OPIC
Office de la propriété intellectuelle du Canada

CIPO
Canadian Intellectual Property Office

(12) (19) (CA) **Demande-Application**

(21) (A1) **2,304,128**
(86) 1998/09/18
(87) 1999/03/25

(72) BECKER, ROBERT, US
(72) GRAY, MARY-ANN, US
(72) PYLE, DEREK, US
(71) PASTEUR MERIEUX CONNAUGHT, US

(51) Int.Cl.[7] C12N 15/31, A61K 39/395, C07K 14/315, C12N 1/21, C07K 16/12, C12N 5/10, A61K 39/09
(30) 1997/09/18 (08/932,982) US

(54) **PROCEDE DE PRODUCTION AMELIOREE DE PROTEINES DE SURFACE PNEUMOCOCCIQUES**
(54) **IMPROVED METHOD OF PRODUCTION OF PNEUMOCOCCAL SURFACE PROTEINS**

(57) Cette invention concerne des procédés améliorés permettant de préparer des compositions immunologiques ou vaccinales contenant des polypeptides du type PspA. D'une manière plus spécifique cette invention concerne des compositions vaccinales dans lesquelles le polypeptide PspA ou du type PspA provenant de Rx1 ou d'autres représentants du clade 2 ou de la famille 1 ont été produits au moyen de procédés de recombinaison. On décrit des structures de recombinaison qui expriment une forme tronquée, sensiblement homogène de polypeptides du type PspA actifs du point de vue immunologique dans des cellules hôtes transformées avec ces mêmes structures.

(57) The present invention provides improved methods for preparation of immunoligical or vaccine compositions comprising PspA-like polypeptides. Specifically, the present invention provides vaccine compositions wherein PspA or PspA-like polypeptide from Rx1 or other representatives of clade 2 or family 1 have been produced by recombinant methods. Recombinant constructs are provided which express a substantially homogeneous, truncated form of immunologically active PspA-like polypeptides in host cells transformed with such constructs.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, C07K 14/315, C12N 1/21, 5/10, C07K 16/12, A61K 39/09, 39/395

A2

(11) International Publication Number: **WO 99/14333**

(43) International Publication Date: 25 March 1999 (25.03.99)

(21) International Application Number: PCT/US98/19740

(22) International Filing Date: 18 September 1998 (18.09.98)

(30) Priority Data:
08/932,982 18 September 1997 (18.09.97) US

(71) Applicant *(for all designated States except US)*: PASTEUR MERIEUX CONNAUGHT [US/US]; Route 611, P.O. Box 187, Swiftwater, PA 186370 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: BECKER, Robert [US/US]; Sylvan Cascades Road, R.R. 1 Box 545A, Henryville, PA 18332 (US). GRAY, Mary–Ann [US/US]; 289Q Green Chapel Road, Cresco, PA 18326 (US). PYLE, Derek [US/US]; 86 East Broad Street, East Stroudsburg, PA 18301 (US).

(74) Agent: SMITH, G., Kenneth; McDonnell Boehnen Hulbert & Berghoff, 300 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, 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), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
Without international search report and to be republished upon receipt of that report.

(54) Title: IMPROVED METHOD OF PRODUCTION OF PNEUMOCOCCAL SURFACE PROTEINS

(57) Abstract

The present invention provides improved methods for preparation of immunoligical or vaccine compositions comprising PspA–like polypeptides. Specifically, the present invention provides vaccine compositions wherein PspA or PspA–like polypeptide from Rx1 or other representatives of clade 2 or family 1 have been produced by recombinant methods. Recombinant constructs are provided which express a substantially homogeneous, truncated form of immunologically active PspA–like polypeptides in host cells transformed with such constructs.

IMPROVED METHOD OF PRODUCTION OF PNEUMOCOCCAL SURFACE PROTEINS

BACKGROUND OF THE INVENTION

Streptococcus pneumoniae is an important cause of otitis media, meningitis, bacteremia and pneumonia, and a leading cause of fatal infections in the elderly and persons with underlying medical conditions, such as pulmonary disease, liver disease, alcoholism, sickle cell, cerebrospinal fluid leaks, acquired immune deficiency syndrome (AIDS), and patients undergoing immunosuppressive therapy. It is also a leading cause of morbidity in young children. Pneumococcal infections cause approximately 40,000 deaths in the U.S. yearly. The most severe pneumococcal infections involve invasive meningitis and bacteremia infections, of which there are 3,000 and 50,000 cases annually, respectively; despite the use of antibiotics and vaccines, the prevalence of pneumococcal infections has declined little over the past twenty-five years.

It is generally accepted that immunity to *Streptococcus pneumoniae* can be mediated by specific antibodies against the polysaccharide capsule of the pneumococcus. However, neonates and young children fail to make adequate immune response against most capsular polysaccharide antigens and can have repeated infections involving the same capsular serotype. Moreover, there are over ninety known capsular serotypes of *S. pneumoniae*, of which twenty-three account for about 95% of the disease.

Pneumococcal vaccines have been developed by combining the 23 different capsular polysaccharides that are the prevalent types of human pneumococcal disease. These 23 polysaccharide types have been used in a licensed pneumococcal vaccine since 1983 (D.S. Fedson and D. M. Musher, Vaccines, S.A. Plotkin and J.E.A. Montimer, eds., 1994, pp. 517-564). The licensed 23-valent polysaccharide vaccine has a reported efficacy of approximately 60% in preventing bacteremia caused vaccine type pneumococci in healthy adults.

An alternative approach for protecting children, and also the elderly, from pneumococcal infection would be to identify protein antigens that could elicit protective immune responses. Such proteins may serve as a vaccine by themselves, may be used in conjunction with successful polysaccharide-protein conjugates, or as carriers for polysaccharides.

McDaniel *et al.* (I), J. Exp. Med. 160:386-397, 1984, describes the production of hybridoma antibodies that recognize cell surface polypeptide(s) on *S. pneumoniae* and protect mice from infection by encapsulated pneumococci. On the basis of these results, a major surface protein antigen, designated "pneumococcal surface protein A", or "PspA", was identified.

McDaniel *et al.* (II), Microbial Pathogenesis 1:519-531, 1986, performed studies to characterize PspA. Considerable diversity in the PspA molecule in different strains was found, as were differences in the epitopes recognized by different antibodies.

McDaniel *et al.* (III), J. Exp. Med. 165:381-394, 1987, demonstrated that immunization of X-linked immunodeficient (XID) mice with non-encapsulated pneumococci expressing PspA, but not isogenic pneumococci lacking PspA, protects mice from subsequent fatal infection with pneumococci.

McDaniel *et al.* (IV), Infect. Immun., 59:222-228, 1991, demonstrated that immunization of mice with a recombinant full length fragment of PspA was able to elicit protection against pneumococcal strains of capsular types 6A and 3.

Analyses of the nucleotide and amino acid sequences indicate that the PspA molecule comprises three major regions. See, e.g., Briles, *et al.*, U.S. Patent No. 5,476,929, the teachings of which are expressly incorporated by reference. The first 288 amino acids at the amino terminal end of the protein are predicted to have a strong alpha helical structure. The adjacent region of 82 amino acids (289 to 370 of PARx1 PspA) has a high density of proline

residues; based on similar regions in other prokaryotic proteins, this region is believed to traverse the bacterial cell wall. The remaining 201 amino acids at the carboxyl-terminal end of the molecule (371 to 571 of PARx1 PspA) have a repeated amino acid sequence that binds to phosphocholine and lipoteichoic acids followed by a seventeen (17) amino acid tail (572-588). Based on this structure, the PspA molecule is thought to associate with the inner membrane and lipoteichoic acids via the repeated region in the middle of the carboxyl-terminal end of the protein. The proline region in the middle of the protein is thought to traverse the cell wall placing the alpha helical region on the outer surface of the *S. pneumoniae* cells. This model is consistent with the demonstration that the alpha helical region, which extends from the surface of the cell, contains the protective epitopes (Yother, J. *et al.*, J. Bacteriol. 1992: 174:601-609; Yother, J. *et al.*, J. Bacteriol. 1994:176:2976-2985; McDaniels, L.S. *et al.*, Microbial. Pathog. 1994; 17:323-337; and Ralph, B.A., *et al.*, Ann. N.Y. Acad. Sci. 1994; 730:361-363).

In addition to the published literature specifically referred to above, the inventors, in conjunction with co-workers, have published further details concerning PspA's, as follows:

1. Abstracts of 89th Annual Meeting of the American Society for Microbiology, p. 125, item D-257, May, 1989;

2. Abstracts of 90th Annual Meeting of the American Society for Microbiology, p. 98, item D-106, May 1990;

3. Abstracts of 3rd International ASM Conference on Streptococcal Genetics, p. 11, item 12, June 1990;

4. Talkington *et al.*, Infect. Immun. 59:1285-1289, 1991;

5. Yother *et al.*, (I), J. Bacteriol. 174:601-609, 1992; and

6. Yother *et al.*, (II), J. Bacteriol. 174:610-618, 1992.

7. McDaniel *et al.*, (V), Microbiol. Pathogenesis, 13:261-268, 1994.

In addition to the above studies, 08/529,055, filed September 15, 1995, 08/470,626, filed June 6, 1995, 08/467, 852, filed June 6, 1995, 08/469,434, filed June 6, 1995,

08/468,718, filed June 6, 1995, 08/247,491, filed May 23, 1994, 08/214,222, filed March 17, 1994, 08/214,164, filed March 17, 1994, 08/246,636, filed May 20, 1994, and 08/319,795, filed October 7, 1994, and U.S. Patent No. 5,476,929, relate to vaccines comprising PspA and fragments thereof, methods for expressing DNA encoding PspA and fragments thereof, DNA encoding PspA and fragments thereof, the amino acid sequences of PspA and fragments thereof, compositions containing PspA and fragments thereof and methods of using such compositions. Each of these applications, as well as each document or reference cited in these applications and each document or reference provided above, is hereby incorporated herein by reference. Documents or references are also cited in the following text, either in the Reference List before the claims or in the text itself; each of these documents or references is hereby expressly incorporated herein by reference.

The prior art fails to provide an immunological composition or vaccination regimen which would elicit protection against various diversified pneumococcal strains, without having to combine a large number of possibly competitive antigens within the same formulation. The present invention fulfills that need by employing recombinant DNA technology to more reproducibly prepare immunogenic polypeptides designed to confer broad protection against diverse pneumococcal strains.

In the specification which follows and the drawings accompanying the same, there are utilized certain accepted abbreviations with respect to the amino acids represented thereby. The following Table 1 identifies those abbreviations:

TABLE I

AMINO ACID ABBREVIATIONS

A=Ala-Alanine	M=Met=Methionine
C=Cys=Cysteine	N=Asn=Asparagine
D=Asp=Aspartic Acid	P=Pro=Proline
E-Glu=Glutamic Acid	Q=Gln=Glutamine

F=Phe=Phenylalanine
G=Gly=Glycine
H=His=Histidine
I=Ile=Isoleucine
K=Lys=Lysine
L=Leu=Leucine
R=Arg=Arginine
S=Ser=Serine
T=Thr=Threonine
V=Val=Valine
W=Trp=Tryptophan
Y=Tyr=Tyrosine

SUMMARY OF THE INVENTION

The present invention relates to improved recombinant PspA constructs which allows for higher yield and more efficient purification of PspA-like polypeptides when expressed in prokaryotic hosts. (As used herein, the term "recombinant construct" or "construct" includes vectors and plasmid encoding, and capable of expressing, foreign genes in transformed host cells.) The wild type *pspA* gene from strain Rx1 of *S. pneumoniae* contains an ATG at codon position 96, which upon expression produces two forms of PspA polypeptide, a first resulting from initiation of translation at position 1 and a second truncated form of PspA, resulting from initiation of translation at position 96. The wild-type *pspA* gene has been truncated and transformed into *E. coli* where it can produce two forms of a truncated PspA polypeptide; a first, 314 aa (50 kD) form, resulting from initiation of translation at position 1 and a second, further truncated form of PspA, resulting from initiation of translation at position 96 (i.e., aa96-314, corresponding to an approximate ms of 35 kD). Recombinants comprising the modified and truncated pspA gene, in which the ATG at position 96 was altered to no longer encode methionine, expressed only one form of PspA, the 50kD aa molecule, PARx1-MI.

Rx1 is a representative of one of four families of *S. pneumoniae*, and vaccine compositions containing PspA-like polypeptides can now be more conveniently prepared using the recombinant PspA construct of the present invention. By eliminating the second, truncated form of PspA at the genetic level, the yield of the recombinant product of interest is improved, and this product can be more readily purified following expression in recombinant hosts. The immunodominant epitopes of native PARx1 are retained in the modified PspA-

like polypeptide of the present invention, making this polypeptide a suitable component in immunological compositions designed to elicit an immune response to representatives of clade 2 of *S. pneumoniae*.

Genes derived from a variety of strains of *S. pneumoniae*, including representatives of other families or clades (as hereinafter defined) expressing PspA or PspA-like polypeptides (including the recently described PspC polypeptides) may contain internal translational initiation sites similar to the ATG sequence at codon position 96 of the *pspA* gene of Rx1. Such sites, regardless of their specific codon position relative to codon position 96 of the pspA gene from strain Rx1, will likely complicate developments of homogeneous preparations of such polypeptides by recombinant methods. The skilled artisan can readily adopt the teachings of the present invention to modify such sites and thereby improve yield and ease of purification of the immunologically active products of interest. While the examples contained herein refer to modifications of recombinant pspA from Rx1, this is meant to be merely illustrative of the more general teachings of the present invention 1.

BRIEF DESCRIPTION OF DRAWINGS

FIG. 1 shows the DNA sequence of the modified *pspA* gene of the present invention and the deduced amino acid sequence of the PspA-like polypeptide encoded by the modified *pspA* gene.

DETAILED DESCRIPTION

A comparison of the amino acid sequences in the C-terminal region of the alpha helix of PspAs from 24 strains of *S. pneumoniae* has revealed that the PspA strains can be grouped into 6 clades with greater than 75% homology, and these clades can be grouped into 4 families with greater than 50% homology.

A clade is defined herein as comprising PspAs which exhibit greater than 75% sequence homology in aligned sequences of the C-terminal region of the alpha helix, and a family is defined herein as those clades which exhibit greater than or equal to 50% homology between member PspA sequences in aligned sequences of the C-terminal region of the alpha helix.

PspAs of strains within the same PspA clade exhibit reciprocal cross-protection from immunization and challenge experiments. The present invention contemplates vaccine compositions comprising two or more, preferably no more than 10, and more preferably a minimum of 4 and a maximum of 6 recombinant forms of PspA-like polypeptide representing a single clade each, and a pharmaceutically acceptable carrier or diluent. In a preferred embodiment, a homogeneous form of a recombinant PspA-like polypeptide from Rx1, a member of clade 2, is the preferred representative of clade 2 and/or family 1 which is optionally included in the vaccine composition of the present invention.

Recombinant methods are preferred since a high yield is desired. The basic steps in the recombinant production of modified PspA include:

a) construction of a synthetic or semi-synthetic DNA encoding the PspA-like polypeptide,

b) integrating said DNA into an expression vector in a manner suitable for the expression of the PspA-like polypeptide either alone or as a fusion protein,

c) transforming an appropriate eukaryotic or prokaryotic host cell with said expression vector,

d) culturing said transformed or transfected host cell, and

e) recovering and purifying the recombinantly produced PspA-like polypeptide.

For recombinant expression, the sequence coding for a PspA-like polypeptide may be wholly synthetic, semi-synthetic or the result of modification of the native *pspA* gene.

Synthetic genes, the in vitro or in vivo transcription and translation of which will result in the production of PspA-like polypeptides may be constructed by techniques well known in the art. Owing to the natural degeneracy of the genetic code, the skilled artisan will recognize that a sizable yet definite number of DNA sequences may be constructed which encode PspA-like polypeptides. The gene encoding the PspA-like polypeptide may be created by synthetic methodology. Such methodology of synthetic gene construction is well known in the art. Brown, E.L., Belagaje, R., Ryan, M.J., and Khorana, H.G. (1979) in Methods in Enzymology, Academic Press, N.Y., Vol. 68, pgs. 109-151, the entire teaching of which is hereby incorporated by reference. The DNA segments corresponding to the PspA gene, or fragments thereof, are generated using conventional DNA synthesizing apparatus such as the Applied Biosystems Model 380A or 380B DNA synthesizers (commercially available from Applied Biosystems, Inc., 850 Lincoln Center Drive, Foster City, CA 94404). The synthetic *pspA* gene may be designed to possess restriction endonuclease cleavage sites at either end of the transcript to facilitate isolation from and integration into expression and amplification plasmids. The choice of restriction sites are chosen so as to properly orient the sequence coding for the PspA-like polypeptide with control sequences to achieve proper in-frame reading and expression of the PspA-like polypeptide. A variety of other such cleavage sites

may be incorporated depending on the particular recombinant constructs employed and may be generated by techniques well known in the art.

DNA sequences designed for recombinant production of PspA are well known. For example, 08/529,055, filed September 15, 1995, 08/470,626, filed June 6, 1995, 08/467,852, filed June 6, 1995, 08/469,434, filed June 6, 1995, 08/468,718, filed June 6, 1995, 08/247,491, filed May 23, 1994, 08/214,222, filed March 17, 1994 and 08/214,164, filed March 17, 1994, 08/246,636, filed May 20, 1994, and 08/319,795, filed October 7, 1994, and U.S. Patent No. 5,476,929, relate to vaccines comprising PspA and fragments thereof, methods for expressing DNA encoding PspA and fragments thereof, DNA encoding PspA and fragments thereof, the amino acid sequences of PspA and fragments thereof, compositions containing PspA and fragments thereof and methods of using such compositions.

The "polymerase chain reaction" or "PCR" refers to a procedure or technique in which amounts of a preselected piece of nucleic acid, RNA and/or DNA, are amplified as described in U.S. Patent No. 4,683,195. Generally, sequence information from the ends of the region of interest or beyond is employed to design oligonucleotide primers. These primers will be identical or similar in sequence to opposite strands of the template to be amplified. PCR can be used to amplify specific RNA sequences, specific DNA sequences from total genomic DNA, and cDNA transcribed from total cellular RNA, bacteriophage or plasmid sequences, and the like. See generally Mullis *et al.*, Cold Spring Harbor Symp. Quant. Biol., 51, 263 (1987); Erlich, ed., PCR Technology, (Stockton Press, NY, 1989). PCR can also be used to conveniently introduce any desired sequence change genes of interest. See generally, Ausubel *et al.*, eds, Current Protocols in Molecular Biology, s8.5.1 (John Wiley & Sons, 1995).

Construction of suitable vectors containing the desired coding and control sequences employ standard ligation techniques. Isolated plasmids or DNA fragments are cleaved, tailored, and religated in the form desired to form the plasmids required.

To effect the translation of the desired PspA-like polypeptide sequence, one inserts the engineered DNA sequence coding for the PspA-like polypeptide in any of a plethora of appropriate recombinant DNA expression vectors through the use of appropriate restriction endonucleases. A synthetic version of the DNA coding sequence is designed to posses restriction endonuclease cleavage sites at either end of the transcript to facilitate isolation from and integration into these expression and amplification plasmids. The coding sequence may be readily modified by the use of synthetic linkers to facilitate the incorporation of this sequence into the desired cloning vectors by techniques well known in the art. The particular endonucleases employed will be dictated by the restriction endonuclease cleavage pattern of the parent expression vector to be employed. The choice of restriction sites are chosen so as to properly orient the DNA coding sequence with control sequences to achieve proper in-frame reading and expression of the PspA-like polypeptide.

In general, plasmid vectors containing promoters and control sequences which are derived from species compatible with the host cell are used with these hosts. The vector ordinarily carries a replication site as well as marker sequences which are capable of providing phenotypic selection in transformed cells. For example, *E. coli* is typically transformed using pBR322, a plasmid derived from an *E. coli* species (Bolivar, *et al.*, Gene 2:95 [1977]), pBR322 contains genes for ampicillin and tetracycline resistance and thus provides easy means for identifying transformed cells. The pBR322 plasmid, or other microbial plasmid must also contain or be modified to contain promoters and other control elements commonly used in recombinant DNA construction.

The DNA sequence coding for the PspA-like polypeptide must be positioned so as to be in proper reading frame with the promoter and ribosome binding site of the expression vector, both of which are functional in the host cell in which the DNA coding sequence for the PspA-like polypeptide is to be expressed. In the preferred practice of the invention, the promoter-operator region is placed in the same sequential orientation with respect to the ATG start codon of DNA sequence encoding the PspA-like polypeptide as the promoter-operator occupies with respect to the ATG-start codon of the gene from which it was derived. Synthetic or modified promoter-operator regions such as the *tac* promoter are well known in the art. When employing such synthetic or modified promoter-operator regions they should be oriented with respect to the ATG start codon of the DNA sequence coding for the PspA-like polypeptide as directed by their creators.

The present invention relates to the alteration of a nucleotide sequence such that internal start sites within the nucleotide sequence are substituted, deleted or otherwise altered to eliminate the internal ATG start sites. Elimination of the internal start sites enables the expression of a single polypeptide rather than multiple polypeptides from a single nucleotide sequence. The single polypeptide expressed can be a fusion protein, can contain a signal sequence or a truncated form of the native polypeptide.

The DNA sequence for a truncated form of the PspA-like polypeptide of strain Rx1 contains at least one such internal ATG sequence. This sequence appears at codon position 127 of the deduced sequence of this PspA-like polypeptide. See Briles, U.S. Patent No. 5,476,929 Figure 3a. Following cleavage of a "leader sequence" corresponding to 31 amino acids at the amino-terminus of the PspA-like polypeptide, a "mature" sequence, beginning with the codon Glu, corresponds to the expressed polypeptide of interest. An internal start site appears at codon position 96 counting from the first amino acid, Glu, of the native pspA polypeptide after cleavage of the leader sequence. (In figure 1, the ATG appears at codon 97

because the truncated pspA molecule represented in that figure has an additional amino acid at the N terminus which is an artifact of expression in a prokaryotic host).

In a preferred embodiment of the present invention the internal start site within the *pspA* gene or truncated forms of the *pspA* gene that include an internal ATG start site are altered such that when the gene or its truncated forms are expressed, a single modified pspA polypeptide or truncated modified pspA polypeptide is produced.

For example, the internal ATG start site could be deleted such that codons 95 and 97 (coding for lys and ile respectively) would be adjacent after deletion. Alternatively, multiple codons surrounding and including the internal start site could be deleted. For example, the codons 92-100 could be deleted such that in the resultant *pspA* gene the codons 91 and 101 would then be adjacent. The deletion should preferably not affect the three dimensional structure of the molecule such that important immunologic epitopes are altered. In a preferred embodiment of the present invention codons 95-97 are deleted from the pspA gene. In a more preferred embodiment codon 96 is deleted.

Alternatively, the internal ATG start codon could be substituted by a different codon that is not recognized by polymerases as a start site for transcription initiation. Therefore any of the 60 codons that are not ATG or stop codons (TGA, TAA, TAG) could be used to replace the ATG internal start site. Therefore the polypeptide produced from such a recombinant DNA molecule could be expressed to produce a recombinant form of the polypeptide wherein the internal met encoded by ATG is replaced with one of nineteen other naturally occurring amino acids.

Preferred embodiments include codons that code for amino acids that share similar size and charge properties as methionine, such as isoleucine, valine, leucine, and phenylalanine.

In general, prokaryotes are used for cloning of DNA sequences in constructing the vectors useful in the invention. For example, *E. coli* K12 strain 294 (ATCC No. 31446) is particularly useful. Other microbial strains which may be used include *E. coli B* and *E. coli* X1776 (ATCC No. 31537), *E. coli* W3110 (phototrophic, ATCC No. 27325), bacilli such as *Bacillus subtilis*, and other enterobacteriaceae such as *Salmonella typhimurium* or *Serratia marcescans*, and various pseudomonas species may be used. Promoters suitable for use with prokaryotic hosts include the β-lactamase (vector pGX2907 [ATCC 39344] contains the replicon and b-lactamase gene) and lactose promoter systems (Chang *et al.*, [1978] Nature 275:615; and Goeddel *et al.*, [1979] Nature 281:544), alkaline phosphatase, the tryptophan (*trp*) promoter system (vector pATH1 [ATCC 37695] is designed to facilitate expression of an open reading frame as a trpE fusion protein under control of the *trp* promoter) and hybrid promoters such as the tac promoter (isolatable from plasmid pDR540 ATCC-37282). However, other functional bacterial promoters, whose nucleotide sequences are generally known, enable one of skill in the art to ligate them to DNA encoding PspA-like polypeptides using linkers or adapters to supply any required restriction sites. Promoters for use in bacterial systems also will contain a Shine-Dalgarno sequence operably linked to the DNA encoding the PspA-like polypeptide. These examples are illustrative rather than limiting.

While the discussion above and the examples provided herein refer to prokaryotic expression, those having skill in the art can readily appreciate that the PspA-like polypeptides of the instant invention may also be recombinantly produced in eukaryotic expression systems.

Host cells may be transformed with the expression vectors of this invention and cultured in conventional nutrient media modified as is appropriate for inducing promoters, selecting transformants or amplifying genes. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be

apparent to the ordinarily skilled artisan. The techniques of transforming cells with the aforementioned vectors are well known in the art and may be found in such general references as Maniatis, *et al.* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York or Current Protocols in Molecular Biology (1989) and supplements.

The PspA-like polypeptide of the present invention may be made either by direct expression or as fusion protein comprising the PspA-like polypeptide followed by enzymatic or chemical cleavage. It is often observed in the production of certain peptides in recombinant systems that expression as a fusion protein prolongs the lifespan and/or increases the yield of the desired peptide. A variety of peptidases (e.g. trypsin) which cleave a polypeptide at specific sites or digest the peptides from the amino or carboxy termini (e.g. diaminopeptidase) of the peptide chain are known. Furthermore, particular chemicals (e.g. cyanogen bromide) will cleave a polypeptide chain at specific sites. The skilled artisan will appreciate the modifications necessary to the amino acid sequence (and synthetic or semi-synthetic coding sequence if recombinant means are employed) to incorporate site-specific internal cleavage sites. See e.g., Carter P. Site Specific Proteolysis of Fusion Proteins, Ch. 13 in Protein Purification: From Molecular Mechanisms to Large Scale Processes, American Chemical Soc., Washington, D.C. (1990).

The present invention provides an immunogenic, immunological or vaccine composition containing recombinant polypeptides derived from pneumococcal strain(s), and a pharmaceutically acceptable carrier or diluent. An immunological composition containing the PspA-like polypeptides elicits an immunological response – local or systemic. The response can, but need not be, protective. An immunogenic composition containing the PspA-like polypeptides likewise elicits a local or systemic immunological response which can, but need not be, protective. A vaccine composition elicits a local or systemic protective response.

Accordingly, the terms "immunological composition" and "immunogenic composition" include a "vaccine composition" (as the two former terms can be protective compositions).

The invention therefore also provides a method of inducing an immunological response in a host mammal comprising administering to the host an immunogenic, immunological or vaccine composition comprising a recombinant PspA-like polypeptide and a pharmaceutically acceptable carrier or diluent.

The determination of the amount of antigen, e.g., PspA. PspA-like polypeptide or truncated portion thereof and optional adjuvant in the inventive compositions and the preparation of those compositions can be in accordance with standard techniques well known to those skilled in the pharmaceutical or veterinary arts. In particular, the amount of antigen and adjuvant in the inventive compositions and the dosages administered are determined by techniques well known to those skilled in the medical or veterinary arts taking into consideration such factors as the particular antigen, the adjuvant (if present), the age, sex, weight, species and condition of the particular patient, and the route of administration. For instance, dosages of particular PspA-like polypeptide antigens for suitable hosts in which an immunological response is desired, can be readily ascertained by those skilled in the art from this disclosure, as is the amount of any adjuvant typically administered herewith. Thus, the skilled artisan can readily determine the amount of antigen and optional adjuvant in compositions and to be administered in methods of the invention. Typically, an adjuvant is commonly used as 0.001 to 50wt% solution in phosphate buffered saline, and the antigen is present on the order of micrograms to milligrams, such as about 0.0001 to about 5 wt%, preferably about 0.0001 to about 1 wt%, most preferably about 0.0001 to about 0.05 wt% (see, e.g., Examples below or in applications cited herein).

Typically, however, the antigen is present in an amount on the order of micrograms to milligrams, or, about 0.001 to about 20 wt%, preferably about 0.01 to about 10 wt%, and most preferably about 0.05 to about 5 wt%.

Of course, for any composition to be administered to an animal or a human, including the components thereof, and for any particular method of administration, it is preferred to determine therefor: toxicity, such as by determining the lethal dose (LD) and LD_{50} in a suitable animal model e.g., rodent such as mouse; and, the dosage of the composition(s), concentration of components therein and timing of administering the composition(s), which elicit a suitable immunological response, such as by titrations of sera and analysis thereof for antibodies or antigens, e.g., by ELISA and/or RFFIT analysis. Such determinations do not require undue experimentation from the knowledge of the skilled artisan, this disclosure and the documents cited herein. And, the time for sequential administrations can be ascertained without undue experimentation.

Those skilled in the art will recognize that the components of the compositions should be selected to be chemically inert with respect to the PspA-like polypeptide antigen and optional adjuvant. This will present no problem to those skilled in chemical and pharmaceutical principles, or problems can be readily avoided by reference to standard texts or by simple experiments, from this disclosure and the documents cited herein. Standard texts, such as "REMINGTON'S PHARMACEUTICAL SCIENCE", 17th edition, 1985, incorporated herein by reference, may be consulted to prepare suitable preparations of PspA polypeptides of the present invention in a form suitable for administration to animals or humans, without undue experimentation. A pharmaceutically acceptable preservative can be employed to increase the shelf-life of the compositions.

Compositions can be administered in dosages and by techniques well known to those skilled in the medical and veterinary arts taking into consideration such factors as the age, sex,

weight, and condition of the particular patient or animal, and the composition form used for administration (e.g., solid vs. liquid). Dosages for humans or other mammals can be determined without undue experimentation by the skilled artisan, from this disclosure and from the documents cited herein.

Suitable regimes for initial administration and booster doses or for sequential administrations also are variable, may include an initial administration followed by subsequent administrations; but nonetheless, may be ascertained by the skilled artisan, from this disclosure and the documents cited herein.

Methods of treating or preventing infection by immunization of susceptible or infected subjects with PspA-like polypeptides have been described above. Such methods induce "active" protection in which the subject, following immunization, mounts an immune response to PspA-like polypeptides which is capable of hastening recovery from, or preventing disease caused by infection with *S. pheumoniae*. It is also possible to treat infected subjects "passively". Passive protection entails the production of antibodies in one host, followed by administration of those antibodies to a second, infected host. In the context of *S. pheumoniae* infection, for example, the PspA-like polypeptides of the present invention can be used to raise antibodies to PspA. Such antibodies could then be administered to infected individuals to hasten recovery from infection.

The following Examples are provided for illustration and are not to be considered a limitation of the invention.

EXAMPLES

Example 1:
Construction of PARx1-MI expression vector pPARx1-MI

(A) Generation of recombinant host cells containing the PARX1-MI coding sequence

pPARx1-MI is a recombinant construct encoding the N-terminal region (aal-314) of the PspA polypeptide of Rx1. This sequence was engineered to change aa96 (ATG-Met) to ATA-Ile using mutagenesis by the polymerase chain reaction (PCR) technique. Two PCR reactions were performed to introduce the sequence change, each using pKSD1014 plasmid DNA containing Rx1 PspA as the template. The first reaction used a pair of Rx1 specific oligonucleotide primers (5'-CCCTAGCATCTG**CATATG**GAAGAATCTCCCGT-AGCC-3' SEQ ID NO.: 1 and 5'-AGCTTCATCTATTATCTTATCTGCTGCGTC-3' SEQ ID NO: 2) to prepare a sequence corresponding to aa1-100 of Rx1 PspA. The second reaction used a pair of Rx1 specific oligonucleotide primers (5'-GACGCAGCAGATAAGATAATAGATGA-AGCT-3' SEQ ID NO: 3 and 5'-CGCAT**GGATCC**TTAAGGAGCCGGCGCTGGTTTTGG-3' SEQ ID NO: 4) to prepare a sequence corresponding to aa91 to 314 of Rx1 PspA. Restriction sites for NdeI and BamH1, respectively, are highlighted in bold.

The products of these two PCR reactions were then purified and used as templates to construct the full length (0.95 kb) sequence encoding a modified form of aa1-314 of Rx1-MI in which Ile has been substituted for Met at codon position 96. This third PCR reaction used two specific oligonucleotide primers (5'-CCCTAGCATCTG**CATATG**GAAG-AATCTCCCGTAGCC-3' SEQ ID NO: 1 and 5'-CGCAT**GGATCC**TTAAGGAGCCGGCG-CTGGTTTTGG-3' SEQ ID NO: 4. This construct was generated such that 5' and 3' ends (with respect to the PspA sense sequence) were bracketed with Nde1 and BamH1 restriction site sequences (noted in bold in the sequences above). A stop codon was introduced immediately upstream from the BamH1 restriction site. The product was purified by agarose

gel electrophoresis and digested with Nde1 and BamH1 restriction enzymes. The restricted DNA was purified and ligated with Nde1-BamH1 digested pET-9a (Novagen).

The ligated DNA was transformed into Subcloning Efficiency $DH5_{\alpha}$ competent cells (Gibco BRL) and plated for isolated colonies. Plasmid DNA was recovered from isolated colonies by conventional methods and analyzed to confirm the presence of the construct of interest. A large-scale plasmid purification was performed from one such colony and the construct sequenced using the dideoxy sequencing method. This plasmid, designated pPARx1-MI, contained a gene fragment identical to the amino-terminal 314 aa of the wild type-Rx1 sequence except for an ATG to ATA change at amino acid position 96.

Example 2:

Expression and characterization of PARx1-MI.

pPARx1-MI was transformed into BL21 (DE3) competent cells (Novagen). Transformed cells were grown to mid-exponential phase and induced with 0.5mM IPTG (Sigma, Inc.). Cells were harvested approximately 2 hours after induction by centrifugation. Cell pellets were resuspended in a PBS-lysozyme lysis buffer (0.5 mg/ml lysozyme in PBS). The cells were lysed by freeze thaw and the lysate was centrifuged.

The supernatant was decanted and adjusted to 70% ammonium sulfate (J.T. Baker) to precipitate proteins. This solution was centrifuged and the supernatant was discarded. The pellet was resuspended in 30% ammonium sulfate and centrifuged. The supernatant was extensively dialyzed with a buffer consisting of 100mM sodium acetate, 1mM EDTA, 10mM NaCl, pH 4.5. The dialysate was centrifuged and the supernatant was applied to a Q Sepharose Fast Flow (Pharmacia Biotech) column. The column flow through and wash were collected and dialyzed against 50mM Tris, 2 mM EDTA, 10mM NaCl, pH 7.5. The dialysate was centrifuged and the supernatant was applied to a Q Sepharose Fast Flow column. The

column was washed with 50mM Tris, 2mM EDTA, 100mM NaCl, pH 7.5, then PARx1-MI was eluted with 50mM Tris, 2mM EDTA, 200 mM NaCl, pH 7.5. This elute was analyzed by SDS-PAGE; proteins were visualized by Coomassie Blue staining and PspA products were identified by Western immunoblot analysis with a monoclonal anti-PspA antibody (XiR278).

Two major bands, migrating at approximately 50 kD and 35 kD, respectively, were observed in lanes loaded with PARx1. In contrast, lanes corresponding to PARx1-MI contained a single major band migrating at approximately 50kD, corresponding to the aa1-314 PspA molecule. The identity of this band was confirmed by Western immunoblot analysis. The immunoblot was probed with monoclonal anti-PARx1 followed by phosphate-labeled goat anti-mouse IgG, then developed with BCIP/NBT. PARx1 contains two bands, migrating at approximately 50 kD and 35 kD, respectively, corresponding to the aa1-314 PspA molecule and the 97-314 form resulting from initiation of translation at position 97. Substitution of Ile for Met at this position in PARx1-MI results in a substantially homogeneous, single form of PspA migrating at approximately 50 kD.

Certain physical properties of PARx1-MI were also examined and compared to those of PARx1. In particular, circular dichroism (CD) and sedimentation analyses were performed. The CD spectropolarimetric method is routinely used to examine the secondary structure or conformation of proteins. The CD technique is based upon the way optically active substances absorb right or left circularly polarized light. See, e.g., Yant, J.T., *et al.*, (1986). Methods in Enzymol. 130:208-269 and Change, C.T., *et al.*, Anal. Biochem. (1978) 91:13-31.

The percentages of the protein secondary-structures could not be determined for the PARx1 preparation because this preparation contains two different forms of the PspA molecule. Analysis of PARx1-MI revealed that the protein has 61.5±5% alpha-helix, no beta-sheet and 37.5% unordered form. These results compare favorably to the predicted structure

of the aa1-314 form of PARx1, indicating that the Met-Ile substitution did not result in any pronounced secondary structure differences. This finding is consistent with the results of immunological analyses, described hereinbelow, which demonstrate that no major immundominant epitopes were destroyed by the substitution, and no new immunodominant epitopes were created.

The sedimentation coefficient of PARx1-MI was determined by velocity sedimentation analysis using conventional techniques. See, e.g., Ralson, G. (1993), Introduction to Analytical Ultracentrifugation, p. 1-87, Beckman Instruments, Inc., Fullerton, CA; Stafford, W.F. III. (1992) Anal. Biochem/ 203:295-301. The sedimentation coefficient calculated for PARx1-MI was 2.15S. The value obtained for PARx1 was 2.13. The results of this analysis indicate that the three-dimensional structure of these molecules is comparable, again consistent with the results of the immunological analyses described hereinbelow.

Example 3:

Immunological characterization of PARx1-MI product

(A) Analysis of PARx1 and PARx1-MI by polyclonal anti-PARx1

To determine whether PARx1-MI retained all the immunodominant epitopes presented by PARx1 the following experiment was performed. Competitive inhibition of anti-PARx1 binding to PARx1 antigen was analyzed using a BIAcore sensory chip coated with PARx1 antigen. Rabbit polyclonal anti-PARx1 (1200 ng/ml) was allowed to react to the chip either alone or in the presence of increasing concentration of PARx1 antigen or PARx1-MI antigen. The concentration of uninhibited antibody able to bind to the PARx1 antigen on the sensory chip surface was measured using mass transport measurements on the BIAcore instrument. The mouse monoclonal IgG anti-PspA antibody, p81-122F10.A11, was used as a standard for these measurements.

Comparison of the inhibition curves generated by this analysis indicated that there was no significant difference in immunological activity of PARx1-MI relative to the wild-type PARx1 antigen. Thus, the substitution of Ile for Met at position 96 in PARx1 had no demonstrable effect on the immunodominant epitopes within PARx1.

(B) Analysis of PARx1 and PARx1-MI by polyclonal anti-PARx1-MI

To determine whether the amino acid substitution within PARx1-MI (Ile for Met at position 96) created any new immunodominant epitopes the following experiment was performed. Rabbit polyclonal antibody was generated following immunization with PARx1-MI. This polyclonal anti-PARx1-MI (1200 ng/ml) was allowed to react with a BIAcore sensory chip coated with PARx1-MI antigen, either alone or in the presence of increasing concentration of PARx1 antigen or PARx1-MI antigen. The concentration of uninhibited antibody able to bind to the PARx1 antigen on the sensory chip surface was measured using mass transport measurements on the BIAcore instrument. As before, the mouse monoclonal IgG anti-PspA antibody P81-122F10.A11 was used as a standard control.

The inhibition curves resulting from this analysis revealed no significant difference between the two recombinant PspA products. Thus, the amino acid substitution within PARx1-MI did not result in the generation of any new immunodominant epitopes not present on PARx1.

Having thus described in detail certain preferred embodiments of the present invention, it is to be understood that the invention defined by the appended claims is not to be limited by particular details set forth in the above description, as many apparent variations thereof are possible without departing from the spirit or scope thereof.

REFERENCES

1. Center of Disease Control. 19484. Pneumococcal polysaccharide vaccine usage, United States, MMWR 33:273-276, 281.

2. Mufson, M.A., G. Oley and D. Hughey, 1982. Pneumococcal disease in a medium-sized community in the United States. JAMA 248:1486-1489.

3. Hook, E.W., C.A. Horton and D.R. Schaberg. 1983. Failure of intensive care unit support to influence mortality from pneumococcal bacteremia. JAMA 249:1055-1057.

4. Breiman, R.F., J.S. Spika, V.J. Navarro, P.M. Darden and C.P. Darby. 1990. Pneumococcal bacteremia in Charleston County, South Carolina. Arch. Intern. Med. 150:1401-1405.

5. Afessa, B., W.L. Greaves and W.R. Frederick. 1995. Pneumococcal bacterimia in adults: a 14-year experience in an inner-city university hospital. Clin. Infec. Diseases 21:345-351.

6. Fang, G.D., M. Fine, J. Orloff, D. Arisumi, V.L. Yu, W. Kapoor, J.T., Grayston, S.P. Wang, R. Kohler, R.R. Muder and *et al.* 1990. New and emerging etiologies for community-acquired pneumonia with implications for therapy. A prospective multicenter study of 359 cases. Medicine (Baltimore) 69: 307-316.

7. Marrie, T.J., H. Durant and L. Yates. 1989. Community-acquired pneumonia requiring hospitalization: 5-year prospective study. Rev. Infect. Dis. 11:586-599.

8. Torres, A., J. Serra-Batlles, A. Ferrer, P. Jimeniz, R. Celis, E. Cobo and R. Rodriquez-Roisin. 1991. Severe community-acquired pneumonia. Epidemiology and prognostic factors. Am. Rev. Respir. Dis. 144:312-318.

9. Bluestone, C.D., J.S. Stephenson and L.M. Martin. 1992. Ten-year review of otitis media pathogens. Pediatr, Infect. Dis. J. 11:S7-11.

10. Teele, D.W., J.O. Klein, B. Rosner and G.B.O.M.S. Group. 1989. Epidemiology of otitis media during the first seven years of life of children in greater Boston: a prospective cohort study. J. Infect. Dis. 160:83-94.

11. Schutze, G.E., S.L. Kaplan and R.F. Jacobs. 1994. Resistant pneumococcus: a worldwide problem. Infection 22:233-237.

12. Privitera, G. 1994. Penicillin resistance among Streptococcus pneumoniae in Europe. Diagnostic Microbiology and Infectious Disease 19:157-161.

13. Bizzozero, O.G., Jr. and V.T. Andriole. 1969. Tetracycline-resistant pneumococcal infection. Incidence, clinical presentation and laboratory evaluation. Arch. Intern. Med. 123:388-393.

14. Workman, M.R., M. Layton, M. Hussein, J. Philpott-Howard and R.C. George. 1993. Nasal carriage of penicillin-resistant pneumococcus in sickle cell patients (letter). Lancet 342:746-747.

15. Koornhof, H.J., A Wasas and K. Klugman. 1992. Antimicrobial resistance in Streptococcus pneumoniae: a South African perspective. Clin. Infect. Dis. 15:84-94.

16. Dagan, R., P. Yagupsky, A. Golbart, A. Wasas and K. Klugman. 1994. Increasing prevalence of penicillin-resistant pneumococcal infections in children in southern Israel: implications for future immunization policies. Pediatr. Infect. Dis. J. 13:782-786.

17. Reichler, M.R., J. Rakovsky, A. Sobotova, M. Slacikova, B. Hlavacova, B. Hill, L. Krajcikova, P. Tarina, R.R. Facklam and R.F. Breiman. 1995. Multiple antimicrobial resistance of pneumococci in children with otitis media, bacteremia, and meningitis in Slovakia. J. Infect. Dis. 171:1491-1496.

18. Freidland, I.R., S. Shelton, M. Paris, S. Rinderknecht, S. Ehrett, K. Krisher, and G.H. McCracken, Jr., 1993. Dilemmas in diagnosis and management of cephalosporin-resistant Streptococcus pneumoniae meningitis. Pediatr. Infect. Dis. J. 12:196-200.

19. Fedson, D.S., and D.M. Musher. 1994. Pneumococcal Vaccine. In Vaccines. S.A. Plotkin and J.E.A. Montimer, Eds. W.B. Saunders Co., Philadelphia, PA, p. 517-564.

20. Takala, A.K., J. Eskola, M. Leinonen, H. Kayhty, A. Nissinen, E. Pekkanen and P.H. Makela. 1991. Reduction of oropharyngeal carriage of Haemophilus influenzae type b (Hib) in children immunized with an Hib conjugate vaccine. J. Infect. Dis. 164:982-986.

21. Takala, A.K., M. Santosham, J. Almeido-Hill, M. Wolff, W. Newcomer, R. Reid, H. Kayhty, E. Esko and P.H. Makela. 1993. Vaccination with Haemophilus influenzae type b meningococcal protein conjugate vaccine reduces oropharyngeal carriage of Haemophilus influenzae type b among American Indian children. Pediatr. Infect. Dis. J. 12:593-599

22. Ward, J., J.M. Lieberman and S.L. Cochi. 1994. Haemophilus influenzae vaccines. In Vaccines. S.A. Plotkin and J.E.A. Montimer, Eds. W.B. Saunders Co., Philadelphia, PA, p. 337-386.

23. Murphy, T.V., P. Pastor, F. Medley, M.T. Osterholm, and D.M. Cranoff. 1993. Decreased Haemophilus colonization in children vaccinated with Haemophilus influenzae type b conjugate vaccine. J. Pediatr. 122:517-523.

24. Mohle-Boetani, J.C., G. Ajello, E. Breneman, K.A., Deaver, C. Harvey, B.D. Plikaytis, M.M. Farley, D.S. Stephens and J.D. Wenger. Carriage of Haemophilus influenzae type b in children after widespread vaccination with conjugate Haemophilus influenzae type b vaccines. Pediatr. Infect. Dis. J. 12:589-593.

25. Watson, D.A. and D.M. Musher. 1990. Interruption of capsule production in Streptococcus pneumonia serotype 3 by insertion of transposon Tn916. Infect. Immun. 58:135-138.

26. Avery, O.T. and R., Dubos. 1931. The protective action of specific enzyme against type III pneumococcus infection in mice. J. Exp. Med. 54:73-89.

27. Alonso DeVelasco, E., A.F.M. Verheul, J. Verhoef and H. Snippe. 1995. Streptococcus pneumoniae; virulence factors, pathogenesis and vaccines. Microbiological Reviews 59:591-603.

28. Butler, J.C., R.F. Breiman, J.F. Campbell, H.B. Lipman, C.V. Broome and R.R. Facklam. 1993. Pneumococcal polysaccharide vaccine efficacy. An evaluation of current recommendations. JAMA 270:1826-1831.

29. Hirschmann, J.V., and B.A. Lipsky. 1994. The pneumococcal vaccine after 15 years of use. Arch. Intern., Med. 154:373-377.

30. Briles, D.E., J. Yother and L.S. McDaniel. 1988. Role of pneumococcal surface protein A in the virulence of Streptococcus pneumoniae. Rev. Infect. Dis. 10:S372-4.

31. Talkington, D.F., D.C. Voellinger, L.S. McDaniel and D.E. Briles. 1992. Analysis of pneumococcal PspA microheterogeneity in SDS-polyacrylamide gels and the association of PspA with the cell membrane. Microb. Pathogen. 13:343-355.

32. Yother, J. and D.E. Briles. 1992. Structural properties and evolutionary relationships of PspA, a surface protein of Streptococcus pneumoniae, as revealed by sequence analysis. J. Bacteriol. 174:601-609.

33. Yother, J. and J.M. White. 1994. Novel surface attachment mechanism of the Streptococcus pneumoniae protein PspaA. J. Bacteriol. 176:2976-85.

34. McDaniel, L.S., B.A. Ralph, D.O. McDaniel and D.E. Briles. 1994. Localization of protection-eliciting epitopes of PspA of Streptococcus pneumoniae between amino acids residues 192 and 260. Microb. Pathog. 17:323-337.

35. Ralph, B.A., D.E. Briles and L.S. McDaniel. 1994. Cross-reactive protection eliciting epitopes of pneumoccal surface protein A. Ann NY Acad. Sci. 730:361-3.

36. Waltman, W.D., L.S. McDaniel, B. Anderson, L. Bland, B.M. Gray, C.S. Eden and D.E. Briles. 1988. Protein serotyping of Streptococcus pneumoniae based on reactivity to six monoclonal antibodies. Microb. Pathog. 5:159-67.

SEQUENCE LISTING

(1) GENERAL INFORMATION

(i) APPLICANT: BECKER, ROBERT S
GRAY, MARY-ANN L
PYLE, DEREK J

(ii) TITLE OF INVENTION: Improved Method of Production of Pneumococcal Surface Proteins

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: CONNAUGHT LABORATORIES, INC.
(B) STREET: ROUTE 611, PO BOX 187
(C) CITY: SWIFTWATER
(D) STATE: PENNSYLVANIA
(E) COUNTRY: USA
(F) ZIP: 18370-0187

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER; IBM PC compatible
(A) OPERATING SYSTEML PC-DOS/MS-DOS
(B) SOFTWARE: PatentIn Release #1.0, Version #1.2

(vi) CURRENT APPLICATION DATA;
(A) APPLICATION NUMBER;
(B) FILING DATE: 16-SEP-1997
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: HOWE, TIMOTHY R.
(B) REGISTRATION NUMBER: 39,228
(C) REFERENCE/DOCKET NUMBER: TH-002

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (717) 839-5027
(B) TELEFAX: (717) 839-0619

(2) INFORMATION FOR SEQ ID NO.: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 36 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(x) SEQUENCE DESCRIPTION: SEQ ID NO.: 1:

CCCTAGCATC TGCATATGGA AGAATCTCCC GTAGCC
36

(3) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

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

AGCTTCATCT ATTATCTTAT CTGCTGCGTC
30

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 30 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO
(xii) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GACGCAGCAG ATAAGATAAT AGATGAAGCT
30

(2) INFORMATION FOR SEQ ID NO.: 4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO.

(iv) ANTI-SENSE: NO

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

CGCATGGATC CTTAAGGAGC CGGCGCTGGT TTTGG
35

(4) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 945 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

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

ATGGAAGAAT CTCCCGTAGC CAGTCAGTCT AAAGCTGAGA AGACTATGA TGCA GCGAAG 60

AAAGATGCTA AGAATGCGAA AAAAGCAGTA GAAGATGCTC AAAAGGCTTT AGAT GATGCA 120

AAAGCTGCTC AGAAAAAATA TGACGAGGAT CAGAAGAAAA CTGAGGAGAA AGCC GCGCTA 180

GAAAAAGCAG CGTCTGAAGA GATGGATAAG GCAGTGGCAG CAGTTCAACA AGCG TATCTA 240

GCCTATCAAC AAGCTACAGA CAAAGCCGCA AAAGACGCAG CAGATAAGAT AATA GATGAA 300

GCTAAGAAAC GCGAAGAAGA GGCAAAAACT AAATTTAATA CTGTTCGAGC AATG GTAGTT 360

CCTGAGCCAG AGCAGTTGGC TGAGACTAAG AAAAAATCAG AAGAAGCTAA ACAA AAAGCA 420

CCAGAACTTA CTAAAAAACT AGAAGAAGCT AAAGCAAAAT TAGAAGAGGC TGAG AAAAAA 480

GCTACTGAAG CCAAACAAAA AGTGGATGCT GAAGAAGTCG CTCCTCAAGC TAAA ATCGCT 540

GAATTGGAAA ATCAAGTTCA TAGACTAGAA CAAGAGCTCA AAGAGATTGA TGAG TCTGAA 600

TCAGAAGATT ATGCTAAAGA AGGTTTCCGT GCTCCTCTTC AATCTAAATT GGAT GCCAAA 660

AAAGCTAAAC TATCAAAACT TGAAGAGTTA AGTGATAAGA TTGATGAGTT AGAC GCTGAA 720

ATTGCAAAAC TTGAAGATCA ACTTAAAGCT GCTGAAGAAA ACAATAATGT AGAA GACTAC 780

TTTAAAGAAG GTTTAGAGAA AACTATTGCT GCTAAAAAAG CTGAATTAGA AAAA ACTGAA 840

GCTGACCTTA AGAAAGCAGT TAATGAGCCA GAAAAACCAG CTCCAGCTCC AGAA ACTCCA 900

GCCCCAGAAG CACCAGCTGA ACAACCAAAA CCAGCGTCCGG CTCCT 945

WE CLAIM:

1. An isolated DNA sequence encoding a pneumococcal surface protein A (PspA) molecule wherein internal, naturally occurring translational initiation sites have been modified or eliminated, such that expression of said DNA sequence results in a single form of PspA

2. The DNA sequence of claim 1, wherein said DNA sequence was derived from strain Rx1 of *S. pneumoniae.*

3. The DNA sequence of claim 1, wherein said DNA sequence is as set forth in SEQ. ID No.: 5.

4. A recombinant construct, comprising the DNA sequence of claim 1 operably linked to an expression control sequence.

5. A unicellular host transformed with the recombinant construct of claim 1.

6. The unicellular host of claim 5, wherein said host is a prokaryote.

7. The unicellular host of claim 6, wherein said prokaryote is *E. coli.*

9. (Amended) A PspA composition characterized in that the PspA is essentially all 50 kD PspA.

10. An immunological composition comprising a therapeutically effective amount of the PspA-like polypeptide of claim 9 and a pharmaceutically acceptable carrier.

11. The immunological composition of claim 10, further comprising an adjuvant.

AMENDED SHEET

12. A method of treating or preventing *S. pneumoniae* infection in a susceptible animal, which method comprises inoculation with an immunoeffective amount of the immunological composition of claim 11.

13. The method of claim 12, wherein said animal is human.

14. The method of claim 13, wherein said human is a child.

15. A method of producing antibodies reactive with PspA, which method comprises vaccination of an animal with an immunoeffective amount of the immunological composition of claim 11.

18. The DNA sequence of claim 1, wherein the internal, naturally occurring translational initiation site at position 96 (met codon) has been replaced by a codon from the group consisting of ATA, ATC, ATT, TTG, CTG or GTG.

19. The DNA sequence of claim 18, wherein the met codon (ATG) at position 96 has been replaced by ATA (ile).

20. The PspA of claim 9, wherein said PspA is comprised of amino acids 1-115 of the mature PspA molecule and in which isoleucine has been substituted for methionine at position 96.

```
1
ATG GAA TCT CCC GTA GCC AGT CAG TCT AAA GCT GAG AAA GAC TAT
met glu ser pro val ala ser gln ser lys ala glu lys asp tyr
   10                  20                  30
GAT GCA GAA AAG AAA GAT GCT AAG GCG AAT GCG AAA GCA GTA GAA GAT
asp ala glu lys lys asp ala lys ala asn ala lys ala val glu asp
   40                  50                  60
GCT CAA AAG AAG GCT TTA GAT GCA AAG GCT CAG AAA GCT GCT CAG GCG
ala gln lys lys ala leu asp ala lys ala gln lys ala ala gln ala
   70                  80                  90
GAG GAT GAG ATG GAT AAG GCA GTG GAG GCA GCA GTT CAA CAA GCG TAT CTA
glu asp glu met asp lys ala val glu ala ala val gln gln ala tyr leu
  100                 110                 120
GCC TAT CAA CAA GCT ACA GAC AAA GCC GCA GCA AAA GAC AAG GCT ATG
ala tyr gln gln ala thy asp lys ala ala ala lys asp lys ala met
  130                 140                 150
ATA GAA GAG GCA AAA ACT AAA TTT GAA GAA GAG GCA AAA ACT AAA TTT
ile glu glu ala lys thr lys phe glu glu glu ala lys thr lys phe
  160                 170                 180
AAT ACT GTT CGA GCA ATG GTA GTT CCT GAG CCA GAG CAG TTG GCT GAG
```

FIG. 1A

CA 02304128 2000-03-17

SUBSTITUTE SHEET (RULE 26)

```
asn thr val arg ala met val val pro glu pro glu gln leu ala glu

    130                                                 140
ACT AAG AAA AAA TCA GAA GAA GCT AAA CAA AAA GCA CCA GAA CTT ACT
thr lys lys lys ser glu glu ala lys gln lys ala pro glu leu thr

                    150                                         160
AAA AAA CTA GAA --U GCT AAA GCA AAA TTA GAA GAG GCT GAG AAA AAA
lys lys leu glu     ala lys ala lys leu glu glu ala glu lys lys

                                        170
GCT ACT GAA GCC AAA CAA AAA GTG GAT GCT GAA GAA GTC GCT CCT CAA
ala thr glu ala lys gln lys val asp ala glu glu val ala pro gln

            180                                         190
GCT AAA ATC GCT GAA TTG GAA AAT CAA GTT CAT AGA CTA GAA CAA GAG
ala lys ile ala glu leu glu asn gln val his arg leu glu gln glu

                            200
CTC AAA GAG ATT GAT GAG TCT GAA TCA GAA GAT TAT GCT AAA GAA GGT
leu lys glu ile asp glu ser glu ser glu asp tyr ala lys glu gly

    210                                             220
TTC CGT GCT CCT CTT CAA TCT AAA TTG GAT GCC AAA AAA GCT AAA CTA
phe arg ala pro leu gln ser lys leu asp ala lys lys ala lys leu

                    230                                         240
TCA AAA CTT GAA GAG TTA AGT GAT AAG ATT GAT GAG TTA GAC GCT GAA
ser lys leu glu glu leu ser asp lys ile asp glu leu asp ala glu

                                    250
ATT GCA AAA CTT GAA GAT CAA CTT AAA GCT GCT GAA GAA AAC AAT AAT
```

FIG. 1B

ile ala lys leu glu asp gln leu lys ala ala glu glu asn asn

ATA GCA AAG CTT GAC CAG CTT AAG GCT GCA GAA GAA AAT

GTA GAA TAC TTT AAA GAA GGT AAA GAC TAC TTT AAA GAA GGT

val glu tyr phe lys glu gly lys asp tyr phe lys glu gly

260 270

AAA GCT GAA TTA GAA AAA ACT GGT GAA TTA GAG AAA ACT ATT

lys ala glu leu glu lys thr gly glu leu glu lys thr ile

GCT CCA GAA AAA ACT CCA GCT GAA AAA ACT CCA GCT GCT GCT

ala pro glu lys thr pro ala glu lys thr pro ala ala ala

280 290 300

GAG CCA GAA ACT CCA GCT GAA ACT CCA GCT GAA ACT CCA GCT

glu pro glu thr pro ala glu thr pro ala glu thr pro ala

CCA GCT CCT CCG GCT CCT GCT CCA GCT CCT CCG GCT CCT GCT

pro ala pro pro ala pro ala pro ala pro pro ala pro ala

310

GAA CCA GCG GAA CAA CCA AAA GAA CAA CCA GCG GAA CAA CCA

glu pro ala glu gln pro lys glu gln pro ala glu gln pro

AAA GAA GAA CCA GTT GCT AAA GAA GAA CCA GTT GCT AAT GCA

lys glu glu pro val ala lys glu glu pro val ala asn ala

320 330 340

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

350 360 370

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

380 390 400

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

410 420 430

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

440 450 460

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

470 480 490

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

500 510 520

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

530 540 550

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

560 570 580

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

590 600 610

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

620 630 640

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

650 660 670

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

680 690 700

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

710 720 730

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

740 750 760

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

770 780 790

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

800 810 820

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

830 840 850

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

860 870 880

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

890 900 910

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

920 930 940

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

950 960 970

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

980 990 1000

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1010 1020 1030

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1040 1050 1060

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1070 1080 1090

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1100 1110 1120

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1130 1140 1150

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1160 1170 1180

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1190 1200 1210

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1220 1230 1240

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1250 1260 1270

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1280 1290 1300

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1310 1320 1330

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1340 1350 1360

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1370 1380 1390

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1400 1410 1420

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1430 1440 1450

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1460 1470 1480

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1490 1500 1510

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1520 1530 1540

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1550 1560 1570

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1580 1590 1600

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1610 1620 1630

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1640 1650 1660

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1670 1680 1690

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1700 1710 1720

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1730 1740 1750

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1760 1770 1780

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1790 1800 1810

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1820 1830 1840

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1850 1860 1870

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1880 1890 1900

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

1910 1920 1930

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

1940 1950 1960

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

1970 1980 1990

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2000 2010 2020

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2030 2040 2050

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2060 2070 2080

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2090 2100 2110

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2120 2130 2140

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2150 2160 2170

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2180 2190 2200

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2210 2220 2230

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2240 2250 2260

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2270 2280 2290

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2300 2310 2320

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2330 2340 2350

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2360 2370 2380

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2390 2400 2410

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2420 2430 2440

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2450 2460 2470

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2480 2490 2500

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2510 2520 2530

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2540 2550 2560

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2570 2580 2590

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2600 2610 2620

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2630 2640 2650

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2660 2670 2680

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2690 2700 2710

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2720 2730 2740

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2750 2760 2770

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2780 2790 2800

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2810 2820 2830

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2840 2850 2860

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2870 2880 2890

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2900 2910 2920

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

2930 2940 2950

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

2960 2970 2980

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

2990 3000 3010

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

3020 3030 3040

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

3050 3060 3070

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

3080 3090 3100

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

3110 3120 3130

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

3140 3150 3160

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

3170 3180 3190

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

3200 3210 3220

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

3230 3240 3250

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

3260 3270 3280

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

3290 3300 3310

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

3320 3330 3340

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

3350 3360 3370

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

3380 3390 3400

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala glu glu

GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA

glu pro val ala glu glu glu pro val ala glu glu glu pro

3410 3420 3430

GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT

val ala glu glu glu pro val ala glu glu glu pro val ala

GAA GAA GAA CCA GTT GCT GAA GAA GAA CCA GTT GCT GAA GAA

glu glu glu pro val ala glu glu glu pro val ala gl