BZ198      GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
P20        GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
H38        GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACACTGCTGT CCGCAACGGT
Z2491      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACGCTGTTGT TTGCAACGGT
H41        GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
EG329      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT
PMC21      GCAACCACAC CAAACGCGCC TCCGCAACCG TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT
EG327      GCAACCACAC CAAACGCGCC TCCGCAACCG TGGCGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT
Consensus  GCAACCACAC CAAACGCGCC TCCGCAACCG TG--GACCGC CGTATTGGCG AC-CTG-TGT --GCAACGGT
                                        C1

           141                                                                       210
H15        TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
BZ10       TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
BZ198      TCAGGCGAAT GCTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
P20        TCAGGCGAAT GCTACCGATA CCGAT..... .GAAGATGAA GAGTTAGAAT CCGTAGCACG CTCTGCTCTG
H38        TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAC CCGTAGTACG CTCTGCTCTG
Z2491      TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAT CCGTACAACG CTCTGTCGTA
H41        TCAGGCGAAT GCTACCGATG AAGAT..... .GAAGAAGAA GAGTTAGAAT CCGTACAACG CTCTG...TC
EG329      TCAGGCAAGT GCTAACAATG AAGAGCAAGA AGAAGATTTA TATTTAGACC CCGTGCTACG CACTGTTGCC
PMC21      TCAGGCAAGT GCTAACAATG AAGAGCAAGA AGAAGATTTA TATTTAGACC CCGTACAACG CACTGTTGCC
EG327      TCAGGCGAGT ACTACCGATG ACGAC..... ....GATTTA TATTTAGAAC CCGTACAACG CACTGCTGTC
Consensus  TCAGGC-A-T -CTA-C-AT- --GA------ ----GA---A -A-TTAGA-- CCGT---ACG C-CTG-----
             C1                                  V1

           211                                                                       280
H15        GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
BZ10       GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
BZ198      GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGGTAC AGAAGA...T TCAAATTGGG
P20        GTGTTGCAAT TCATGATCGA TAAAGAAGGC AATGGAGAAA TCGAATCTAC AGGAGA...T ATAGGTTGGA
H38        GTGTTGCAAT TCATGATCGA TAAAGAAGGC AATGGAGAAA ACGAATCTAC AGGAAA...T ATAGGTTGGA
Z2491      GGG..AGCAT TCAAG.CCAG TATGGAAGGC AGCGGCGAAT TGGAAACGAT ATCAT....T ATCAATGACT
H41        GTAGGGAGCA TTCAAGCCAG TATGGAAGGC AGCGTCGAAT TGGAAACGAT A......... TCATTATCAA
EG329      GTGTTGATAG TCAATTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAAAAGT AGAAGAAAAT TCAGATTGGG
PMC21      GTGTTGATAG TCAATTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAAAAGT AGAAGAAAAT TCAGATTGGG
EG327      GTGTTGAGCT TCCGTTCCGA TAAAGAAGGC ACGGGAGAAA AAGAAGTTAC AGAAGA...T TCAAATTGGG
Consensus  G--------- T------C-- TA--GAAGGC A--G--GAA- --GAA----- A--------- ----------
                                             V1
```

FIG. 2

```
           281                                                                350
       H15 CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
      BZ10 CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
     BZ198 CAGTATATTT CGACGAGAAA AGAGTACTAA AAGCCGGAGC AATCACCCTC AAAGCCGGCG ACAACCTGAA
       P20 GTATATATTA CGACGATCAC AACACTCTAC ACGGCGCAAC CGTTACCCTC AAAGCCGGCG ACAACCTGAA
       H38 GTATATATTA CGACAATCAC AACACTCTAC ACGGCGCAAC CGTTACCCTC AAAGCCGGCG ACAACCTGAA
     Z2491 AACGACAGCA AGGAATTTGT AGACCCATAC ATAGTA.... .GTTACCCTC AAAGCCGGCG ACAACCTGAA
       H41 TGACTAACGA CAGCAAGGAA TTTGTAGACC CATACATAGT AGTTACCCTC AAAGCCGGCG ACAACCTGAA
     EG329 CAGTATATTT CAACGAGAAA GGAGTACTAA CAGCCAGAGA AATCACCCTC AAAGCCGGCG ACAACCTGAA
     PMC21 CAGTATATTT CAACGAGAAA GGAGTACTAA CAGCCAGAGA AATCACCCTC AAAGCCGGCG ACAACCTGAA
     EG327 GAGTATATTT CGACAAGAAA GGAGTACTAA CAGCCGGAAC AATCACCCTC AAAGCCGGCG ACAACCTGAA
 Consensus ------A--- ---------- ---------- ---------- --T-ACCCTC AAAGCCGGCG ACAACCTGAA
                                          V1                      C2

           351                                                                420
       H15 AATCAAACAA AACACCAATG AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCCCTGAAA
      BZ10 AATCAAACAA AACACCAATG AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCCCTGAAA
     BZ198 AATCAAACAA AACACCAATG AAAACACC.. .......... AATGACAGTA GCTTCACCTA CTCCCTGAAA
       P20 AATCAAACAA AGCGGCAAAG A......... .......... .......... .CTTCACCTA CTCGCTGAAA
       H38 AATCAAACAA AACACCAATA AAAACACCAA TGAAAACACC AATGACAGTA GCTTCACCTA CTCGCTGAAA
     Z2491 AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
       H41 AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
     EG329 AATCAAACAA AAC....... .......G.. .......... ....GCACAA ACTTCACCTA CTCGCTGAAA
     PMC21 AATCAAACAA AAC....... .......G.. .......... ....GCACAA ACTTCACCTA CTCGCTGAAA
     EG327 AATCAAACAA AACACCAATG AAAACACC.. .......... AATGCCAGTA GCTTCACCTA CTCGCTGAAA
 Consensus AATCAAACAA A-C------- ---------- ---------- ---------- -CTTCACCTA CTC-CTGAAA
             C2                   V2                                           C3

           421                                                                490
       H15 AAAGACCTCA CAGATCTGAC CAGTGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGT AATAAAGTCA
      BZ10 AAAGACCTCA CAGATCTGAC CAGTGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGT AATAAAGTCA
     BZ198 AAAGACCTCA CAGATCTGAC CAGTGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGT AATAAAGTCA
       P20 AAAGAGCTGA AAGACCTGAC CAGTGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGT AATAAAGTCA
       H38 AAAGACCTCA CAGATCTGAC CAGTGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGC AATAAAGTCA
     Z2491 AAAGACCTCA CAGGCCTGAT CAATGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGC AAGAAAGTCA
       H41 AAAGACCTCA CAGGCCTGAT CAATGTTGAA ACTGAAAAAT TATCGTTTGG CGCAAACGGC AAGAAAGTCA
     EG329 AAAGACCTCA CAGATCTGAC CAGTGTTGGA ACTGAAAAAT TATCGTTTAG CGCAAACGGC AATAAAGTCA
     PMC21 AAAGACCTCA CAGATCTGAC CAGTGTTGGA ACTGAAAAAT TATCGTTTAG CGCAAACGGC AATAAAGTCA
     EG327 AAAGACCTCA CAGATCTGAC CAGTGTTGGA ACTGAAAAAT TATCGTTTAG CGCAAACAGC AATAAAGTCA
 Consensus AAAGA-CT-A -AG--CTGA- CA-TGTTG-A ACTGAAAAAT TATCGTTT-G CGCAAAC-G- AA-AAAGTCA
                                          C3

           491                                                                560
       H15 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACCCCACGGT
      BZ10 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACCCCACGGT
     BZ198 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACCCCACGGT
       P20 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACCCCACGGT
       H38 ACATCACAAG CGACACCAAA GGCTTGAATT TCGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
     Z2491 ACATCATAAG CGACACCAAA GGCTTGAATT TCGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
       H41 ACATCATAAG CGACACCAAA GGCTTGAATT TCGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
     EG329 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
     PMC21 ACATCACAAG CGACACCAAA GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
     EG327 ACATCACAAG CGACACCAAA GGCTTGAATT TCGCGAAAAA AACGGCTGAG ACCAACGGCG ACACCACGGT
 Consensus ACATCA-AAG CGACACCAAA GGCTTGAATT T-GCGAAA-A AACGGCTG-G AC-AACGGCG AC-CCACGGT
                                          C3

           561                                                                630
       H15 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
      BZ10 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
     BZ198 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
       P20 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTT GCGGGTTCTT CTGCTTCTCA CGTTGATGCG
       H38 TCATCTGAAC GGTATTGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
     Z2491 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTT GCGGGTTCTT CTGCTTCTCA CGTTGATGCG
       H41 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATATGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
     EG329 TCATCTGAAC GGTATTGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
     PMC21 TCATCTGAAC GGTATTGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
     EG327 TCATCTGAAC GGTATCGGTT CGACTTTGAC CGATACGCTG CTGAATACCG GAGCGACCAC AAACGTAACC
 Consensus TCATCTGAAC GGTAT-GGTT CGACTTTGAC CGATA-GCT- --G--T-C-- --GC--C--- ----G---C-
                   C3                                         V3
```

FIG. 2 cont.

```
           631                                                                  700
H15        AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
BZ10       AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
BZ198      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
P20        GGTAACCAAA GTACACATTA C......ACT CGTGCAGCAA GTATTAAGGA TGTGTTGAAT GCGGGTTGGA
H38        AACGACAACG TTACCGATGA CAAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
Z2491      GGTAACCAAA GTACACATTA C......ACT CGTGCAGCAA GTATTAAGGA TGTGTTGAAT GCGGGTTGGA
H41        AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
EG329      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCTGGCTGGA
PMC21      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCTGGCTGGA
EG327      AACGACAACG TTACCGATGA CGAGAAAAAA CGTGCGGCAA GCGTTAAAGA CGTATTAAAC GCAGGCTGGA
Consensus  ----AC-A-- -TAC--AT-A C------A-- CGTGC-GCAA G--TTAA-GA -GT-TT-AA- GC-GG-TGGA
                     V3                                            C4

           701                                                                  770
H15        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
BZ10       ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTC GATTTCGTCC GCACTTACGA
BZ198      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
P20        ATATTAAGGG TGTTAAAACT GGCTCAACAA CTGGTCAATC AGAAAATGTC GATTTCGTCC GCACTTACGA
H38        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC ACACTTACGA
Z2491      ATATTAAGGG TGTTAAAACT GGCTCAACAA CTGGTCAATC AGAAAATGTC GATTTCGTCC GCACTTACGA
H41        ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
EG329      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
PMC21      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
EG327      ACATTAAAGG CGTTAAACCC GGTACAACAG CT......TC CGATAACGTT GATTTCGTCC GCACTTACGA
Consensus  A-ATTAA-GG -GTTAAA-C- GG--CAACA- CT------TC -GA-AA-GT- GATTTCGTCC -GACTTACGA
               C4                V4                                C5

           771                                                                  840
H15        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
BZ10       CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
BZ198      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
P20        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
H38        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
Z2491      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
H41        CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
EG329      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
PMC21      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
EG327      CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
Consensus  CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT GTTAATGTGG AAAGCAAAGA CAACGGCAAG
                                                C5

           841                                                                  910
H15        AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
BZ10       AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
BZ198      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
P20        AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
H38        AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
Z2491      AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
H41        AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
EG329      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
PMC21      AAAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
EG327      AGAACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTATCA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
Consensus  A-AACCGAAG TTAAAATCGG TGCGAAGACT TCTGTTAT-A AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
                                                C5

           911                                                                  980
H15        AAGGCAAAGA CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
BZ10       AAGGCAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
BZ198      AAGGCAAAGA CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
P20        AAGGCAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
H38        AAGGCAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
Z2491      AAGGCAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
H41        AAGGCAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
EG329      AAGACAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
PMC21      AAGACAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
EG327      AAGACAAAGG CGAGAATGAT TCTTCTACAG ACAAAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
Consensus  AAG-CAAAG- CGAGAATG-T TCTTCTACAG AC-AAGGCGA AGGCTTAGTG ACTGCAAAAG AAGTGATTGA
                                                C5
```

FIG. 2 cont.

```
           981                                                                   1050
      H15  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
     BZ10  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
    BZ198  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
      P20  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
      H38  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
    Z2491  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
      H41  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
    EG329  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
    PMC21  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
    EG327  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
Consensus  TGCAGTAAAC AAGGCTGGTT GGAGAATGAA AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG
                                              C5

           1051                                                                  1120
      H15  TTTGAAACCG TTACATCAGG CACAAAAGTA ACCTTTGCTA GTGGTAATGG TACAACTGCG ACTGTAAGTA
     BZ10  TTTGAAACCG TTACATCAGG CACAAAAGTA ACCTTTGCTA GTGGTAATGG TACAACTGCG ACTGTAAGTA
    BZ198  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
      P20  TTTGAAACCG TTACATCAGG CACAAAAGTA ACCTTTGCTA GTGGTAATGG TACAACTGCG ACTGTAAGTA
      H38  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
    Z2491  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
      H41  TTTGAAACCG TTACATCAGG CACAAAAGTA ACCTTTGCTA GTGGTAATGG TACAACTGCG ACTGTAAGTA
    EG329  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
    PMC21  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
    EG327  TTTGAAACCG TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG ACTGTAAGTA
Consensus  TTTGAAACCG TTACATCAGG CACAAA-GTA ACCTTTGCTA GTGGTAA-GG TACAACTGCG ACTGTAAGTA
                                              C5

           1121                                                                  1190
      H15  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
     BZ10  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
    BZ198  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
      P20  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
      H38  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
    Z2491  AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
      H41  AAGATGATCA AGGCAACATC ACTGTTAAGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
    EG329  AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
    PMC21  AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
    EG327  AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
Consensus  AAGATGATCA AGGCAACATC ACTGTTA-GT ATGATGTAAA TGTCGGCGAT GCCCTAAACG TCAATCAGCT
                                              C5

           1191                                                                  1260
      H15  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
     BZ10  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    BZ198  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      P20  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      H38  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    Z2491  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
      H41  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    EG329  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    PMC21  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
    EG327  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
Consensus  GCAAAACAGC GGTTGGAATT TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
                                              C5

           1261                                                                  1330
      H15  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
     BZ10  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    BZ198  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
      P20  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
      H38  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    Z2491  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTAGCC
      H41  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    EG329  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    PMC21  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
    EG327  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTACCC
Consensus  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG CAACAACATC GAGATTA-CC
                                              C5
```

FIG. 2 cont.

```
          1331                                                                   1400
      H15 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
     BZ10 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
    BZ198 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGGCGCC GCAGTTTTCC AGCGTTTCGC TCGGTGCGGG
      P20 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
      H38 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    Z2491 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGGCGCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
      H41 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
    EG329 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    PMC21 GCAACGGTAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG
    EG327 GCAACGGCAA AAATATCGAC ATCGCCACTT CGATGACCCC GCAATTTTCC AGCGTTTCGC TCGGCGCGGG
Consensus GCAACGG-AA AAATATCGAC ATCGCCACTT CGATG-C-CC GCA-TTTTCC AGCGTTTCGC TCGG-GCGGG
                                          C5

          1401                                                                   1470
      H15 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
     BZ10 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    BZ198 GGCGGATGCG CCCACTTTGA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TACCAACAAA
      P20 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
      H38 GGCGGATGCG CCCACTTTGA GCGTGGATGA CAAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    Z2491 GGCAGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
      H41 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
    EG329 GGCGGATGCG CCCACTTTGA GCGTGGAT.. .GGGGACGCA TTGAATGTCG GCAGCAAGAA GGACAACAAA
    PMC21 GGCGGATGCG CCCACTTTGA GCGTGGAT.. .GGGGACGCA TTGAATGTCG GCAGCAAGAA GGACAACAAA
    EG327 GGCGGATGCG CCCACTTTAA GCGTGGATGA CGAGGGCGCG TTGAATGTCG GCAGCAAGGA TGCCAACAAA
Consensus GGC-GATGCG CCCACTTT-A GCGTGGAT-- ---GG-CGC- TTGAATGTCG GCAGCAAG-A ---CAACAAA
                                          C5

          1471                                                                   1540
      H15 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
     BZ10 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
    BZ198 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
      P20 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
      H38 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
    Z2491 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
      H41 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCG CAACTTAAAG
    EG329 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
    PMC21 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
    EG327 CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGCA CAACTTAAAG
Consensus CCCGTCCGCA TTACCAATGT CGCCCCGGGC GTTAAAGAGG GGGATGTTAC AAACGTCGC- CAACTTAAAG
                                          C5

          1541                                                                   1610
      H15 GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
     BZ10 GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
    BZ198 GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
      P20 GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGAA CGGCAACGCG CGCGCGGGTA TCGCCCAAGC
      H38 GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
    Z2491 GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
      H41 GTGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGAA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
    EG329 GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
    PMC21 GCGTGGCGCA AAACTTGAAC AACCGCATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
    EG327 GCGTGGCGCA AAACTTGAAC AACCACATCG ACAATGTGGA CGGCAACGCG CGTGCGGGCA TCGCCCAAGC
Consensus G-GTGGCGCA AAACTTGAAC AACC-CATCG ACAATGTG-A CGGCAACGCG CG-GCGGG-A TCGCCCAAGC
                                          C5

          1611                                                                   1680
      H15 GATTGCAACC GCAGGTTTGG CTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
     BZ10 GATTGCAACC GCAGGTTTGG CTCAGGCCTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
    BZ198 GATTGCAACC GCAGGTCTAG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGACACT
      P20 GATTGCAACC GCAGGTTTGG CTCAGGCCTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGTACT
      H38 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
    Z2491 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
      H41 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
    EG329 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
    PMC21 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TTTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
    EG327 GATTGCAACC GCAGGTCTGG TTCAGGCGTA TCTGCCCGGC AAGAGTATGA TGGCGATCGG CGGCGGCACT
Consensus GATTGCAACC GCAGGT-T-G -TCAGGC-TA T-TGCCCGGC AAGAGTATGA TGGCGATCGG CGGCG--ACT
                                          C5
```

FIG. 2 cont.

```
           1681                                                                          1750
       H15 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
      BZ10 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
     BZ198 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCAAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
       P20 TATCTCGGCG AAGCCGGTTA CGCCATCGGC TACTCGAGCA TTTCTGACAC TGGGAATTGG GTTATCAAGG
       H38 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
     Z2491 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
       H41 TATCTCGGCG AAGCCGGTTA TGCCATCGGC TACTCAAGCA TTTCCGCCGG CGGAAATTGG ATTATCAAAG
     EG329 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
     PMC21 TATCGCGGCG AAGCCGGTTA CGCCATCGGC TACTCCAGTA TTTCCGACGG CGGAAATTGG ATTATCAAAG
     EG327 TATCGCGGCG AAGCCGGTTA TGCCATCGGC TACTCAAGCA TTTCCGACGG CGGAAATTGG ATTATCAAAG
 Consensus TATC-CGGCG AAGCCGGTTA -GCCATCGGC TACTC-AG-A TTTC-G-C-- -GG-AATTGG -TTATCAA-G
                                                   C5

           1751                                                                1815
       H15 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
      BZ10 GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTACTTC CGCATCTGTC GGTTATCAGT GGTAA
     BZ198 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAAT GGTAA
       P20 GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTACTTC CGCATCTGTC GGTTATCAGT GGTAA
       H38 GCACGGCTTC CGGCAATTCG CGCGGTCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
     Z2491 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
       H41 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
     EG329 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
     PMC21 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
     EG327 GCACGGCTTC CGGCAATTCG CGCGGCCATT TCGGTGCTTC CGCATCTGTC GGTTATCAGT GGTAA
 Consensus GCACGGCTTC CGGCAATTCG CGCGG-CATT TCGGT-CTTC CGCATCTGTC GGTTATCA-T GGTAA
                                                   C5
```

FIG. 2 cont.

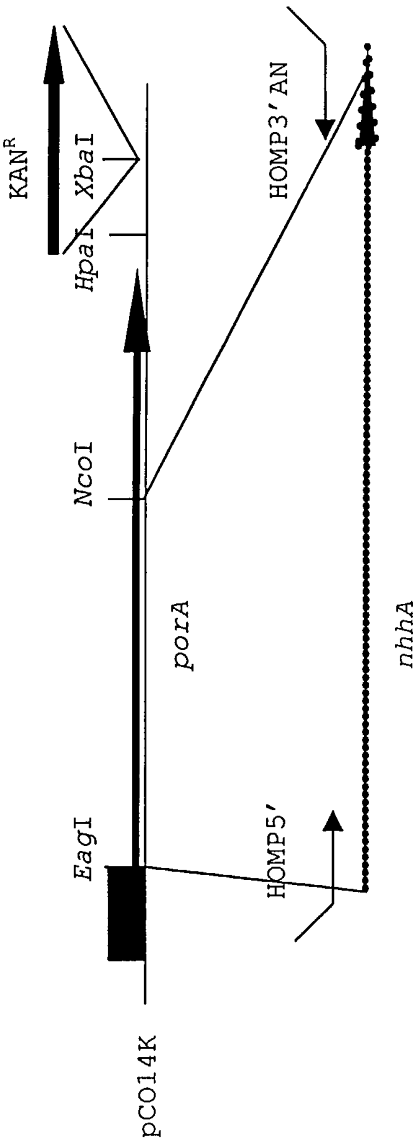

FIG. 3A

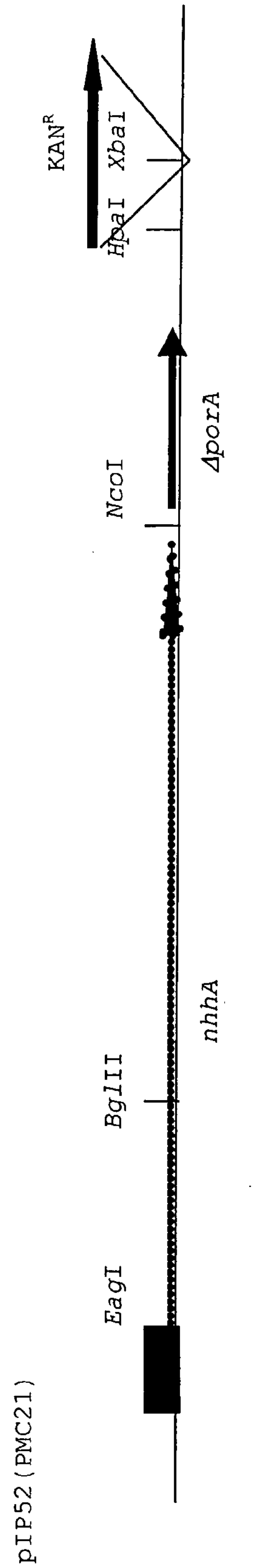

FIG. 3B

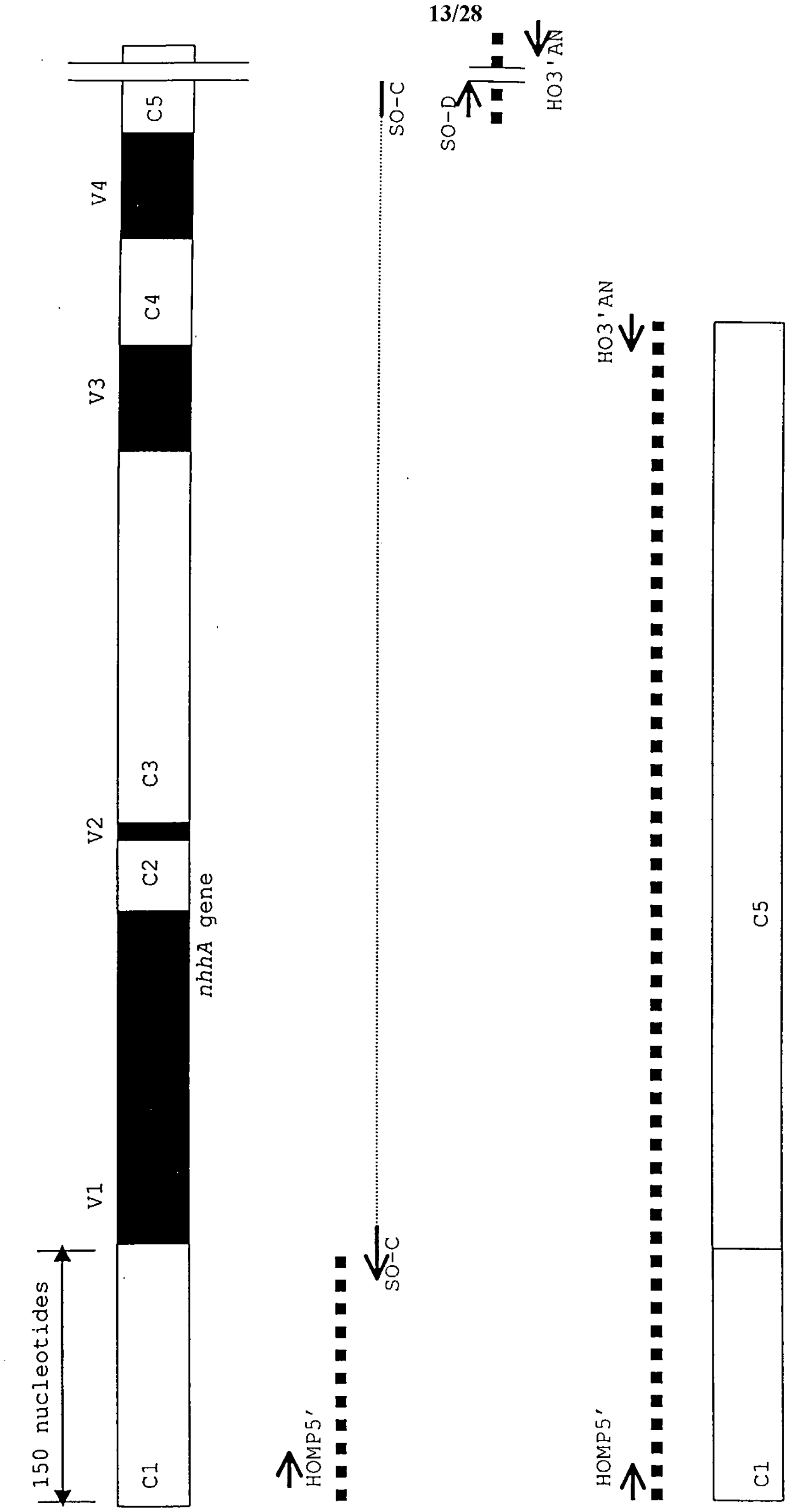
C5
V4
C4
V3
C3
V2
C2
nhhA gene
V1
150 nucleotides
C1
SO-C
SO-D
HO3'AN
SO-C
HOMP5'
HO3'AN
HOMP5'
C5
C1

FIG. 4A

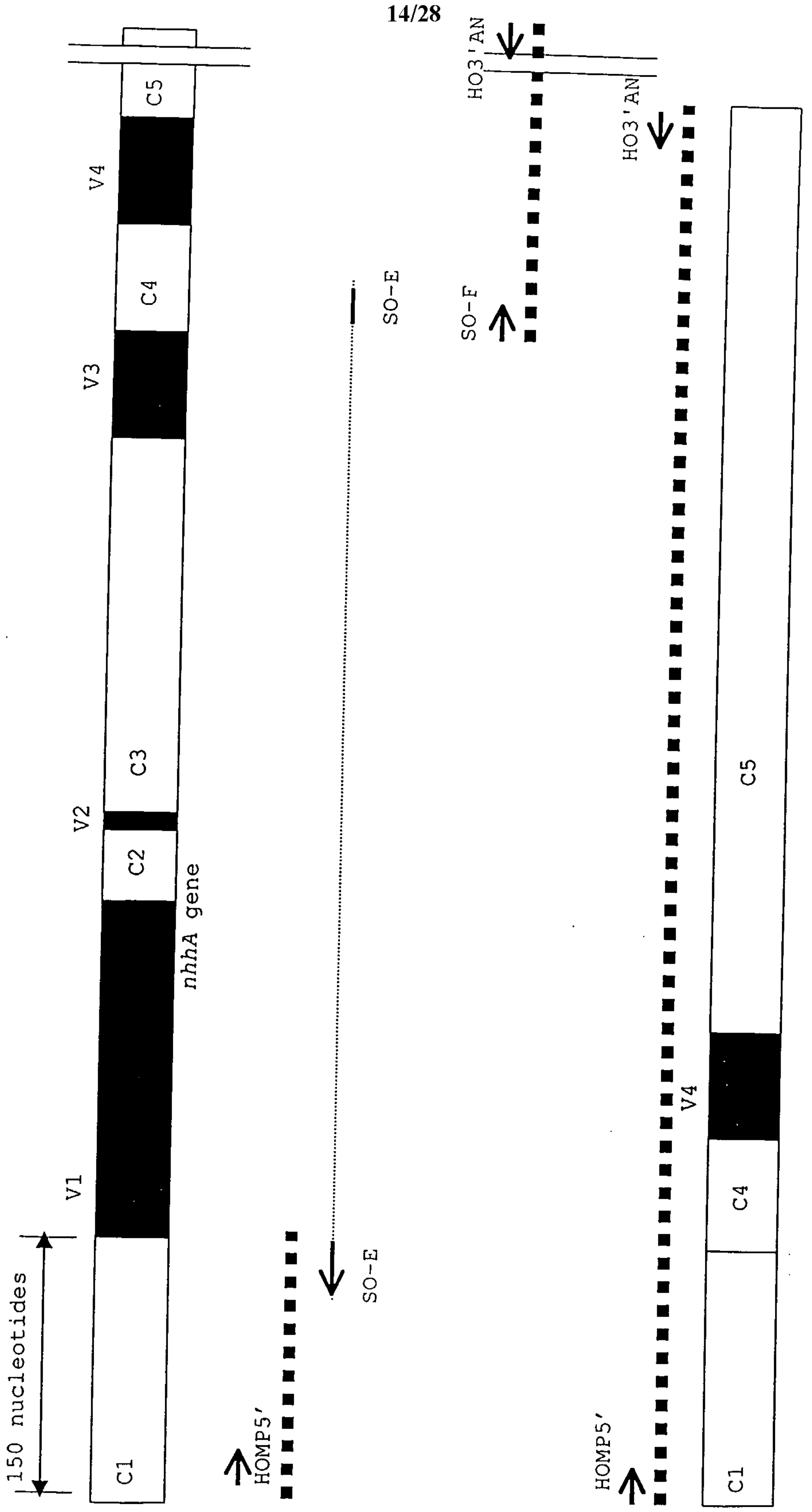
C5
V4
C4
V3
V2
C3
C2
nhhA gene
V1
150 nucleotides
C1
SO-E
SO-E
HOMP5'
HO3'AN
SO-F
HO3'AN
C5
V4
C4
C1
HOMP5'

FIG. 4B

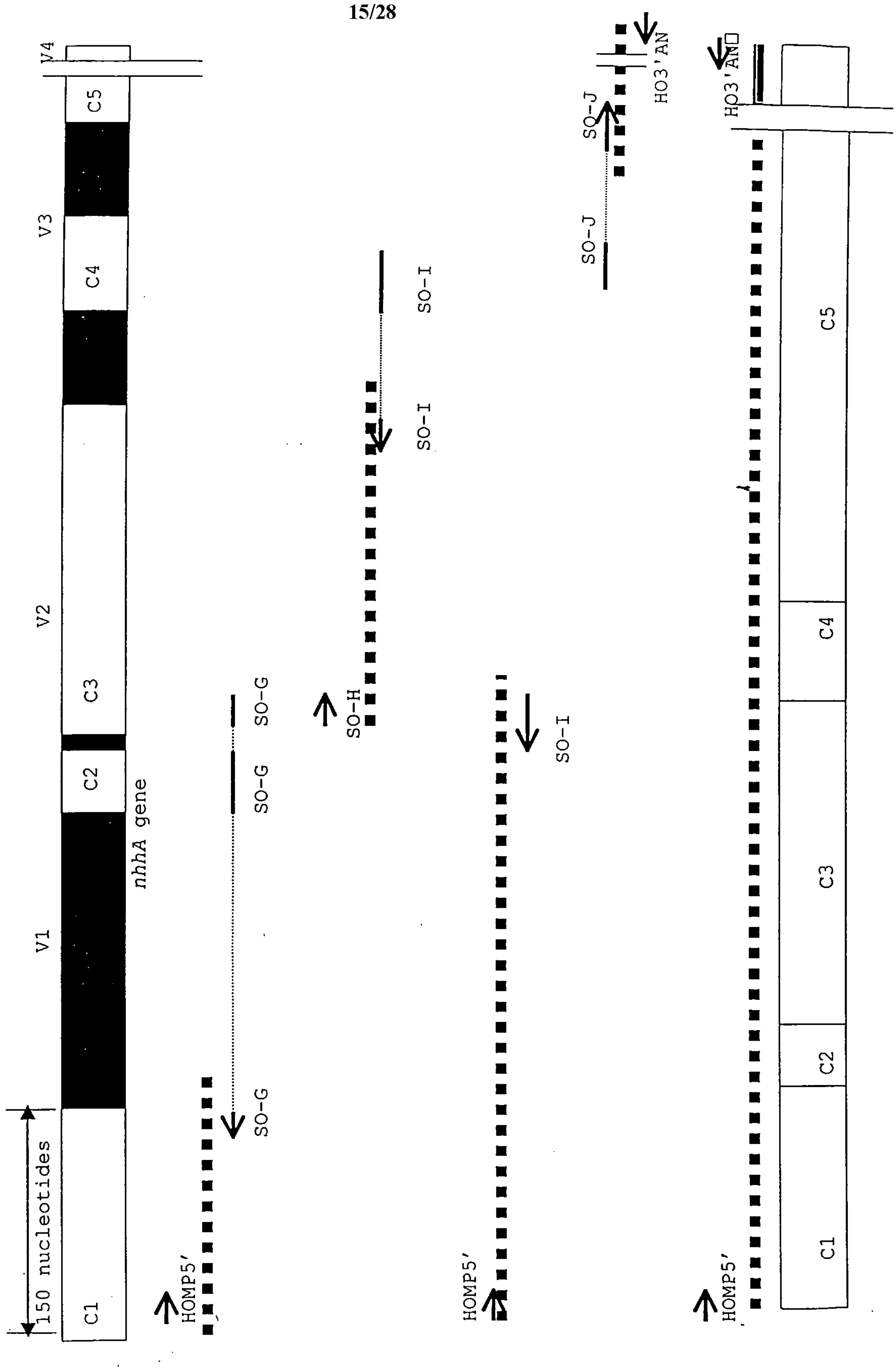
V4
C5
V3
C4
V2
C3
C2
V1
C1
150 nucleotides
nhhA gene
HOMP5'
SO-G
SO-G
SO-G
SO-H
SO-I
SO-I
SO-I
HOMP5'
SO-J
SO-J
HO3'AN
HOMP5'
C1
C2
C3
C4
C5

FIG. 4C

```
  1 MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51 ANNETDLTSV GTEKLSFSAN GNKVNITSDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDT LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV
401 RITNVAPGVK EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG
451 LVQAYLPGKS MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG
501 HFGASASVGY QW*
```

A

```
   1 ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51 CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101 TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151 GCTAACAATG AAACAGATCT GACCAGTGTT GGAACTGAAA AATTATCGTT
 201 TAGCGCAAAC GGCAATAAAG TCAACATCAC AAGCGACACC AAAGGCTTGA
 251 ATTTTGCGAA AGAAACGGCT GGGACGAACG GCGACACCAC GGTTCATCTG
 301 AACGGTATTG GTTCGACTTT GACCGATACG CTGCTGAATA CCGGAGCGAC
 351 CACAAACGTA ACCAACGACA ACGTTACCGA TGACGAGAAA AAACGTGCGG
 401 CAAGCGTTAA AGACGTATTA AACGCTGGCT GGAACATTAA AGGCGTTAAA
 451 CCCGGTACAA CAGCTTCCGA TAACGTTGAT TTCGTCCGCA CTTACGACAC
 501 AGTCGAGTTC TTGAGCGCAG ATACGAAAAC AACGACTGTT AATGTGGAAA
 551 GCAAAGACAA CGGCAAGAAA ACCGAAGTTA AAATCGGTGC GAAGACTTCT
 601 GTTATTAAAG AAAAAGACGG TAAGTTGGTT ACTGGTAAAG ACAAAGGCGA
 651 GAATGGTTCT TCTACAGACG AAGGCGAAGG CTTAGTGACT GCAAAAGAAG
 701 TGATTGATGC AGTAAACAAG GCTGGTTGGA GAATGAAAAC AACAACCGCT
 751 AATGGTCAAA CAGGTCAAGC TGACAAGTTT GAAACCGTTA CATCAGGCAC
 801 AAATGTAACC TTTGCTAGTG GTAAAGGTAC AACTGCGACT GTAAGTAAAG
 851 ATGATCAAGG CAACATCACT GTTATGTATG ATGTAAATGT CGGCGATGCC
 901 CTAAACGTCA ATCAGCTGCA AAACAGCGGT TGGAATTTGG ATTCCAAAGC
 951 GGTTGCAGGT TCTTCGGGCA AAGTCATCAG CGGCAATGTT TCGCCGAGCA
1001 AGGGAAAGAT GGATGAAACC GTCAACATTA ATGCCGGCAA CAACATCGAG
1051 ATTACCCGCA ACGGTAAAAA TATCGACATC GCCACTTCGA TGACCCCGCA
1101 GTTTTCCAGC GTTTCGCTCG GCGCGGGGGC GGATGCGCCC ACTTTGAGCG
1151 TGGATGGGGA CGCATTGAAT GTCGGCAGCA AGAAGGACAA CAAACCCGTC
1201 CGCATTACCA ATGTCGCCCC GGGCGTTAAA GAGGGGGATG TTACAAACGT
1251 CGCACAACTT AAAGGCGTGG CGCAAAACTT GAACAACCGC ATCGACAATG
1301 TGGACGGCAA CGCGCGTGCG GGCATCGCCC AAGCGATTGC AACCGCAGGT
1351 CTGGTTCAGG CGTATTTGCC CGGCAAGAGT ATGATGGCGA TCGGCGGCGG
1401 CACTTATCGC GGCGAAGCCG GTTACGCCAT CGGCTACTCC AGTATTTCCG
1451 ACGGCGGAAA TTGGATTATC AAAGGCACGG CTTCCGGCAA TTCGCGCGGC
1501 CATTTCGGTG CTTCCGCATC TGTCGGTTAT CAGTGGTAA
```

B

FIG. 5

```
  1 MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
 51 ATDETGLINV ETEKLSFGAN GKKVNIISDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDM LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKGKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTKVT FASGNGTTAT VSKDDQGNIT VKYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDDEGAL NVGSKDANKP
401 VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN RIDNVNGNAR AGIAQAIATA
451 GLVQAYLPGK SMMAIGGGTY LGEAGYAIGY SSISAGGNWI IKGTASGNSR
501 GHFGASASVG YQW*
```

A

```
   1 ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCCTGGGT
  51 CGCCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101 TGAAGACCGC CGTATTGGCG ACACTGTTGT TTGCAACGGT TCAGGCGAAT
 151 GCTACCGATG AAACAGGCCT GATCAATGTT GAAACTGAAA AATTATCGTT
 201 TGGCGCAAAC GGCAAGAAAG TCAACATCAT AAGCGACACC AAAGGCTTGA
 251 ATTTCGCGAA AGAAACGGCT GGGACGAACG GCGACACCAC GGTTCATCTG
 301 AACGGTATCG GTTCGACTTT GACCGATATG CTGCTGAATA CCGGAGCGAC
 351 CACAAACGTA ACCAACGACA ACGTTACCGA TGACGAGAAA AAACGTGCGG
 401 CAAGCGTTAA AGACGTATTA AACGCAGGCT GGAACATTAA AGGCGTTAAA
 451 CCCGGTACAA CAGCTTCCGA TAACGTTGAT TTCGTCCGCA CTTACGACAC
 501 AGTCGAGTTC TTGAGCGCAG ATACGAAAAC AACGACTGTT AATGTGGAAA
 551 GCAAAGACAA CGGCAAGAAA ACCGAAGTTA AAATCGGTGC GAAGACTTCT
 601 GTTATTAAAG AAAAAGACGG TAAGTTGGTT ACTGGTAAAG GCAAAGGCGA
 651 GAATGGTTCT TCTACAGACG AAGGCGAAGG CTTAGTGACT GCAAAAGAAG
 701 TGATTGATGC AGTAAACAAG GCTGGTTGGA GAATGAAAAC AACAACCGCT
 751 AATGGTCAAA CAGGTCAAGC TGACAAGTTT GAAACCGTTA CATCAGGCAC
 801 AAAAGTAACC TTTGCTAGTG GTAATGGTAC AACTGCGACT GTAAGTAAAG
 851 ATGATCAAGG CAACATCACT GTTAAGTATG ATGTAAATGT CGGCGATGCC
 901 CTAAACGTCA ATCAGCTGCA AAACAGCGGT TGGAATTTGG ATTCCAAAGC
 951 GGTTGCAGGT TCTTCGGGCA AAGTCATCAG CGGCAATGTT TCGCCGAGCA
1001 AGGGAAAGAT GGATGAAACC GTCAACATTA ATGCCGGCAA CAACATCGAG
1051 ATTACCCGCA ACGGCAAAAA TATCGACATC GCCACTTCGA TGACCCCGCA
1101 ATTTTCCAGC GTTTCGCTCG GCGCGGGGGC GGATGCGCCC ACTTTAAGCG
1151 TGGATGACGA GGGCGCGTTG AATGTCGGCA GCAAGGATGC CAACAAACCC
1201 GTCCGCATTA CCAATGTCGC CCCGGGCGTT AAAGAGGGGG ATGTTACAAA
1251 CGTCGCGCAA CTTAAAGGTG TGGCGCAAAA CTTGAACAAC CGCATCGACA
1301 ATGTGAACGG CAACGCGCGT GCGGGCATCG CCCAAGCGAT TGCAACCGCA
1351 GGTCTGGTTC AGGCGTATCT GCCCGGCAAG AGTATGATGG CGATCGGCGG
1401 CGGCACTTAT CTCGGCGAAG CCGGTTATGC CATCGGCTAC TCAAGCATTT
1451 CCGCCGGCGG AAATTGGATT ATCAAAGGCA CGGCTTCCGG CAATTCGCGC
1501 GGCCATTTCG GTGCTTCCGC ATCTGTCGGT TATCAGTGGT AA
```

B

FIG. 6

```
  1  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51  ANNVDFVRTY DTVEFLSADT KTTTVNVESK DNGKKTEVKI GAKTSVIKEK
101  DGKLVTGKDK GENGSSTDEG EGLVTAKEVI DAVNKAGWRM KTTTANGQTG
151  QADKFETVTS GTNVTFASGK GTTATVSKDD QGNITVMYDV NVGDALNVNQ
201  LQNSGWNLDS KAVAGSSGKV ISGNVSPSKG KMDETVNINA GNNIEITRNG
251  KNIDIATSMT PQFSSVSLGA GADAPTLSVD GDALNVGSKK DNKPVRITNV
301  APGVKEGDVT NVAQLKGVAQ NLNNRIDNVD GNARAGIAQA IATAGLVQAY
351  LPGKSMMAIG GGTYRGEAGY AIGYSSISDG GNWIIKGTAS GNSRGHFGAS
401  ASVGYQW*
```

A

```
   1  ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101  TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151  GCTAACAACG TTGATTTCGT CCGCACTTAC GACACAGTCG AGTTCTTGAG
 201  CGCAGATACG AAAACAACGA CTGTTAATGT GGAAAGCAAA GACAACGGCA
 251  AGAAAACCGA AGTTAAAATC GGTGCGAAGA CTTCTGTTAT TAAAGAAAAA
 301  GACGGTAAGT TGGTTACTGG TAAAGACAAA GGCGAGAATG GTTCTTCTAC
 351  AGACGAAGGC GAAGGCTTAG TGACTGCAAA AGAAGTGATT GATGCAGTAA
 401  ACAAGGCTGG TTGGAGAATG AAAACAACAA CCGCTAATGG TCAAACAGGT
 451  CAAGCTGACA AGTTTGAAAC CGTTACATCA GGCACAAATG TAACCTTTGC
 501  TAGTGGTAAA GGTACAACTG CGACTGTAAG TAAAGATGAT CAAGGCAACA
 551  TCACTGTTAT GTATGATGTA AATGTCGGCG ATGCCCTAAA CGTCAATCAG
 601  CTGCAAAACA GCGGTTGGAA TTTGGATTCC AAAGCGGTTG CAGGTTCTTC
 651  GGGCAAAGTC ATCAGCGGCA ATGTTTCGCC GAGCAAGGGA AAGATGGATG
 701  AAACCGTCAA CATTAATGCC GGCAACAACA TCGAGATTAC CCGCAACGGT
 751  AAAAATATCG ACATCGCCAC TTCGATGACC CCGCAGTTTT CCAGCGTTTC
 801  GCTCGGCGCG GGGGCGGATG CGCCCACTTT GAGCGTGGAT GGGGACGCAT
 851  TGAATGTCGG CAGCAAGAAG GACAACAAAC CCGTCCGCAT TACCAATGTC
 901  GCCCCGGGCG TTAAAGAGGG GGATGTTACA AACGTCGCAC AACTTAAAGG
 951  CGTGGCGCAA AACTTGAACA ACCGCATCGA CAATGTGGAC GGCAACGCGC
1001  GTGCGGGCAT CGCCCAAGCG ATTGCAACCG CAGGTCTGGT TCAGGCGTAT
1051  TTGCCCGGCA AGAGTATGAT GGCGATCGGC GGCGGCACTT ATCGCGGCGA
1101  AGCCGGTTAC GCCATCGGCT ACTCCAGTAT TTCCGACGGC GGAAATTGGA
1151  TTATCAAAGG CACGGCTTCC GGCAATTCGC GCGGCCATTT CGGTGCTTCC
1201  GCATCTGTCG GTTATCAGTG GTAA
```

B

FIG. 7

```
  1  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51  ANRAASVKDV LNAGWNIKGV KPGTTASDNV DFVRTYDTVE FLSADTKTTT
101  VNVESKDNGK KTEVKIGAKT SVIKEKDGKL VTGKDKGENG SSTDEGEGLV
151  TAKEVIDAVN KAGWRMKTTT ANGQTGQADK FETVTSGTNV TFASGKGTTA
201  TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS GWNLDSKAVA GSSGKVISGN
251  VSPSKGKMDE TVNINAGNNI EITRNGKNID IATSMTPQFS SVSLGAGADA
301  PTLSVDGDAL NVGSKKDNKP VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN
351  RIDNVDGNAR AGIAQAIATA GLVQAYLPGK SMMAIGGGTY RGEAGYAIGY
401  SSISDGGNWI IKGTASGNSR GHFGASASVG YQW*
```

A

```
   1  ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101  TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151  GCTAACCGTG CGGCAAGCGT TAAAGACGTA TTAAACGCTG GCTGGAACAT
 201  TAAAGGCGTT AAACCCGGTA CAACAGCTTC CGATAACGTT GATTTCGTCC
 251  GCACTTACGA CACAGTCGAG TTCTTGAGCG CAGATACGAA AACAACGACT
 301  GTTAATGTGG AAAGCAAAGA CAACGGCAAG AAAACCGAAG TTAAAATCGG
 351  TGCGAAGACT TCTGTTATTA AAGAAAAAGA CGGTAAGTTG GTTACTGGTA
 401  AAGACAAAGG CGAGAATGGT TCTTCTACAG ACGAAGGCGA AGGCTTAGTG
 451  ACTGCAAAAG AAGTGATTGA TGCAGTAAAC AAGGCTGGTT GGAGAATGAA
 501  AACAACAACC GCTAATGGTC AAACAGGTCA AGCTGACAAG TTTGAAACCG
 551  TTACATCAGG CACAAATGTA ACCTTTGCTA GTGGTAAAGG TACAACTGCG
 601  ACTGTAAGTA AAGATGATCA AGGCAACATC ACTGTTATGT ATGATGTAAA
 651  TGTCGGCGAT GCCCTAAACG TCAATCAGCT GCAAAACAGC GGTTGGAATT
 701  TGGATTCCAA AGCGGTTGCA GGTTCTTCGG GCAAAGTCAT CAGCGGCAAT
 751  GTTTCGCCGA GCAAGGGAAA GATGGATGAA ACCGTCAACA TTAATGCCGG
 801  CAACAACATC GAGATTACCC GCAACGGTAA AAATATCGAC ATCGCCACTT
 851  CGATGACCCC GCAGTTTTCC AGCGTTTCGC TCGGCGCGGG GGCGGATGCG
 901  CCCACTTTGA GCGTGGATGG GGACGCATTG AATGTCGGCA GCAAGAAGGA
 951  CAACAAACCC GTCCGCATTA CCAATGTCGC CCCGGGCGTT AAAGAGGGGG
1001  ATGTTACAAA CGTCGCACAA CTTAAAGGCG TGGCGCAAAA CTTGAACAAC
1051  CGCATCGACA ATGTGGACGG CAACGCGCGT GCGGGCATCG CCCAAGCGAT
1101  TGCAACCGCA GGTCTGGTTC AGGCGTATTT GCCCGGCAAG AGTATGATGG
1151  CGATCGGCGG CGGCACTTAT CGCGGCGAAG CCGGTTACGC CATCGGCTAC
1201  TCCAGTATTT CCGACGGCGG AAATTGGATT ATCAAAGGCA CGGCTTCCGG
1251  CAATTCGCGC GGCCATTTCG GTGCTTCCGC ATCTGTCGGT TATCAGTGGT
1301  AA
```

B

FIG. 8

```
  1  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
 51  ANTLKAGDNL KIKQFTYSLK KDLTDLTSVG TEKLSFSANG NKVNITSDTK
101  GLNFAKETAG TNGDTTVHLN GIGSTLTDRA ASVKDVLNAG WNIKGVKNVD
151  FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS VIKEKDGKLV
201  TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA NGQTGQADKF
251  ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA LNVNQLQNSG
301  WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE ITRNGKNIDI
351  ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV RITNVAPGVK
401  EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG LVQAYLPGKS
451  MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG HFGASASVGY
501  QW*
```

A

```
   1  ATGAACAAAA TATACCGCAT CATTTGGAAT AGTGCCCTCA ATGCATGGGT
  51  CGTCGTATCC GAGCTCACAC GCAACCACAC CAAACGCGCC TCCGCAACCG
 101  TGAAGACCGC CGTATTGGCG ACTCTGTTGT TTGCAACGGT TCAGGCAAGT
 151  GCTAACACCC TCAAAGCCGG CGACAACCTG AAAATCAAAC AATTCACCTA
 201  CTCGCTGAAA AAAGACCTCA CAGATCTGAC CAGTGTTGGA ACTGAAAAAT
 251  TATCGTTTAG CGCAAACGGC AATAAAGTCA ACATCACAAG CGACACCAAA
 301  GGCTTGAATT TTGCGAAAGA AACGGCTGGG ACGAACGGCG ACACCACGGT
 351  TCATCTGAAC GGTATTGGTT CGACTTTGAC CGATCGTGCG GCAAGCGTTA
 401  AAGACGTATT AAACGCTGGC TGGAACATTA AAGGCGTTAA AAACGTTGAT
 451  TTCGTCCGCA CTTACGACAC AGTCGAGTTC TTGAGCGCAG ATACGAAAAC
 501  AACGACTGTT AATGTGGAAA GCAAAGACAA CGGCAAGAAA ACCGAAGTTA
 551  AAATCGGTGC GAAGACTTCT GTTATTAAAG AAAAAGACGG TAAGTTGGTT
 601  ACTGGTAAAG ACAAAGGCGA GAATGGTTCT TCTACAGACG AAGGCGAAGG
 651  CTTAGTGACT GCAAAAGAAG TGATTGATGC AGTAAACAAG GCTGGTTGGA
 701  GAATGAAAAC AACAACCGCT AATGGTCAAA CAGGTCAAGC TGACAAGTTT
 751  GAAACCGTTA CATCAGGCAC AAATGTAACC TTTGCTAGTG GTAAAGGTAC
 801  AACTGCGACT GTAAGTAAAG ATGATCAAGG CAACATCACT GTTATGTATG
 851  ATGTAAATGT CGGCGATGCC CTAAACGTCA ATCAGCTGCA AAACAGCGGT
 901  TGGAATTTGG ATTCCAAAGC GGTTGCAGGT TCTTCGGGCA AAGTCATCAG
 951  CGGCAATGTT TCGCCGAGCA AGGGAAAGAT GGATGAAACC GTCAACATTA
1001  ATGCCGGCAA CAACATCGAG ATTACCCGCA ACGGTAAAAA TATCGACATC
1051  GCCACTTCGA TGACCCCGCA GTTTTCCAGC GTTTCGCTCG GCGCGGGGGC
1101  GGATGCGCCC ACTTTGAGCG TGGATGGGGA CGCATTGAAT GTCGGCAGCA
1151  AGAAGGACAA CAAACCCGTC CGCATTACCA ATGTCGCCCC GGGCGTTAAA
1201  GAGGGGGATG TTACAAACGT CGCACAACTT AAAGGCGTGG CGCAAAACTT
1251  GAACAACCGC ATCGACAATG TGGACGGCAA CGCGCGTGCG GGCATCGCCC
1301  AAGCGATTGC AACCGCAGGT CTGGTTCAGG CGTATTTGCC CGGCAAGAGT
1351  ATGATGGCGA TCGGCGGCGG CACTTATCGC GGCGAAGCCG GTTACGCCAT
1401  CGGCTACTCC AGTATTTCCG ACGGCGGAAA TTGGATTATC AAAGGCACGG
1451  CTTCCGGCAA TTCGCGCGGC CATTTCGGTG CTTCCGCATC TGTCGGTTAT
1501  CAGTGGTAA
```

B

FIG. 9

```
             1                                                   50
H41          MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
PMC21        MNKIYRIIWN SALNAWVVVS DLTRNHTKRA SATVNTAVLA TLLFATVQAS
H41Studel    MNKIYRIIWN SALNAWVAVS ELTRNHTKRA SATVKTAVLA TLLFATVQAN
PMC21Bgldel  MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
PMC21C1C5    MNKIYRIIWN SALNAWVVVS ELTRNHTKRA SATVKTAVLA TLLFATVQAS
                                   C1

             51                                                 100
H41          ATDED..EEE ELESVQRS.V VGSIQASMEG SVELET...I SLSMTNDSKE
PMC21        ANNEEQEEYL YLHPVQRTVA VLIVNSDKEG AGEKEKVEEN SDWAVYFNEK
H41Studel    ATDE...... .......... .......... .......... ..........
PMC21Bgldel  ANNE...... .......... .......... .......... ..........
PMC21C1C5    AN ....... .......... .......... .......... ..........
                                   V1

             101                                                150
H41          FVDPYIVVTL KAGDNLKIKQ N.TNENTNAS SFTYSLKKDL TGLINVETEK
PMC21        GVLTAREITL KAGDNLKIKQ NGTN...... .FTYSLKKDL TDLTSVGTEK
H41Studel    .......... .......... .......... .......... TGLINVETEK
PMC21Bgldel  .......... .......... .......... .......... TDLTSVGTEK
PMC21C1C5    .......... .......... .......... .......... ..........
             V1           C2           V2                C3

             151                                                200
H41          LSFGANGKKV NIISDTKGLN FAKETAGTNG DTTVHLNGIG STLTDMLLNT
PMC21        LSFSAHGNKV NITSDTKGLN FAKETAGTNG DTTVHLNGIG STLTDTLLNT
H41Studel    LSFGANGKKV NIISDTKGLN FAKETAGTNG DTTVHLNGIG STLTDMLLNT
PMC21Bgldel  LSFSANGNKV NITSDTKGLN FAKETAGTNG DTTVHLNGIG STLTDTLLNT
PMC21C1C5    .......... .......... .......... .......... ..........
                            C3                                V3

             201                                                250
H41          GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21        GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
H41Studel    GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21Bgldel  GATTNVTNDN VTDDEKKRAA SVKDVLNAGW NIKGVKPGTT ASDNVDFVRT
PMC21C1C5    .......... .......... .......... .......... ...NVDFVRT
                 V3                    C4                V4   C5

             251                                                300
H41          YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKG
PMC21        YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
H41Studel    YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKG
PMC21Bgldel  YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
PMC21C1C5    YDTVEFLSAD TKTTTVNVES KDNGKKTEVK IGAKTSVIKE KDGKLVTGKD
                                   C5

             301                                                350
H41          KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21        KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
H41Studel    KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21Bgldel  KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
PMC21C1C5    KGENGSSTDE GEGLVTAKEV IDAVNKAGWR MKTTTANGQT GQADKFETVT
                                   C5

             351                                                400
H41          SGTKVTFASG NGTTATVSKD DQGNITVKYD VNVGDALNVN QLQNSGWNLD
PMC21        SGTNVTFASG KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
H41Studel    SGTKVTFASG NGTTATVSKD DQGNITVKYD VNVGDALNVN QLQNSGWNLD
PMC21Bgldel  SGTNVTFASG KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
PMC21C1C5    SGTNVTFASG KGTTATVSKD DQGNITVMYD VNVGDALNVN QLQNSGWNLD
                                   C5
```

FIG. 10

```
             401                                                    450
         H41 SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
       PMC21 SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
   H41Studel SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
PMC21Bgldel SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
  PMC21C1C5 SKAVAGSSGK VISGNVSPSK GKMDETVNIN AGNNIEITRN GKNIDIATSM
                                     C5

             451                                                    500
         H41 TPQFSSVSLG AGADAPTLSV DDEGALNVGS KDANKPVRIT NVAPGVKEGD
       PMC21 TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
   H41Studel TPQFSSVSLG AGADAPTLSV DDEGALNVGS KDANKPVRIT NVAPGVKEGD
PMC21Bgldel TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
  PMC21C1C5 TPQFSSVSLG AGADAPTLSV DG.DALNVGS KKDNKPVRIT NVAPGVKEGD
                                     C5

             501                                                    550
         H41 VTNVAQLKGV AQNLNNRIDN VNGNARAGIA QAIATAGLVQ AYLPGKSMMA
       PMC21 VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
   H41Studel VTNVAQLKGV AQNLNNRIDN VNGNARAGIA QAIATAGLVQ AYLPGKSMMA
PMC21Bgldel VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
  PMC21C1C5 VTNVAQLKGV AQNLNNRIDN VDGNARAGIA QAIATAGLVQ AYLPGKSMMA
                                     C5

             551                                                    600
         H41 IGGGTYLGEA GYAIGYSSIS AGGNWIIKGT ASGNSRGHFG ASASVGYQW.
       PMC21 IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
   H41Studel IGGGTYLGEA GYAIGYSSIS AGGNWIIKGT ASGNSRGHFG ASASVGYQW.
PMC21Bgldel IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
  PMC21C1C5 IGGGTYRGEA GYAIGYSSIS DGGNWIIKGT ASGNSRGHFG ASASVGYQW.
                                     C5
```

FIG. 10 cont'd

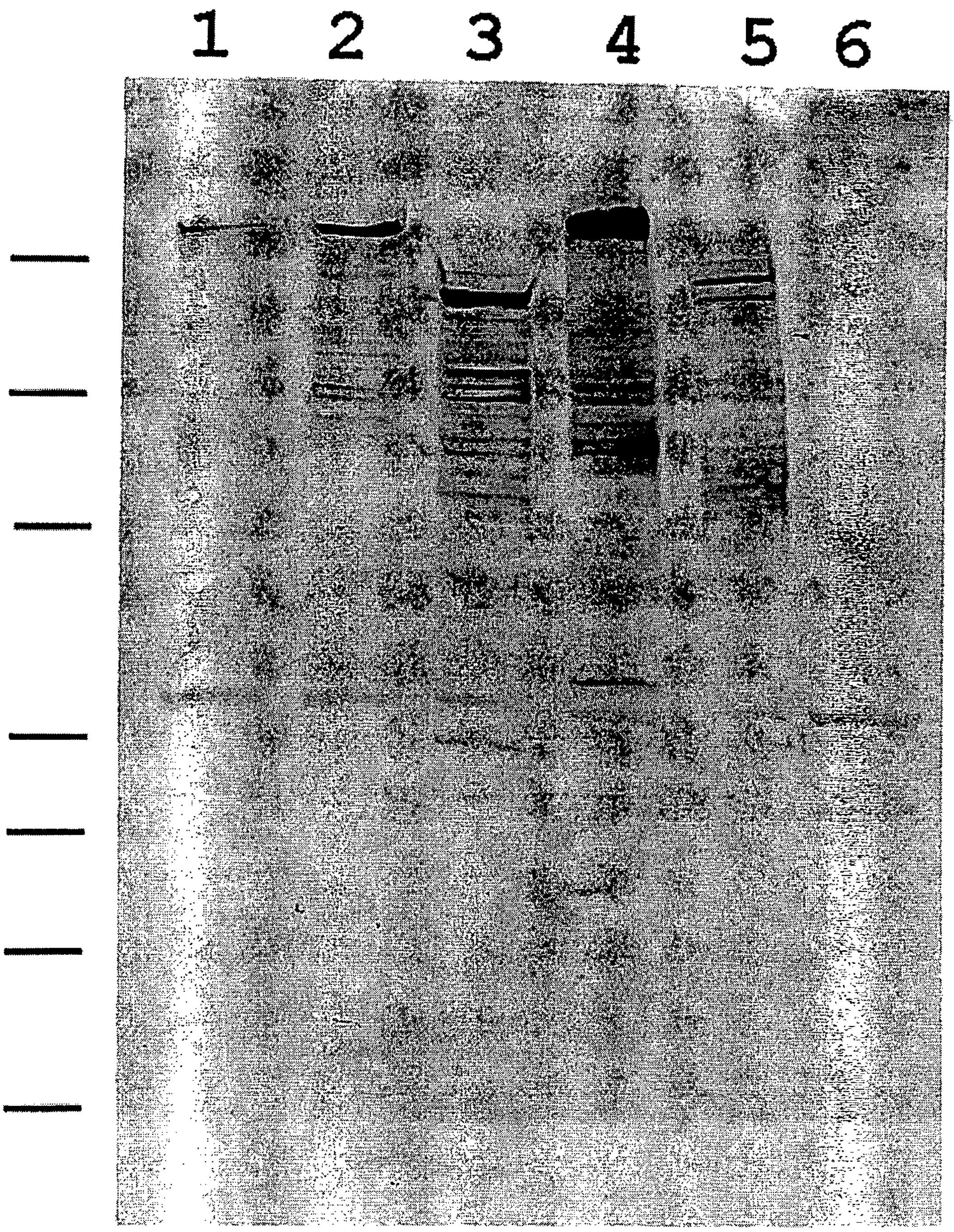

FIG. 11

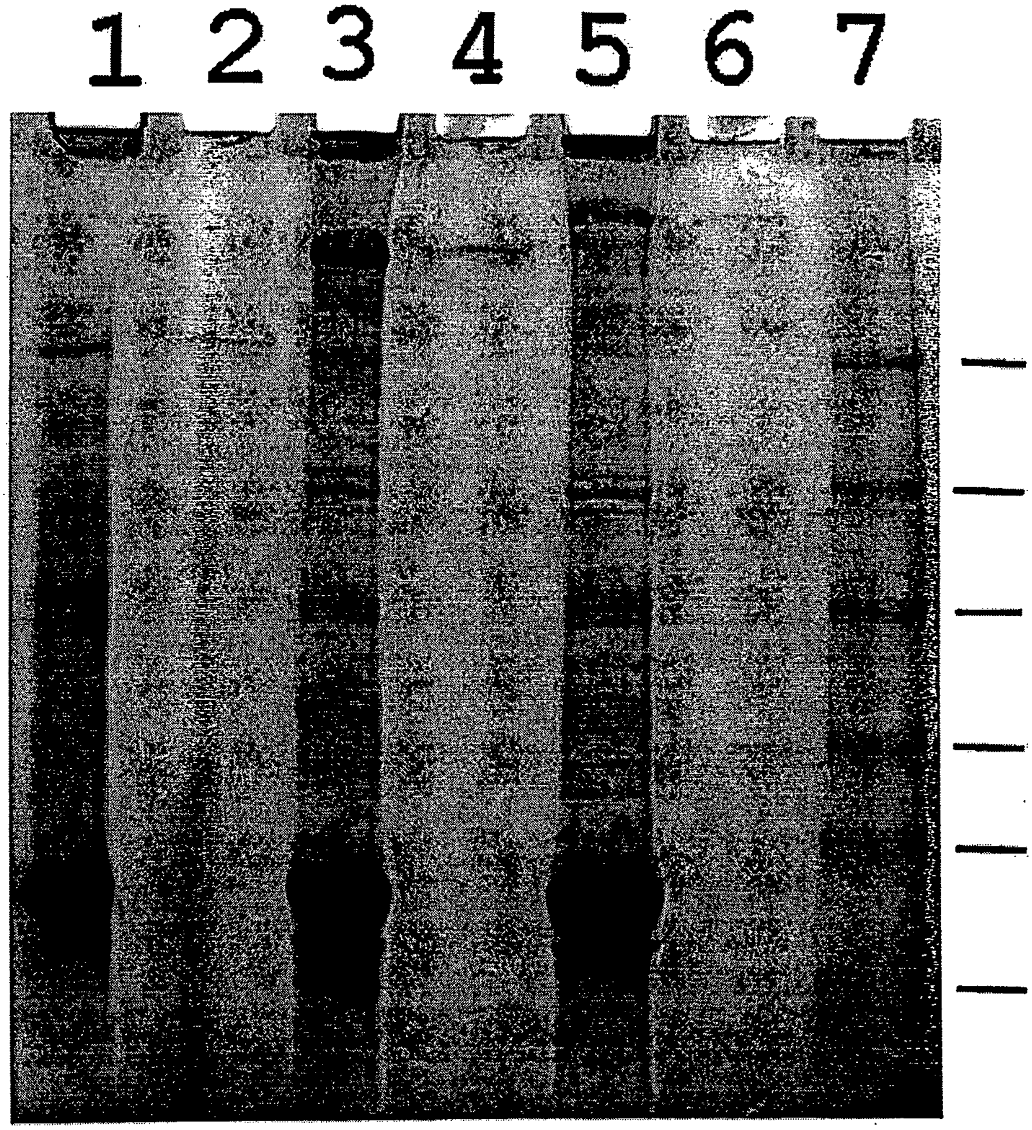

FIG. 12

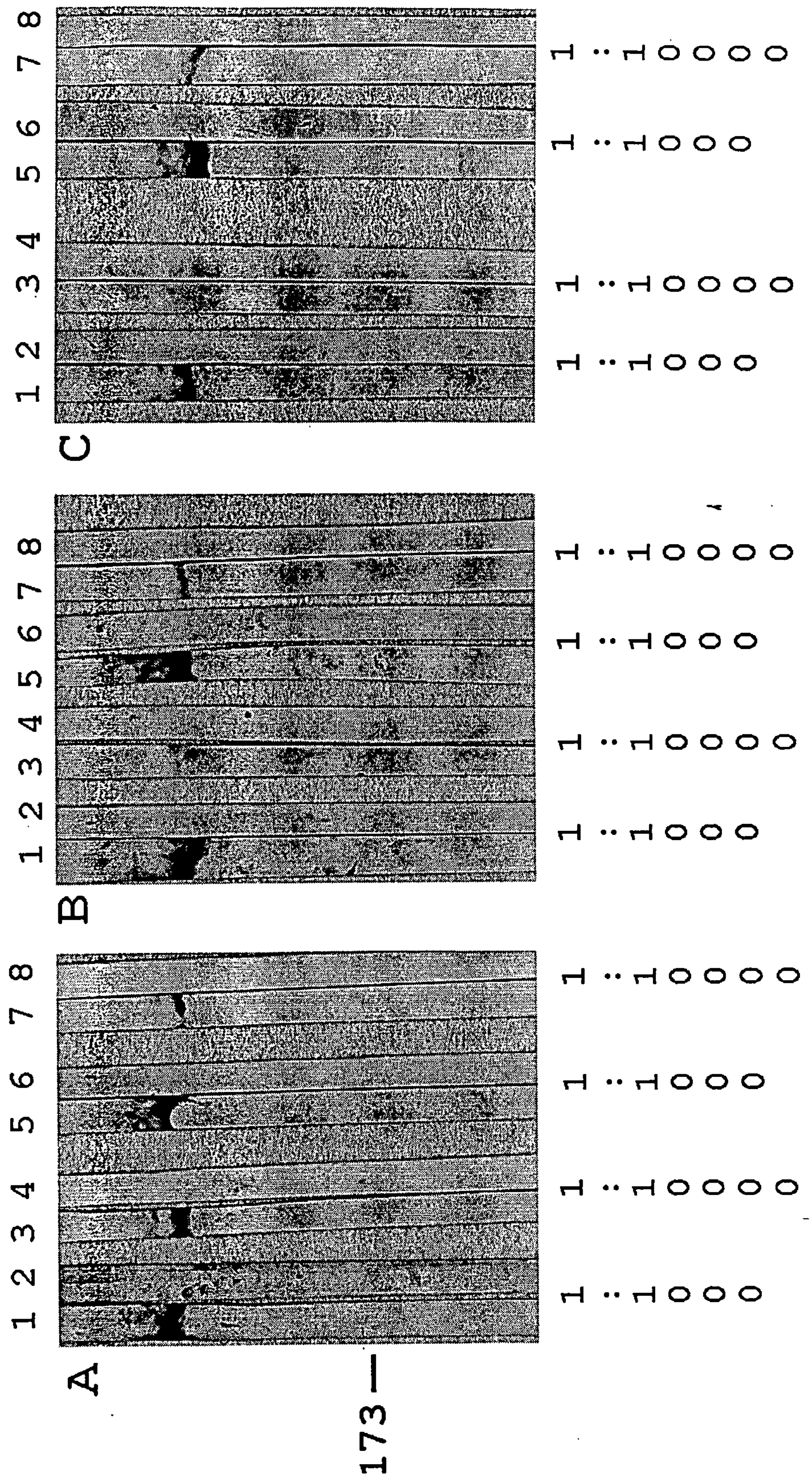
C
1 2 3 4 5 6 7 8
1:1000
1:10000
1:1000
1:10000
B
1 2 3 4 5 6 7 8
1:1000
1:10000
1:1000
1:10000
A
1 2 3 4 5 6 7 8
1:1000
1:10000
1:1000
1:10000
173

FIG. 13

A

52 NNEEQEEYL YLHPVQRTVA VLIVNSDKEG AGEKEKVEEN SDWAVYFNEK

101 GVLTAREITL KAGDNLKIKQ NGTNFTYSLK KDLTDLTSVG TEKLSFSAHG

151 NKVNITSDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDTL LNTGATTNVT

201 NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF VRTYDTVEFL

251 SADTKTTTVN VESKDNGKKT EVKIGAKTSV IKEKDGKLVT GKDKGENGSS

301 TDEGEGLVTA KEVIDAVNKA GWRMKTTTAN GQTGQADKFE TVTSGTNVTF

351 ASGKGTTATV SKDDQGNITV MYDVNVGDAL NVNQLQNSGW NLDSKAVAGS

401 SGKVISGNVS PSKGKMDETV NINAGNNIEI TRNGKNIDIA TSMTPQFSSV

451 SLGAGADAPT LSVDGDALNV GSKKDNKPVR ITNVAPGVKE GDVTNVAQLK

501 GVAQNLNNRI DNVDGNARAG IAQAIATAGL VQAYLPGKSM MAIGGGTYRG

551 EAGYAIGYSS ISDGGNWIIK GTASGNSRGH FGASASVGYQ W*

B

52 TDEDEEEEL ESVQRSVVGS IQASMEGSVE LETISLSMTN DSKEFVDPYI

101 VVTLKAGDNL KIKQNTNENT NASSFTYSLK KDLTGLINVE TEKLSFGANG

151 KKVNIISDTK GLNFAKETAG TNGDTTVHLN GIGSTLTDML LNTGATTNVT

201 NDNVTDDEKK RAASVKDVLN AGWNIKGVKP GTTASDNVDF VRTYDTVEFL

251 SADTKTTTVN VESKDNGKKT EVKIGAKTSV IKEKDGKLVT GKGKGENGSS

301 TDEGEGLVTA KEVIDAVNKA GWRMKTTTAN GQTGQADKFE TVTSGTKVTF

351 ASGNGTTATV SKDDQGNITV KYDVNVGDAL NVNQLQNSGW NLDSKAVAGS

401 SGKVISGNVS PSKGKMDETV NINAGNNIEI TRNGKNIDIA TSMTPQFSSV

451 SLGAGADAPT LSVDDEGALN VGSKDANKPV RITNVAPGVK EGDVTNVAQL

501 KGVAQNLNNR IDNVNGNARA GIAQAIATAG LVQAYLPGKS MMAIGGGTYL

551 GEAGYAIGYS SISAGGNWII KGTASGNSRG HFGASASVGY QW*

FIG. 14

C

52 NNETDLTSV GTEKLSFSAN GNKVNITSDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDT LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV
401 RITNVAPGVK EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG
451 LVQAYLPGKS MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG
501 HFGASASVGY QW*

D

52 TDETGLINV ETEKLSFGAN GKKVNIISDT KGLNFAKETA GTNGDTTVHL
101 NGIGSTLTDM LLNTGATTNV TNDNVTDDEK KRAASVKDVL NAGWNIKGVK
151 PGTTASDNVD FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS
201 VIKEKDGKLV TGKGKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA
251 NGQTGQADKF ETVTSGTKVT FASGNGTTAT VSKDDQGNIT VKYDVNVGDA
301 LNVNQLQNSG WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE
351 ITRNGKNIDI ATSMTPQFSS VSLGAGADAP TLSVDDEGAL NVGSKDANKP
401 VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN RIDNVNGNAR AGIAQAIATA
451 GLVQAYLPGK SMMAIGGGTY LGEAGYAIGY SSISAGGNWI IKGTASGNSR
501 GHFGASASVG YQW*

FIG. 14

E

52 NNVDFVRTY DTVEFLSADT KTTTVNVESK DNGKKTEVKI GAKTSVIKEK
101 DGKLVTGKDK GENGSSTDEG EGLVTAKEVI DAVNKAGWRM KTTTANGQTG
151 QADKFETVTS GTNVTFASGK GTTATVSKDD QGNITVMYDV NVGDALNVNQ
201 LQNSGWNLDS KAVAGSSGKV ISGNVSPSKG KMDETVNINA GNNIEITRNG
251 KNIDIATSMT PQFSSVSLGA GADAPTLSVD GDALNVGSKK DNKPVRITNV
301 APGVKEGDVT NVAQLKGVAQ NLNNRIDNVD GNARAGIAQA IATAGLVQAY
351 LPGKSMMAIG GGTYRGEAGY AIGYSSISDG GNWIIKGTAS GNSRGHFGAS
401 ASVGYQW*

F

52 NRAASVKDV LNAGWNIKGV KPGTTASDNV DFVRTYDTVE FLSADTKTTT
101 VNVESKDNGK KTEVKIGAKT SVIKEKDGKL VTGKDKGENG SSTDEGEGLV
151 TAKEVIDAVN KAGWRMKTTT ANGQTGQADK FETVTSGTNV TFASGKGTTA
201 TVSKDDQGNI TVMYDVNVGD ALNVNQLQNS GWNLDSKAVA GSSGKVISGN
251 VSPSKGKMDE TVNINAGNNI EITRNGKNID IATSMTPQFS SVSLGAGADA
301 PTLSVDGDAL NVGSKKDNKP VRITNVAPGV KEGDVTNVAQ LKGVAQNLNN
351 RIDNVDGNAR AGIAQAIATA GLVQAYLPGK SMMAIGGGTY RGEAGYAIGY
401 SSISDGGNWI IKGTASGNSR GHFGASASVG YQW*

G

50 SANTLKAGDNL KIKQFTYSLK KDLTDLTSVG TEKLSFSANG NKVNITSDTK
101 GLNFAKETAG TNGDTTVHLN GIGSTLTDRA ASVKDVLNAG WNIKGVKNVD
151 FVRTYDTVEF LSADTKTTTV NVESKDNGKK TEVKIGAKTS VIKEKDGKLV
201 TGKDKGENGS STDEGEGLVT AKEVIDAVNK AGWRMKTTTA NGQTGQADKF
251 ETVTSGTNVT FASGKGTTAT VSKDDQGNIT VMYDVNVGDA LNVNQLQNSG
301 WNLDSKAVAG SSGKVISGNV SPSKGKMDET VNINAGNNIE ITRNGKNIDI
351 ATSMTPQFSS VSLGAGADAP TLSVDGDALN VGSKKDNKPV RITNVAPGVK
401 EGDVTNVAQL KGVAQNLNNR IDNVDGNARA GIAQAIATAG LVQAYLPGKS
451 MMAIGGGTYR GEAGYAIGYS SISDGGNWII KGTASGNSRG HFGASASVGY
501 QW*

FIG. 14

INTERNATIONAL SEARCH REPORT

International application No.
PCT/AU01/00069

A.	**CLASSIFICATION OF SUBJECT MATTER**
Int. Cl. 7:	C07K 14/22; C12N 15/31

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC (WPIDS) AND CHEMICAL ABSTRACTS - KEYWORDS BELOW

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
SEE BELOW

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
Swissprot, Genpept, Pir, Trembl, Genbank, Embl, Wpids, CA, Medline. Keywords: neisseria, adhesin, surface, protein, antigen, nhha, hianm, modif?, mutat?, conserv?

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 99/36544 A (CHIRON S.P.A.) 22 July 1999 (See SEQ ID 5)	1-3, 9, 18, 21-24
X	WO 99/58683 A (SMITHKLINE BEECHAM BIOLOGICALS S.A.) 18 November 1999 (See SEQ ID 1, 3)	1-3, 9, 18, 21-24
P,X	WO 00/66791 A (CHIRON CORPORATION et al.) 9 November 2000 (See SEQ ID 110)	1-3, 9, 18, 21-24

[X] Further documents are listed in the continuation of Box C [X] See patent family annex

* Special categories of cited documents:

- "A" document defining the general state of the art which is not considered to be of particular relevance
- "E" earlier application or patent but published on or after the international filing date
- "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- "O" document referring to an oral disclosure, use, exhibition or other means
- "P" document published prior to the international filing date but later than the priority date claimed
- "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
- "&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
6 June 2001	13 June 2001

Name and mailing address of the ISA/AU	Authorized officer
AUSTRALIAN PATENT OFFICE PO BOX 200, WODEN ACT 2606, AUSTRALIA E-mail address: pct@ipaustralia.gov.au Facsimile No. (02) 6285 3929	**CHRISTOPHER LUTON** Telephone No : (02) 6283 2256

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GenPept database accession number AAC43721, 21 March 1996, Barenkamp and St Geme	1-3, 9, 18, 21-23
X	GenPept database accession number AAC44560, 27 October 1999, St Geme et al.	1-3, 9, 18, 21-23
X	PIR database accession number I64138, 24 October 1997, Fleischmann et al.	1-3, 9, 18, 21-23
X	WO 99/31132 A (THE UNIVERSITY OF QUEENSLAND AND ISIS INNOVATION LIMITED) 24 June 1999 (See figures 6 and 7)	

This Annex lists the known "A" publication level patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent Document Cited in Search Report		Patent Family Member					
WO	9931132	AU	16495/99	BR	9814276	EP	1045859
		NO	20002990	PL	341160	GB	9726398
WO	9936544	AU	19795/99	EP	1047784	GB	9800760
WO	9958683	AU	41420/99	EP	1078063	NO	20005696
WO	200066791	AU	200012022	AU	200032492	WO	200022430

END OF ANNEX