



(19)
Bundesrepublik Deutschland
Deutsches Patent- und Markenamt

(10) **DE 600 37 413 T2** 2008.06.19

(12) **Übersetzung der europäischen Patentschrift**

(97) **EP 1 180 119 B1**

(21) Deutsches Aktenzeichen: **600 37 413.0**

(86) PCT-Aktenzeichen: **PCT/CA00/00289**

(96) Europäisches Aktenzeichen: **00 910 463.9**

(87) PCT-Veröffentlichungs-Nr.: **WO 2000/055191**

(86) PCT-Anmeldetag: **16.03.2000**

(87) Veröffentlichungstag
der PCT-Anmeldung: **21.09.2000**

(97) Erstveröffentlichung durch das EPA: **20.02.2002**

(97) Veröffentlichungstag
der Patenterteilung beim EPA: **12.12.2007**

(47) Veröffentlichungstag im Patentblatt: **19.06.2008**

(51) Int Cl.⁸: **C07K 14/285** (2006.01)

C12N 15/00 (2006.01)

A61K 38/00 (2006.01)

(30) Unionspriorität:

268347 16.03.1999 US

(73) Patentinhaber:

Sanofi Pasteur Ltd., Toronto, Ontario, CA

(74) Vertreter:

Meyer & Partner GbR, 20354 Hamburg

(84) Benannte Vertragsstaaten:

**AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT,
LI, LU, MC, NL, PT, SE**

(72) Erfinder:

**LOOSMORE, Sheena M., Aurora, Ontario L4G 4R4,
CA; YANG, Yan-Ping, Willowdale, Ontario M2R
3N7, CA; KLEIN, Michel H., Willowdale, Ontario
M2P 1B9, CA**

(54) Bezeichnung: **REKOMBINANTE ADHESIN-PROTEINE AUS HAEMOPHILUS INFLUENZAE**

Anmerkung: Innerhalb von neun Monaten nach der Bekanntmachung des Hinweises auf die Erteilung des europäischen Patents kann jedermann beim Europäischen Patentamt gegen das erteilte europäische Patent Einspruch einlegen. Der Einspruch ist schriftlich einzureichen und zu begründen. Er gilt erst als eingelegt, wenn die Einspruchsgebühr entrichtet worden ist (Art. 99 (1) Europäisches Patentübereinkommen).

Die Übersetzung ist gemäß Artikel II § 3 Abs. 1 IntPatÜG 1991 vom Patentinhaber eingereicht worden. Sie wurde vom Deutschen Patent- und Markenamt inhaltlich nicht geprüft.

Beschreibung**GEBIET DER ERFINDUNG**

[0001] Die vorliegende Erfindung betrifft das Gebiet der Molekulargenetik und insbesondere trunkierte Haemophilus influenzae Adhäsine(Hia)-Proteine.

HINTERGRUND DER ERFINDUNG

[0002] Haemophilus influenzae ist die Ursache von mehreren ernsten Erkrankungen des Menschen wie z. B. Meningitis, Epiglottitis, Septikämie und Otitis media. Es gibt sechs Serotypen von H. influenzae, die mit a bis f bezeichnet werden, welche durch ihr Kapselpolysaccharid identifiziert werden. H. influenzae Typ b (Hib) war bis zu der Einführung von mehreren Hib-Konjugatimpfstoffen in den 1980ern eine Hauptursache von bakterieller Meningitis (Ref. 1. In der gesamten Anmeldung werden verschiedene Referenzen in Klammern angegeben, um den Stand der Technik, zu welchem diese Erfindung gehört, vollständiger zu beschreiben. Die volle bibliographische Information für jedes Zitat findet man am Ende der Beschreibung, unmittelbar vor den Ansprüchen.)

[0003] WO 96/30519 bezieht sich auf Haemophilus Adhäsionsproteine, Nukleinsäuren und abgeleitete Impfstoffe.

[0004] Impfstoffe, die auf H. influenzae Typ b-Kapselpolysaccharid, konjugiert mit Diphtherietoxoid (Ref. 2), Tetanustoxoid (Ref. 3 und US-Patent 4,496,538) oder Neisseria meningitidis Außenmembranprotein (Ref. 4) basieren, waren bei der Linderung einer von H. influenzae Typ b hervorgerufenen Meningitis wirksam. Die anderen Serotypen von H. influenzae sind mit niedrigen Häufigkeiten mit einer invasiven Erkrankung verbunden, auch wenn es eine Zunahme bei dem Auftreten der Erkrankung, die durch diese Stämme verursacht wird, zu geben scheint, wenn das Auftreten der Hib-Erkrankung abnimmt (Ref. 5; Ref. 6). Nicht eingekapselte oder nicht typisierbare H. influenzae (NTHi) sind ebenfalls für einen weiten Bereich menschlicher Erkrankungen einschließlich Otitis media, Epiglottitis, Pneumonie und Tracheobronchitis verantwortlich. Das Auftreten der NTHi-induzierten Erkrankung ist durch die Einführung der Hib-Impfstoffe nicht beeinflusst worden (Ref. 7).

[0005] Otitis media ist die häufigste Krankheit der frühen Kindheit, wobei 60 bis 70% aller Kinder mit einem Alter von weniger als 2 Jahren zwischen einer und drei Ohreninfektionen erleiden (Ref. 8). Chronische Otitis media ist bei Kindern für Hör-, Sprach- und kognitive Schwächen verantwortlich. H. influenzae-Infektionen machen ca. 30% der Fälle von akuter Otitis media und ca. 60% von chronischer Otitis media aus. Allein in den Vereinigten Staaten kostet die Behandlung von Otitis media zwischen 1 und 2 Milliarden Dollar pro Jahr für Antibiotika und operative Verfahren wie z. B. Tonsillektomien, Adenoidektomien und die Einführung von Tympanostomie-Röhrchen. Es wird geschätzt, dass zusätzliche 30 Milliarden \$ pro Jahr für begleitende Therapien wie z. B. Sprachtherapie und spezielle Schulungsklassen ausgegeben werden. Weiterhin werden viele der Erregerorganismen von Otitis media gegen eine Antibiotikabehandlung resistent. Ein wirksamer prophylaktischer Impfstoff gegen Otitis media ist somit wünschenswert.

[0006] Während der natürlichen Infektion durch NTHi sind an der Oberfläche exponierte äußere Membranproteine, welche eine Antikörperantwort stimulieren, potentiell wichtige Targets für bakterizide und/oder schützende Antikörper und daher mögliche Impfstoffkandidaten. Eine Familie von Proteinen mit hohem Molekulargewicht (HMW1 und HMW2), die bei der Anhaftung von NTHi an Epithelzellen wichtig sind, ist bei ca. 70 bis 75% der NTHi-Stämme identifiziert worden (Ref. 9; Ref. 10). Von diesen Adhäsinen mit hohem Molekulargewicht ist gezeigt worden, dass sie in dem Chinchillamodell von Otitis media einen gewissen Schutz liefern (Ref. 11). Eine zweite Familie von Adhäsionsproteinen mit hohem Molekulargewicht ist in ca. 25% der NTHi und in eingekapselten H. influenzae-Stämmen identifiziert worden (Ref. 12; Ref. 13, Ref. 14). Das NTHi-Mitglied dieser zweiten Familie wird Haemophilus influenzae Adhäsine oder Hia genannt, und das homologe Protein, das in eingekapselten Stämmen gefunden wird, wird Haemophilus influenzae Oberflächenfibrillenprotein oder Hsf genannt. Das hia-Gen wurde ursprünglich aus einer Expressionsbibliothek kloniert, indem Rekonvaleszenten-serum aus einem Otitis media-Patienten verwendet wurde, was anzeigt, dass es ein wichtiges Immunogen während der Erkrankung ist.

[0007] Die Prototyp Hia- und Hsf-Proteine zeigen ca. 82% Sequenzähnlichkeit, auch wenn das Hsf-Protein beträchtlich größer ist. Die Proteine umfassen konservierte Amino- und Carboxytermini und mehrere Wiederholungsmotive, wobei Hsf mehr Wiederholungssequenzen als Hia enthält. Ein Protein mit hohem Molekulargewicht (200 kDa) wurde ebenfalls aus Moraxella catarrhalis identifiziert, welches eine gewisse Sequenzho-

mologie mit den Hsf- und Hia-Proteinen aufweist (U.S.-Patent Nr. 5,808,024).

[0008] Da Hia oder Hsf bei eingekapselten Stämmen von *Haemophilus influenzae* und ca. 20 bis 25% der nicht eingekapselten Stämme konserviert ist und gezeigt wurde, dass dieses ein Adhäsin ist, weist das Protein einen Nutzen bei der Diagnose von und einer Impfung gegen eine Erkrankung, die von *H. influenzae* oder anderen bakteriellen Pathogenen, die Hia oder ein Protein erzeugen, das in der Lage ist, Antikörper hervorzurufen, die spezifisch mit Hia reaktiv sind, verursacht wird, auf.

[0009] Ein Nachteil von Hia zur Verwendung als ein Antigen in der Diagnose, zur Erzeugung von anti-Hia-Antikörpern, die in der Diagnose nützlich sind, und als ein Immunogen bei der Impfung ist die geringe Ausbeute des nativen Proteins aus *Haemophilus influenzae*-Spezies.

[0010] Es wäre vorteilhaft, rekombinantes Hia-Protein zur Verwendung als Antigen, in immunogenen Präparationen einschließlich Impfstoffen, als Träger für andere Immunogene und bei der Erzeugung von diagnostischen Reagenzien bereitzustellen.

ZUSAMMENFASSUNG DER ERFINDUNG

[0011] Die vorliegende Erfindung ist auf die Bereitstellung von trunkierten rekombinanten *H. influenzae* Adhäsin(rHia)-Proteinen gerichtet.

[0012] In Verbindung mit der Bereitstellung solcher rekombinanten Proteine stellt die vorliegende Erfindung gewisse isolierte und gereinigte Nukleinsäuremoleküle bereit. Gemäß einem ersten Aspekt der vorliegenden Erfindung wird ein isoliertes und gereinigtes Nukleinsäuremolekül bereitgestellt, welches ein N-trunkiertes *Haemophilus influenzae* Adhäsin(Hia)-Protein eines Stammes von *Haemophilus influenzae* codiert, das in Form von Einschlusskörpern exprimiert wird, wobei das N-trunkierte Protein die Fähigkeit aufweist, gegen eine Besiedlung zu schützen, und dieses das N-trunkierte *Haemophilus influenzae* Adhäsin(Hia)-Protein des *Haemophilus influenzae*-Stamms 33 ist, das an der Aminosäure V38 von SEQ ID No: 24 beginnt.

[0013] Ein solches Nukleinsäuremolekül kann in einen Vektor eingeschlossen werden, welcher ein Plasmidvektor sein kann.

[0014] Unter einem anderen Aspekt der vorliegenden Erfindung wird ein Vektor zum Transformieren eines Wirtes bereitgestellt, welcher ein Nukleinsäuremolekül, welches das N-trunkierte *Haemophilus influenzae* Adhäsin(Hia)-Protein codiert, und einen Promotor, der operativ mit dem Nukleinsäuremolekül verbunden ist, zur Expression des trunkierten Hia-Proteins und gegebenenfalls das *cer*-Gen von *E. coli* umfasst. Der Vektor kann ein Plasmidvektor sein, welcher die Identifizierungscharakteristika eines Plasmidvektors aufweist, der ausgewählt ist aus der Gruppe, bestehend aus:

DS-2340-2-3 wie in [Fig. 8A](#) gezeigt; und

DS-2448-17 wie in [Fig. 9B](#) gezeigt.

[0015] Die hierin bereitgestellten Vektoren können einen Replikationsvektor, einschließlich eines Vektors aus *Salmonella*, BCG, Adenovirus, Pockenvirus, Vaccinia- oder Poliovirus beinhalten.

[0016] Jeder der hierin bereitgestellten Vektoren kann eingesetzt werden, um eine geeignete Wirtszelle zur Expression eines schützenden *Haemophilus influenzae* Adhäsin (Hia)-Proteins eines nicht typisierbaren Stammes von *Haemophilus*, welches in trunkierter Form vorliegt, in dieser zu transformieren. Ein solcher Wirt kann geeigneterweise *E. coli* sein. Eine solche Expression kann unter der Kontrolle des T7-Promotors stattfinden, und die Expression des rekombinanten Hia aus dem transformierten Wirt kann bewirkt werden, indem in einer induzierenden Konzentration von Lactose oder einem anderen geeigneten Induktionsmittel kultiviert wird.

[0017] Ebenfalls bereitgestellt von der vorliegenden Erfindung wird eine Wirtszelle, die durch den Vektor der vorliegenden Erfindung transformiert ist und ein schützendes N-trunkiertes *Haemophilus influenzae* Adhäsin(Hia)-Protein eines nicht typisierbaren Stammes von *Haemophilus* exprimiert.

[0018] Weiterhin wird von der vorliegenden Erfindung ein rekombinantes schützendes N-trunkiertes Hia-Protein bereitgestellt, welches bei der Aminosäure V38 der SEQ ID No: 24 (*Haemophilus influenzae* Stamm 33) beginnt.

[0019] Die rekombinanten N-trunkierten Hia-Proteine, die hierin bereitgestellt werden, sind als Antigene in im-

munogenen Zusammensetzungen, Träger für andere Immunogene, diagnostische Mittel und bei der Erzeugung von diagnostischen Mitteln nützlich. Die Nukleinsäuremoleküle, welche das N-trunkierte Hia-Protein codieren, sind ebenfalls als Sonden für eine diagnostische Verwendung und ebenfalls in immunogenen Zusammensetzungen nützlich.

[0020] Die vorliegende Erfindung stellt unter einem zusätzlichen Aspekt von dieser eine immunogene Zusammensetzung bereit, welche wenigstens eine immunologisch aktive Komponente, welche ausgewählt ist aus der Gruppe, bestehend aus einem isolierten und gereinigten Nukleinsäuremolekül, wie es hierin bereitgestellt wird, und einem rekombinanten schützenden N-trunkierten Hia-Protein eines Stammes von *Haemophilus*, wie es hierin bereitgestellt wird, und einen pharmazeutisch verträglichen Träger für diese umfasst.

[0021] Die hierin bereitgestellten immunogenen Zusammensetzungen können als ein Impfstoff (eine Vakzine) zur in vivo-Verabreichung an einen Wirt formuliert werden, um einen Schutz vor einer Erkrankung, die von *H. influenzae* verursacht wird, zu bieten. Für einen solchen Zweck können die Zusammensetzungen als eine Mikropartikel-, Kapsel-, ISCOM- oder Liposomenpräparation formuliert werden. Die immunogene Zusammensetzung kann in Kombination mit einem Zielmolekül zur Beförderung zu spezifischen Zellen des Immunsystems oder zu Schleimhautoberflächen bereitgestellt werden.

[0022] Die immunogenen Zusammensetzungen der Erfindung (einschließlich Impfstoffen) können weiterhin wenigstens ein anderes immunogenes oder immunstimulierendes Material umfassen, und das immunstimulierende Material kann wenigstens ein Hilfsmittel oder wenigstens ein Zytokin sein. Geeignete Hilfsmittel beinhalten (sind aber nicht beschränkt auf) Aluminiumphosphat, Aluminiumhydroxid, QS21, Quil A, Derivate und Komponente von diesen, eine ISCOM-Matrix, Calciumphosphat, Calciumhydroxid, Zinkhydroxid, ein Glycolipidanalogen, einen Octadecylester einer Aminosäure, ein Muramyl-dipeptid, Polyphosphazen, ISCOPREP, DC-Chol, DDBA und ein Lipoprotein und andere Hilfsmittel.

[0023] Vorteilhafte Kombinationen von Hilfsmitteln werden in den parallelen U.S.-Patentanmeldungen, Aktenzeichen 08/261,194, angemeldet am 16. Juni 1994, und 08/483,856, angemeldet am 7. Juni 1995, welche der WO 95/34308 entsprechen, die am 21. November 1995 veröffentlicht wurde, beschrieben.

[0024] Die vorliegende Erfindung beinhaltet unter noch einem zusätzlichen Aspekt von dieser ein in vitro-Verfahren zur Erzeugung eines schützenden N-trunkierten *Haemophilus influenzae* Adhäsins(Hia)-Proteins eines nicht typisierbaren Stammes von *Haemophilus influenzae*, welches umfasst:

das Transformieren eines Wirtes, wie z. B. *E. coli*, mit einem Vektor, welcher ein Nukleinsäuremolekül, das eine N-trunkierte Form des *Haemophilus influenzae* Adhäsins-Proteins, wie es hierin bereitgestellt wird, codiert, umfasst,

das Wachsenlassen des Wirtes, um das codierte trunkierte Hia zu exprimieren, und

das Isolieren und Reinigen des exprimierten Hia-Proteins.

[0025] Das codierte trunkierte Hia kann in Einschlusskörpern exprimiert werden. Der Isolations- und Reinigungsschritt kann bewirkt werden, indem die kultivierten transformierten Zellen aufgebrochen werden, um einen Überstand und die Einschlusskörper, welche das Hia enthalten, zu erzeugen, die Einschlusskörper nach Trennung von dem Überstand solubilisiert werden, um eine Lösung des rekombinanten Hia zu erzeugen, die Lösung des rekombinanten Hia chromatographisch gereinigt wird, so dass sie frei von Zelltrümmern ist, und das gereinigte rekombinante Hia-Protein isoliert wird.

[0026] Der Vektor, welcher die Wirtszelle wie z. B. *E. coli* transformiert, kann den T7-Promotor einschließen, und die *E. coli*- oder eine andere Wirtszelle kann in der Gegenwart einer induzierenden Menge von Lactose oder eines anderen geeigneten Induktionsmittels kultiviert werden.

[0027] Vorzugsweise wird die Isolation und Reinigung des exprimierten Hia bewirkt durch:

das Aufbrechen der kultivierten transformierten Zellen, um einen Überstand und die Einschlusskörper zu erzeugen,

das Solubilisieren der Einschlusskörper, um eine Lösung des rekombinanten Hia zu erzeugen,

das chromatographische Reinigen der Lösung des rekombinanten Hia, so dass sie frei von Zelltrümmern ist, und

das Isolieren des gereinigten rekombinanten Hia-Proteins.

[0028] Der Stamm von *Haemophilus influenzae* besteht aus dem nicht typisierbaren Stamm 33. Die spezifische Nukleinsäuresequenz für das Gen, welches das Hia-Protein aus einem solchen Stamm codiert, wird

nachstehend beschrieben.

[0029] Die hierin bereitgestellten Nukleinsäuremoleküle sind in diagnostischen Anwendungen nützlich.

[0030] Die immunogenen Zusammensetzungen der Erfindung sind in Impfstoffen zur in vivo-Verabreichung nützlich, um vor einer Erkrankung, die durch *Haemophilus* verursacht wird, zu schützen.

KURZE BESCHREIBUNG DER ZEICHNUNGEN

[0031] Die vorliegende Erfindung wird besser verstanden werden aus der folgenden Beschreibung mit Bezug auf die Figuren, in welchen:

[0032] [Fig. 1A](#) eine Restriktionskarte für das Plasmid DS-2008-2-3 zeigt, welches den T7-Promotor und das hia-Gen in voller Länge des NTHi-Stamms 11 enthält.

[0033] [Fig. 1B](#) die Oligonukleotide zeigt, die verwendet wurden, um das hia-Gen des Stamms 11 mit PCR zu amplifizieren. Sinn-Strang (5038.SL): SEQ ID No: 1, codierte Aminosäuren SEQ ID No: 2; Gegensinn-Strang (5039.SL): SEQ ID No: 3, komplementärer SEQ ID No: 4, codierte Aminosäuren SEQ ID No: 5. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; H, HindIII; N, NdeI; Ps, PstI; Sty, Styl. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz.

[0034] [Fig. 2](#) einen Immunoblot der Erkennung des rHia-Proteins in voller Länge durch einen anti-natives Adhäsins mit hohem Molekulargewicht von *Moraxella catarrhalis*-Antikörper zeigt. Bahn 1, DS-2043-1, nicht induziert; Bahn 2, DS-2043-1, für 4 h induziert; Bahn 3, DS-2043-2, nicht induziert; Bahn 4, DS-2043-2, für 4 h induziert; Bahn 5, Molekulargewichtsmarker. DS-2043-1 und DS-2043-2 sind unabhängige Klone von pT7-hia (11) in BL21 (DE3).

[0035] [Fig. 3](#) die Konstruktion der Plasmide DS-2092-1 und DS-2092-40 zeigt, die Tandemkopien der T7-hia-Genkassette für das hia-Gen des Stamms 11 enthalten. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; H, HindIII; Ps, PstI; Xb, XbaI. Die anderen Abkürzungen sind: CAP, alkalische Kalbsphosphatase; T7p, T7-Promotor; ApR, Ampicillinresistenz.

[0036] [Fig. 4](#) die Stellen der Trunkierung für das Hia-Protein des Stamms 11 (SEQ ID No: 6) zeigt.

[0037] [Fig. 5A](#) die Konstruktion von Plasmiden zeigt, welche trunkierte hia-Gene aus dem Stamm 11 exprimieren. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; H, HindIII; N, NdeI; Nhe, NheI; Ps, PstI; R, EcoRI; Sty, Styl; Xb, XbaI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz.

[0038] [Fig. 5B](#) die Oligonukleotide zeigt, die verwendet wurden, um die 5'-Fragmente für die trunkierten Gene mit PCR zu amplifizieren. E21-Trunkierung: Sinn (5524.SL): SEQ ID No: 7, codierte Aminosäuren SEQ ID No: 8; T33-Trunkierung: Sinn (5525.SL) SEQ ID No: 9, codierte Aminosäuren SEQ ID No: 10; V38-Trunkierung: Sinn (5526.SL): SEQ ID No: 11, codierte Aminosäuren, SEQ ID No: 12; N52-Trunkierung: Sinn (5527.SL): SEQ ID No: 13, codierte Aminosäuren SEQ ID No: 14; Gegensinn (5528.SL): SEQ ID No: 15; komplementär SEQ ID No: 16, codierte Aminosäuren SEQ ID No: 17.

[0039] [Fig. 6A](#) die Konstruktion des Plasmids BK-96-2-11 zeigt, welches das V38-hia-Gen aus dem NTHi-Stamm 11 und das *E. coli* cer-Gen enthält. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; K, KpnI; N, NdeI; P, PstI; R, EcoRI; S, Sell; Sm, SmaI; Sty, Styl; Xb, XbaI; Xho, XhoI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz; CAP, alkalische Kalbsphosphatase; tt1 Transkriptionsterminator 1 aus trpA; tt2, Transkriptionsterminator 2 aus dem T7-Gen 10.

[0040] [Fig. 6B](#) die Oligonukleotide zeigt, die verwendet wurden, um die multiple Klonierungsstelle und die Transkriptionsterminatoren zu konstruieren. "R" und "Ps" zeigen Enden an, die mit EcoRI- oder PstI-Enden überlappen werden aber die Stellen nicht neu erzeugen werden. Oberer Strang (SEQ ID No: 50), unterer Strang (SEQ ID No: 51).

[0041] [Fig. 7A](#) die Konstruktion der Plasmide DS-2242-1 und DS 2242-2 zeigt, welche den T7-Promotor und das hia-Gen des NTHi-Stamms 33 in voller Länge, das *E. coli* cer-Gen und das Kanamycinresistenzgen enthalten. Die Restriktionsenzymstellen sind: A, AlwNI; B, BamHI; Bg, BglII; H, HindIII; K, KpnI; N, NdeI; Ps, PstI;

R, EcoRI; S, Sall; Sm, SmaI; Xb, XbaI; Xho, XhoI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz; tt1 Transkriptionsterminator 1 aus trpA; tt2, Transkriptionsterminator 2 aus dem T7-Gen 10.

[0042] [Fig. 7B](#) die Oligonukleotide, die verwendet wurden, um das 5'-Ende des das hia-Gen des Stamms 33 codierenden Strangs (SEQ ID No: 52) zu erzeugen, den komplementären Strang (SEQ ID No: 53) und die codierte Aminosäuresequenz (SEQ ID No: 54) zeigt.

[0043] [Fig. 8A](#) die Konstruktion des Plasmids DS-2340-2-3 zeigt, welches den T7-Promotor und das V38-hia-Gen aus dem Stamms 33, das E. coli cer-Gen und das Kanamycinresistenzgen enthält. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; H, HindIII; N, NdeI; Ps, PstI; R, EcoRI; S, Sall; Sn, SnaBI; Xb, XbaI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz; tt1 Transkriptionsterminator 1 aus trpA; tt2, Transkriptionsterminator 2 aus dem T7-Gen 10.

[0044] [Fig. 8B](#) die Oligonukleotide zeigt, die verwendet wurden, um das 5'-Ende des trunkierten hia-Gens mit PCR zu amplifizieren. Sinn (6286.SL): SEQ ID No: 60, codierte Aminosäuren SEQ ID No: 61; Gegensinn (6287.SL) SEQ ID No: 18, komplementär SEQ ID No: 19, codierte Aminosäuren SEQ ID No: 20.

[0045] [Fig. 9A](#) und [Fig. 9B](#) die Konstruktion der Plasmide DS-2447-2 und DS 2448-17 zeigen, welche Tandemkopien der T7-V38-hia(11)- bzw. T7-V38-hia(33)-Gene enthalten. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; H, HindIII; Ps, PstI; R, EcoRI; S, Sall; Xb, XbaI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz; CAP, alkalische Kalbsphosphatase; tt1 Transkriptionsterminator 1 aus trpA; tt2, Transkriptionsterminator 2 aus dem T7-Gen 10.

[0046] [Fig. 10](#) die Expression von rHia zeigt. Feld A: Bahn 1, rHia (11) in voller Länge, keine Induktion; Bahn 2, rHia (11) in voller Länge; Bahn 3, E21-rHia (11); Bahn 4, T33-rHia (11); Bahn 5, V38-rHia (11); Bahn 6, N52-rHia (11). Feld B: Bahn 1, V38-rHia (11), keine Induktion; Bahn 2, V38-rHia (11); Bahn 3, V38-rHia (11)/cer.

[0047] [Fig. 11](#) ein Reinigungsschema für rHia-Proteine zeigt. Die Abkürzungen sind: SP, Überstand; PPT, Präzipitat; DTT, Dithiothreitol; OG, Octylglucosid; (x) bedeutet verworfen.

[0048] [Fig. 12](#), mit den Feldern A und B, die SDS-PAGE-Analyse von gereinigtem rHia zeigt. Feld A zeigt gereinigtes V38-rHia-Protein aus dem Stamm 11, und Feld B zeigt gereinigtes V38-rHia-Protein aus dem Stamm 33. Bahn 1, Molekulargewichtsmarker; Bahn 2, Ganzzelllysate; Bahn 3, Rohextrakt; Bahn 4, gereinigtes rHia-Protein.

[0049] [Fig. 13](#), mit den Feldern A, B und C, die Stabilität von V38-rHia (11) zeigt. Feld A zeigt Proben, die bei 4°C ohne Glycerol gelagert wurden. Feld B zeigt Proben, die bei 4°C in der Gegenwart von 20% Glycerol gelagert wurden. Feld C zeigt Proben, die bei -20°C in der Gegenwart von 20% Glycerol gelagert wurden. Bahn 0 zeigt t₀; die Bahnen 1 bis 8 zeigen Proben, die für 1 bis 8 Wochen gelagert wurden.

[0050] [Fig. 14](#), mit den Feldern A und B, die Immunogenität von V38-rHia (11) oder V38-rHia (33) in CD-1-Mäusen zeigt. Feld A zeigt die Antwort nach einer einzigen Immunisierung, und Feld B zeigt die Antwort einer Primär/Verstärkungs-Immunisierung.

[0051] [Fig. 15A](#) und [Fig. 15B](#) die Immunogenität von V38-rHia (11) in BALB/c-Mäusen und Meerschweinchen zeigen. [Fig. 15A](#) zeigt die Antikörperantwort in Mäusen, und [Fig. 15B](#) zeigt die Antwort in Meerschweinchen.

[0052] [Fig. 16](#) das Schutzvermögen von V38-rHia (33) vor einer nasopharyngealen Besiedlung in einem Chinchillamodell darstellt.

[0053] [Fig. 17](#) die Oligonukleotide zeigt, die verwendet wurden, um zusätzliche hia-Gene mit PCR zu amplifizieren. Sinn (5040.SL), SEQ ID No: 21, codierte Aminosäuren SEQ ID No: 22; Gegensinn (5039.SL), SEQ ID No: 3, komplementär SEQ ID No: 4, codierte Aminosäuren SEQ ID No: 5.

[0054] [Fig. 18](#) die Nukleotidsequenz (SEQ ID No: 23) und die abgeleitete Aminosäuresequenz (SEQ ID No: 24) des hia-Gens aus dem NTHi-Stamm 33 zeigt.

[0055] [Fig. 19](#) die Nukleotidsequenz (SEQ ID No: 25) und die abgeleitete Aminosäuresequenz (SEQ ID No:

26) des hia-Gens aus dem NTHi-Stamm 32 zeigt.

[0056] Fig. 20 die Nukleotidsequenz (SEQ ID No: 27) und die abgeleitete Aminosäuresequenz (SEQ ID No: 28) des hia-Gens aus dem NTHi-Stamm 29 zeigt.

[0057] Fig. 21 die Nukleotidsequenz (SEQ ID No: 29) und die abgeleitete Aminosäuresequenz (SEQ ID No: 30) des hia-Gens aus dem NTHi-Stamm M4071 zeigt.

[0058] Fig. 22 die Nukleotidsequenz (SEQ ID No: 31) und die abgeleitete Aminosäuresequenz (SEQ ID No: 32) des hia-Gens aus dem NTHi-Stamm K9 zeigt.

[0059] Fig. 23 die Nukleotidsequenz (SEQ ID No: 33) und die abgeleitete Aminosäuresequenz (SEQ ID No: 34) des hia-Gens aus dem NTHi-Stamm K22 zeigt.

[0060] Fig. 24 die Nukleotidsequenz (SEQ ID No: 35) und die abgeleitete Aminosäuresequenz (SEQ ID No: 36) des hia-Gens aus dem Typ c-Stamm API zeigt.

[0061] Fig. 25 die Nukleotidsequenz (SEQ ID No: 37) und die abgeleitete Aminosäuresequenz (SEQ ID No: 38) des hia-Locus aus dem NTHi-Stamm 12 zeigt. Die überstrichenen oder unterstrichenen Sequenzen zeigen Oligonukleotide an, die verwendet wurden, um über die Verbindung der zwei orfs hinweg mit PCR zu amplifizieren. Sinn (6431.SL) SEQ ID No: 39, (6432.SL) SEQ ID No: 40; Gegensinn (6295.SL) SEQ ID No: 41, (6271.SL) SEQ ID No: 42.

[0062] Fig. 26 die Nukleotidsequenz (SEQ ID No: 43) und die abgeleitete Aminosäuresequenz (SEQ ID No: 44) des hia-Locus aus dem NTHi-Stamm 11 zeigt, wie sie in dem U.S.-Patent Nr. 5,646,259 veröffentlicht ist.

[0063] Fig. 27 die Ausrichtung (das Alignment) des stromaufwärts liegenden ORF aus dem hia-Locus des Stamms 12 (SEQ ID No: 45) mit einem Teil des HI1732-Proteins (SEQ ID No: 46) aus dem H. influenzae Typ b-Stamm Rd zeigt.

[0064] Fig. 28 die Ausrichtung von Aminosäuresequenzen von Hia (SEQ ID No: 24, 26, 28, 34, 30, 44, 32), Hsf (SEQ ID No: 47) und partiellen Sequenzen von Moraxella catarrhalis-Proteinen mit hohem Molekulargewicht (200 kDa) aus den Stämmen 4223 und LES-1 (SEQ ID No: 48, 49) zeigt. Sternchen innerhalb der Sequenzen zeigen Stoppcodons an, aber unterhalb der Sequenz zeigten sie eine Sequenzhomologie an. Punkte zeigen identische Reste an. Die Sequenzalignierungen wurden durch direkten Vergleich der Aminosäuresequenzen der entsprechenden Proteine aufgestellt.

[0065] Fig. 29 die Oligonukleotide zeigt, die verwendet wurden, um das 5'-Ende des hia-Gens an der S44-trunkierten Position mit PCR zu amplifizieren. Sinn (6817.SL) SEQ ID No: 55, codierende Aminosäuren SEQ ID No: 56; Gegensinn (6818.SL) SEQ ID No: 57, komplementär SEQ ID No: 58, codierte Aminosäuren SEQ ID No: 59.

[0066] Fig. 30 die Konstruktion des Plasmids JB-2930-3 zeigt, welches das 544-hia-Gen aus dem NTHi-Stamm 11 und das E. coli cer-Gen und den T7-Promotor enthält. Die Restriktionsenzymstellen sind: B, BamHI; Bg, BglII; K, KpnI; N, NdeI; P, PstI; R, EcoRI; S, Sall; Sm, SmaI; Sty, StyI; Xb, XbaI; Xho, XhoI. Die anderen Abkürzungen sind: T7p, T7-Promotor; ApR, Ampicillinresistenz; KanR, Kanamycinresistenz; CAP, alkalische Kalbsphosphatase; tt1 Transkriptionsterminator 1 aus trpA; tt2, Transkriptionsterminator 2 aus dem T7-Gen 10.

[0067] Fig. 31 die SDS-PAGE-Analyse der Expression von rHia ausgehend von S44 zeigt. Bahn 1, Expression aus dem pET-S44-Vektor zur Zeit 0 (keine Induktion); Bahn 2, Expression aus dem pET-S44-Vektor nach 4 Stunden Induktion; Bahn 3, Expression aus JB-2930-3 nach 4 Stunden Induktion.

[0068] Fig. 32 eine schematische Darstellung der zwei Vektoren, JB-2930-3 und IA-191-3-1, die für die Expressionsstudie des S44-trunkierten rHia verwendet wurden, zeigt.

ALLGEMEINE BESCHREIBUNG DER ERFINDUNG

[0069] Da H. influenzae-Stämme geringe Mengen der Hia- und Hsf-Proteine erzeugen, wurde das hia-Gen aus NTHi-Stämmen in einen Expressionsvektor zur Überproduktion des rekombinanten Proteins in E. coli klo-

niert. Wenn das rekombinante Hia(rHia)-Protein in voller Länge exprimiert wurde, wurde es in relativ geringen Mengen erzeugt. Um zu bestätigen, dass es eine Expression des rekombinanten Proteins gab, wurde ein Immunoblot durchgeführt, welcher einen Antikörper verwendete, der gegen ein Adhäsinsprotein mit hohem Molekulargewicht von *Moraxella catarrhalis*, welches in dem U.S.-Patent Nr. 5,808,024 als 200 kDa identifiziert wurde, hervorgerufen wurde.

[0070] Antikörper gegen das gelgereinigte native 200 kDa-Protein erkannten eine spezifische induzierte Bande in der rHia-Proteinprobe. Die Ausbeute an rHia wurde durch ein Erhöhen der Genkopienzahl der T7-hia-Genkassette nicht signifikant verbessert.

[0071] Es wurde gezeigt, dass das *E. coli* cer-Gen Plasmide stabilisiert, welche große Inserts enthalten (Ref. 15), aber die Ausbeute von rHia wurde nicht signifikant verbessert, indem das *E. coli* cer-Gen zu dem Expressionsvektor hinzugefügt wurde. Jedoch wurde beobachtet, dass die *E. coli*-Zellen während der Kultivierung verklumpten, was nahe legt, dass es eine Expression an der Oberfläche des Hia-Adhäsinsproteins gab. Die apparente Toxizität des rHia-Proteins könnte überwunden werden, wenn es als Einschlusskörper erzeugt würde, und so wurden Trunkierungen an dem 5'-Ende des Gens durchgeführt, um mögliche Signalsequenzen zu entfernen. Diese Modifikation führte zu einer guten Produktion und Ausbeute von trunkiertem rHia, das an der V38-Position begann.

[0072] Das V38-trunkierte rHia-Protein und das in voller Länge waren immunogen, und die resultierenden anti-rHia-Antikörper waren in passiven Babyrattenmodellen der Bakteriämie aufgrund von *H. influenzae* Typ a- oder Typ b-Stämmen schützend. Zusätzlich wurde gefunden, dass das trunkierte V38-rHia-Protein gegenüber einer nasopharyngealen Besiedlung in einem aktiven Herausforderungsmodell in Chinchillas teilweise schützend ist. Der Schutz, der von rHia geboten wurde, das von einem NTHi-Stamm abgeleitet ist, vor einer Erkrankung, die durch NTHi und eingekapselte Typ a- oder Typ b-Stämme verursacht wird, zeigt, dass es gemeinsame schützende Epitope geben könnte. Die Klonierung und Sequenzanalyse von zusätzlichen hia-Genen könnte helfen, konservierte Bereiche zu identifizieren. Die N-terminal trunkierten rHia-Proteine und die in voller Länge könnten als Impfstoffkomponenten verwendet werden, um vor einer *Haemophilus influenzae*-Erkrankung zu schützen.

[0073] Der nicht typisierbare *Haemophilus*-Stamm 33 kann geeigneterweise verwendet werden, um die gereinigten und isolierten Nukleinsäuremoleküle bereitzustellen (welche in der Form von DNA-Molekülen vorliegen können), wie durch Ausführungsformen der vorliegenden Erfindung beispielhaft dargestellt wird. Ein solcher Stamm ist im Allgemeinen aus Klinikquellen und aus Bakterienkultursammlungen wie z. B. der American Type Culture Collection erhältlich.

[0074] In dieser Anmeldung wird der Ausdruck "Hia"-Protein verwendet, um eine Familie von Hia-Proteinen zu definieren, welche jene einschließt, die natürlich auftretende Variationen in ihren Aminosäuresequenzen aufweisen, wie sie in verschiedenen Stämmen von *Haemophilus* gefunden werden.

[0075] Was [Fig. 1A](#) betrifft, so wird dort eine Restriktionskarte des Plasmids DS-2008-2-3 dargestellt, welches ein hia-Gen in voller Länge aus dem nicht typisierbaren *Haemophilus influenzae*-Stamm 11, unter dem Einfluss des T7-Promotors, enthält. Die Nukleotidsequenz (SEQ ID No: 43) und die abgeleitete Aminosäuresequenz (SEQ ID No: 44) des hia-Gens aus Stamm 11 werden in dem vorstehend erwähnten U.S.-Patent Nr. 5,646,259 beschrieben (und darin als "HA1" identifiziert). Die Oligonukleotide, die verwendet wurden, um das hia-Gen ausgehend von dem ATG-Startcodon des Gens von Stamm 11 mit PCR zu amplifizieren, sind in [Fig. 1B](#) gezeigt.

[0076] Was [Fig. 2](#) betrifft, so wird dort ein Immunoblot dargestellt, welcher die Erkennung des rHia(11)-Proteins durch einen anti-natives Adhäsins mit hohem Molekulargewicht von *Moraxella catarrhalis*-Antikörper demonstriert. Das Adhäsins mit hohem Molekulargewicht von *M. catarrhalis* oder das 200 kDa-Protein, das in dem vorstehend erwähnten U.S.-Patent Nr. 5,808,024 beschrieben wird, weist eine gewisse Sequenzhomologie mit den Hia- und Hsf-Proteinen, insbesondere an dem Carboxyterminus (**Fig. 28**), auf.

[0077] Was [Fig. 3](#) betrifft, so wird dort ein Konstruktionsschema für die Plasmide DS-2092-1 und DS-2092-40 dargestellt, welche Tandemkopien der T7-hia-Genkassetten enthalten, welche das hia-Gen aus dem NTHi-Stamm 11 in voller Länge umfassen. Solche Plasmide, die erhöhte Kopienzahlen von Genen enthalten, weisen oft erhöhte Produktionsspiegel für rekombinante Proteine auf. Jedoch wurde, wie man unten sieht, die niedrige Ausbeute des rekombinanten Hia durch ein Erhöhen der Genkopienzahl nicht signifikant verbessert.

[0078] Was [Fig. 4](#) betrifft, so wird dort die N-terminale Sequenz des Proteins des NTHi-Stamms 11 und die Position von N-terminal trunkeerten Hia-Proteinen dargestellt. Die N-terminale Trunkierung bis zu Position E21 deletiert einen langen hydrophoben Bereich, welcher einen Teil einer Signalsequenz für Hia bilden könnte. Die Deletion bis zu Position T33 beinhaltet einen langen hydrophoben Bereich und folgt einer potentiellen Ala-X-Ala-Signal-Spaltungsstelle. Die Deletion bis zu Position V38 beinhaltet einen langen hydrophoben Bereich und folgt einer potentiellen Ala-X-Ala-Signal-Spaltungsstelle. Das rekombinante Hia-Protein, das bei Position S44 beginnt, beinhaltet einen langen hydrophoben Bereich und folgt einer potentiellen Ala-X-Ala-Signal-Spaltungsstelle. Das rekombinante Hia-Protein, das bei Position N52 beginnt, ahmt den ungefähren Start des verwandten Adhäsins mit hohem Molekulargewicht (200 kDa) aus *Moraxella catarrhalis* nach, das in dem vorstehend erwähnten U.S.-Patent Nr. 5,808,024 beschrieben wird, wobei das rekombinante Protein überproduziert wird, wenn es an seinem N-Terminus so trunkeert wird, dass es bei V56 beginnt.

[0079] Was [Fig. 5A](#) betrifft, so wird dort das Konstruktionsschema für die Erzeugung der Plasmide DS-2186-1-1, DS-2201-1, DS-2186-2-1 und DS-2168-2-6 dargestellt, welche vier der N-terminal trunkeerten rHia-Proteine erzeugen. Die Oligonukleotide, die verwendet wurden, um die 5'-Fragmente mit PCR zu amplifizieren, sind in [Fig. 5B](#) gezeigt. In [Fig. 30](#) wird das Konstruktionsschema für die Erzeugung des Plasmids JB-2930-3 dargestellt, welches die S44-Deletion erzeugt. Die Oligonukleotide, die verwendet wurden, um die 5'-Fragmente mit PCR zu amplifizieren, sind in [Fig. 29](#) gezeigt.

[0080] Was [Fig. 6A](#) betrifft, so wird dort ein Konstruktionsschema für die Erzeugung des Plasmids BK-96-2-11 dargestellt, welches das V38-hia-Gen aus dem NTHi-Stamm 11 wie auch das *E. coli* cer-Gen enthält, von welchem gezeigt wurde, dass es Plasmide stabilisiert. Von der Einführung des cer-Gens in Plasmide, die toxische Proteine erzeugen, wurde vorhergesagt, dass dieses die Proteinproduktion erhöht. Es wurde eine Änderung in der Morphologie der *E. coli*-Zellen, welche rHia in voller Länge in der Gegenwart des cer-Gens erzeugten, beobachtet, indem diese verklumpten. Dieses legt nahe, dass es eine verstärkte Expression des Adhäsins an der Oberfläche der Zellen gab, welche die Verklumpung verursachte. Das Expressionsplasmid BK 96-2-11 enthält ebenfalls Transkriptionsterminatoren stromaufwärts und stromabwärts der T7-V38-hia-Genkassette, von denen vorhergesagt wurde, dass sie die Stabilität des Gens erhöhen. Die Oligonukleotide, die verwendet wurden, um die multiple Klonierungsstelle und die Transkriptionsterminatoren zu erzeugen, sind in [Fig. 6B](#) gezeigt.

[0081] Was [Fig. 7A](#) betrifft, so wird dort ein Konstruktionsschema für die Plasmide DS-2242-1 und DS-2242-2 dargestellt, welche ein hia-Gen in voller Länge aus dem nicht typisierbaren *Haemophilus influenzae*-Stamm 33 unter dem Einfluss des T7-Promotors enthalten. Die Expressionsplasmide enthalten ebenfalls das *E. coli* cer-Gen und Transkriptionsterminatoren stromaufwärts und stromabwärts der T7-hia(33)-Genkassette. Bei DS-2242-1 sind die Terminatoren auf demselben Strang codiert wie das T7-hia(33)-Gen. Jedoch gab es keinen zu beobachtenden Unterschied bei der Expression von rHia aus den zwei Plasmiden. Die Oligonukleotide, die verwendet wurden, um das authentische 5'-Ende des Gens aus dem NTHi-Stamm 33 zu konstruieren, sind in [Fig. 7B](#) gezeigt.

[0082] Was [Fig. 8A](#) betrifft, so ist dort ein Konstruktionsschema für das Plasmid DS-2340-2-3 dargestellt, welches das V38-hia-Gen aus dem NTHi-Stamm 33 wie auch das *E. coli* cer-Gen enthält. Es sind ebenfalls Transkriptionsterminatoren stromaufwärts und stromabwärts der T7-V38-hia-Genkassette auf demselben Strang lokalisiert. Die Oligonukleotide, die verwendet wurden, um das hia-Gen des NTHi-Stamms 33 ausgehend von dem V38-Codon mit PCR zu amplifizieren, sind in [Fig. 8B](#) gezeigt.

[0083] Was [Fig. 9](#) betrifft, so ist dort die Konstruktion der Plasmide DS-2447-2 und DS-2448-17 gezeigt, welche Tandemkopien der T7-V38-hia(11)- bzw. T7-V38-hia(33)-Genkassetten enthalten.

[0084] Was [Fig. 10](#) betrifft, so ist dort in Feld A die Produktion von rHia-Proteinen aus Plasmiden dargestellt, welche hia-Gene aus dem NTHi-Stamm 11 in voller Länge oder trunkeert codieren. Die Produktion von rHia(11)-Protein in voller Länge war sehr gering. Es wurde ebenfalls eine geringe Expression für die E21- und T33-trunkeerten rHia-Proteine beobachtet. Jedoch weisen die V38- und N52-trunkeerten rHia-Proteine signifikant verbesserte Expressionsspiegel auf. Wie in [Fig. 10](#), Feld B, gezeigt ist, scheint die Produktion von V38 rHia (11) verstärkt zu werden, wenn das *E. coli* cer-Gen zu dem Expressionsplasmid zugefügt wird.

[0085] Was [Fig. 11](#) betrifft, so wird dort ein Reinigungsschema für rHia-Proteine dargestellt, welche als Einschlusskörper erzeugt werden. Zellen wurden durch Beschallung lysiert, und die Einschlusskörper durch Reihenextraktionen gereinigt. Die Einschlusskörper wurden in Guanidiniumchlorid solubilisiert und Verunreinigungen durch die Zugabe von Polyethylenglycol (PEG) ausgefällt. Die Zugabe von $(\text{NH}_4)_2\text{SO}_4$ führte zu der Aus-

fällung von rHia, und das rohe rHia wurde weiter durch Gelfiltration gereinigt.

[0086] Was [Fig. 12](#) betrifft, so werden dort die gereinigten V38-rHia-Proteine aus den Stämmen 11 und 33 dargestellt. Die Einschlusskörper sind in Bahn 3 und das endgültige gereinigte Protein in Bahn 4 gezeigt. Die abgeschätzte Reinheit des gereinigten Proteins ist größer als ca. 90%, wie durch SDS-PAGE-Densitometrie bestimmt wurde.

[0087] Was [Fig. 13](#) betrifft, so ist dort die SDS-PAGE-Analyse der Stabilität von rHia-Proteinen, die wie hierin beschrieben erzeugt wurden, während 8 Wochen Lagerung mit oder ohne Glycerol bei 4°C und mit Glycerol bei -20°C gezeigt. Das Protein ist unter jeder dieser Bedingungen stabil.

[0088] Was [Fig. 14](#) betrifft, so wird dort die Immunogenität von V38-rHia-Proteinen aus den Stämmen 11 und 33 in CD-1-Mäusen dargestellt. Bei Dosen von 0,3 bis 10 µg gibt es bei jedem Protein nach ein oder zwei Dosen eine starke Immunantwort. Bei diesen Spiegeln gibt es keine offensichtliche Dosisantwort. Ähnliche Ergebnisse wurden in BALB/c-Mäusen ([Fig. 15A](#)) und in Meerschweinchen ([Fig. 15B](#)) beobachtet, was zeigt, dass rHia selbst bei 0,3 µg pro Dosis sehr immunogen war.

[0089] Was [Fig. 16](#) betrifft, so ist dort der Schutz dargestellt, der von V38-rHia (33) vor einer Besiedlung durch den NTHi-Stamm 33 geboten wird. Wie von Yang et al. (Ref. 20) beschrieben wird, ist ein nasopharyngeales Besiedlungsmodell bei Chinchillas entwickelt worden, um einen Schutz vor diesem frühesten Stadium der Erkrankung festzustellen. Das Modell wurde ursprünglich für NTHi-Stämme eingerichtet, die hmw-Gene exprimieren, und musste für NTHi-Stämme, die hia-Gene exprimieren, angepasst werden. Für den Prototyp des hmw exprimierenden Stamms (NTHi 12) konnten 10² bis 10⁸ cfu verwendet werden, um eine Infektion einzuführen, aber 5 × 10⁸ cfu des NTHi-Stamms 33 wurden benötigt, und selbst bei diesem hohen Spiegel konnte mit dem Prototyp des hia exprimierenden Stamms 11 keine Infektion eingeführt werden. Bei einer Dosis von 100 µg ist es offensichtlich, dass es in der immunisierten Gruppe einen partiellen Schutz gibt, auch wenn es bei einer Dosis von 50 µg keinen Schutz gibt. Ein solcher Schutz vor den frühen Stadien der Erkrankung veranschaulicht den Nutzen der rHia-Adhäsine als Impfstoffantigene.

[0090] Was [Fig. 17](#) betrifft, so werden dort die Oligonukleotide dargestellt, die verwendet wurden, um zusätzliche Haemophilus influenzae-hia-Gene mit PCR zu amplifizieren. Die Sequenzen basieren auf den konservierten amino- und carboxyterminalen Sequenzen der Hia- und Hsf-Proteine.

[0091] Was [Fig. 18](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms 33 dargestellt. Was [Fig. 19](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms 32 dargestellt. Was [Fig. 20](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms 29 dargestellt. Was [Fig. 21](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms M4071 dargestellt. Was [Fig. 22](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms K9 dargestellt. Was [Fig. 23](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des NTHi-Stamms K22 dargestellt. Was [Fig. 24](#) betrifft, so wird dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens des Haemophilus influenzae Typ c-Stamms API dargestellt. Was [Fig. 25](#) betrifft, so werden dort die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Locus aus dem NTHi-Stamm 12 dargestellt. Das PCR-amplifizierte Fragment enthält das 3'-Ende eines Gens, das mit dem HI1733-Gen aus dem Haemophilus influenzae Typ d-Stamm Rd-Genom verwandt ist, verbunden mit dem 3'-Ende eines hia-Gens. Eine Ausrichtung des stromaufwärts liegenden ORF mit dem HI1733-Protein ist in [Fig. 27](#) gezeigt.

[0092] [Fig. 26](#) zeigt die vollständige Nukleotidsequenz und die abgeleitete Aminosäuresequenz des hia-Gens aus dem NTHi-Stamm 11, wie sie in der vorstehend erwähnten U.S. 5,646,259 veröffentlicht sind.

[0093] Was [Fig. 28](#) betrifft, so wird dort eine Ausrichtung der abgeleiteten Proteinsequenzen von Hsf, Hia und der partiellen Sequenzen des M. catarrhalis 200 kDa-Proteins dargestellt.

[0094] Es ist für einen Fachmann klar ersichtlich, dass die verschiedenen Ausführungsformen der vorliegenden Erfindung Verwendung finden bei Anwendungen auf den Gebieten der Impfung, Diagnose, Behandlung einer Haemophilus-Infektion und Erzeugung von immunologischen Mitteln. Eine weitere nicht beschränkende Diskussion solcher Verwendungen wird ferner nachstehend angegeben.

[0095] Immunogene Zusammensetzungen, die geeignet sind, um als Impfstoff verwendet zu werden, können aus immunogenen rekombinanten Haemophilus influenzae Adhäsion (rHia)-Proteinen von nicht typisierbaren Haemophilus-Stämmen, immunogenen Analogen und Fragmenten von diesen und/oder immunogenen Peptiden, wie sie hierin offenbart sind, hergestellt werden. Der Impfstoff ruft eine Immunantwort hervor, welche Antikörper, einschließlich anti-rHia-Antikörpern und Antikörpern, die opsonisierend oder bakterizid sind, erzeugt.

[0096] Immunogene Zusammensetzungen, einschließlich Impfstoffen, können zur Injektion, als flüssige Lösungen oder Emulsionen, hergestellt werden. Das rHia-Protein, immunogene Analoge und Fragmente von diesem und/oder immunogene Peptide können mit pharmazeutisch verträglichen Arzneimittelträgern gemischt werden, welche mit dem rHia-Protein, immunogenen Fragmenten, Analogen oder immunogenen Peptiden kompatibel sind. Solche Arzneimittelträger können Wasser, Salzlösung, Dextrose, Glycerol, Ethanol oder Kombinationen von diesen beinhalten.

[0097] Die immunogenen Zusammensetzungen und Impfstoffe können weiterhin Hilfssubstanzen wie z. B. Benetzungs- oder Emulgiermittel, pH-puffernde Mittel oder Hilfsmittel, um die Wirksamkeit der Impfstoffe zu erhöhen, enthalten.

[0098] Die immunogenen Zusammensetzungen und Impfstoffe können parenteral, durch subkutane oder intramuskuläre Injektion verabreicht werden. Alternativ können die immunogenen Zusammensetzungen, die gemäß der vorliegenden Erfindung gebildet wurden, in einer Weise formuliert und befördert werden, dass sie eine Immunantwort an Schleimhautoberflächen hervorrufen. So können die immunogenen Zusammensetzungen beispielsweise über den nasalen oder oralen (intragastischen) Weg an Schleimhautoberflächen verabreicht werden.

[0099] Die immunogene Zusammensetzung kann in Kombination mit einem Zielmolekül zur Beförderung zu spezifischen Zellen des Immunsystems oder zu Schleimhautoberflächen bereitgestellt werden. Einige solcher Zielmoleküle beinhalten Vitamin B12 und Fragmente von bakteriellen Toxinen, wie in der WO 92/17167 (Biotech Australia Pty. Ltd.) beschrieben wird, und monoklonale Antikörper, wie in dem U.S.-Patent Nr. 5,194,254 (Barber et al.) beschrieben wird.

[0100] Alternativ können andere Arten der Verabreichung, einschließlich Suppositorien und oraler Formulierungen, wünschenswert sein. Für Suppositorien können Bindemittel und Träger beispielsweise Polyalkylenglycole oder Triglyceride einschließen. Orale Formulierungen können normalerweise eingesetzte Träger (incipients) wie z. B. Saccharin, Cellulose und Magnesiumcarbonat von pharmazeutischer Güte einschließen. Diese Zusammensetzungen nehmen die Form von Lösungen, Suspensionen, Tabletten, Pillen, Kapseln, verzögert freisetzenden Formulierungen oder Pulvern an und enthalten ca. 1 bis 95% des rHia-Proteins, der Fragmentanalogen und/oder Peptide.

[0101] Die Impfstoffe können in einer Weise verabreicht werden, die mit der Dosisformulierung kompatibel ist, und in einer solchen Menge, wie sie therapeutisch wirksam, schützend und immunogen sein wird. Die zu verabreichende Menge hängt von dem zu behandelnden Subjekt, einschließlich z. B. der Kapazität des Immunsystems des Individuums ab, Antikörper zu synthetisieren, und wenn nötig, eine zellvermittelte Immunantwort zu erzeugen. Die genauen Mengen des aktiven Bestandteils, die verabreicht werden müssen, hängen von der Beurteilung des praktizierenden Arztes ab. Jedoch sind geeignete Dosisbereiche von einem Fachmann leicht zu bestimmen und können in der Größenordnung von Mikrogramm des rHia, der Analogen und Fragmente von diesem und/oder Peptide liegen. Geeignete Regimes zur Primärverabreichung und für Verstärkungsdosen sind ebenfalls variabel, können aber eine Primärverabreichung, gefolgt von anschließenden Verabreichungen, beinhalten. Die Dosis des Impfstoffs kann ebenfalls von dem Weg der Verabreichung abhängen und wird entsprechend der Größe des Wirtes variieren.

[0102] Die Nukleinsäuremoleküle, welche die rHia-Proteine von nicht typisierbaren Haemophilus codieren, können ebenfalls direkt zur Immunisierung durch Verabreichung der DNA direkt, beispielsweise durch Injektion zur genetischen Immunisierung oder durch Konstruieren eines Lebendvektors wie z. B. Salmonella, BCG, Adenovirus, Pockenvirus, Vaccinia- oder Poliovirus, welche das Nukleinsäuremolekül enthalten, verwendet werden. Eine Diskussion von gewissen Lebendvektoren, die verwendet worden sind, um heterologe Antigene in das Immunsystem zu tragen, ist beispielsweise in O'Hagen (1992) (Ref. 16) enthalten. Prozesse für die direkte Injektion der DNA in Testsubjekte zur genetischen Immunisierung werden z. B. in Ulmer et al. 1993 (Ref. 17) beschrieben.

[0103] Die Immunogenität kann signifikant verbessert werden, wenn die Antigene zusammen mit Hilfsmitteln verabreicht werden, die gewöhnlich als eine 0,05 bis 1,0 Prozent-Lösung in phosphatgepufferter Salzlösung verwendet werden. Hilfsmittel verstärken die Immunogenität eines Antigens, sind aber nicht notwendigerweise selbst immunogen. Hilfsmittel können wirken, indem sie das Antigen lokal nahe der Verabreichungsstelle zurückhalten, um eine Depotwirkung zu erzeugen, welche eine langsame, verzögerte Freisetzung des Antigens an Zellen des Immunsystems erleichtert. Hilfsmittel können ebenfalls Zellen des Immunsystems zu einem Antigenepot hinziehen und solche Zellen stimulieren, um Immunantworten hervorzurufen.

[0104] Immunstimulatorische Mittel oder Hilfsmittel sind über viele Jahre verwendet worden, um die Immunantworten eines Wirts auf beispielsweise Impfstoffe zu verbessern. Intrinsische Hilfsmittel wie z. B. Lipopolysaccharide sind normalerweise die Komponenten von abgetöteten oder abgeschwächten Bakterien, die als Impfstoffe verwendet werden. Extrinsische Hilfsmittel sind Immunmodulatoren, welche typischerweise nicht kovalent mit Antigenen verknüpft sind und formuliert sind, um die Immunantworten eines Wirts zu verstärken. So wurden Hilfsmittel identifiziert, welche die Immunantwort auf Antigene, die parenteral gegeben werden, verstärken. Einige dieser Hilfsmittel sind jedoch toxisch und können unerwünschte Nebenwirkungen verursachen, was diese zur Verwendung in Menschen und vielen Tieren ungeeignet macht. In der Tat werden routinemäßig nur Aluminiumhydroxid und Aluminiumphosphat (zusammen üblicherweise als Alaun bezeichnet) als Hilfsmittel in Human- und Veterinärimpfstoffen verwendet. Die Effizienz von Alaun bei der Verstärkung von Antikörperantworten auf Diphtherie- und Tetanustoxoide ist gut etabliert.

[0105] Ein weiter Bereich von extrinsischen Hilfsmitteln kann starke Immunantworten auf Antigene hervorruufen. Diese beinhalten die spezifischen Hilfsmittel, die oben detailliert angegeben sind, wie auch Saponine, die mit Membranproteinantigenen komplexiert sind (immunstimulierende Komplexe), pluronische Polymere mit Mineralöl, abgetötete Mycobakterien und Mineralöl, Freund's vollständiges Adjuvans, bakterielle Produkte wie z. B. Muramyl-dipeptid (MDP) und Lipopolysaccharid (LPS), wie auch Lipid A und Liposomen.

[0106] Um effizient humorale Immunantworten (HIR) und eine zellvermittelte Immunität (CMI) zu induzieren, werden Immunogene in Hilfsmitteln emulgiert. Viele Hilfsmittel sind toxisch, induzieren Granulome, akute und chronische Entzündungen (Freund's vollständiges Adjuvans, FCA), Zytolyse (Saponine und pluronische Polymere) und Pyrogenizität, Arthritis und Uveitis anterior (LPS und MDP). Obwohl FCA ein ausgezeichnetes Hilfsmittel ist und in der Forschung weithin verwendet wird, ist es zur Verwendung in Human- oder Veterinärimpfstoffen aufgrund seiner Toxizität nicht zugelassen.

[0107] Wünschenswerte Charakteristika von idealen Hilfsmitteln beinhalten:

- (1) das Fehlen von Toxizität;
- (2) die Fähigkeit, eine lang andauernde Immunantwort zu stimulieren;
- (3) die Einfachheit der Herstellung und die Stabilität bei einer Langzeitlagerung;
- (4) die Fähigkeit, sowohl CMI als auch HIR auf Antigene, die wenn nötig über verschiedene Wege verabreicht werden, hervorzurufen;
- (5) die Synergie mit anderen Hilfsmitteln;
- (6) die Fähigkeit, selektiv mit Populationen antigenpräsentierender Zellen (APC) wechselzuwirken;
- (7) die Fähigkeit, spezifisch geeignete T_H1 - oder T_H2 -zellspezifische Immunantworten hervorzurufen; und
- (8) die Fähigkeit, die Spiegel geeigneter Antikörperisotypen (z. B. IgA) gegen Antigene selektiv zu erhöhen.

[0108] Das U.S.-Patent Nr. 4,855,283, das für Lockhoff et al. am 8. August 1989 erteilt wurde, lehrt Glycolipidanalogue einschließlich N-Glycosylamiden, N-Glycosylharnstoffen und N-Glycosylcarbamaten, welche jeweils in dem Zuckerrest durch eine Aminosäure substituiert sind, als Immunmodulatoren oder Hilfsmittel. So berichteten Lockhoff et al. 1991 (Ref. 18), dass N-Glycolipidanalogue, die strukturelle Ähnlichkeiten mit den natürlich vorkommenden Glycolipiden, wie z. B. Glycosphingolipiden und Glycoglycerolipiden, zeigen, in der Lage sind, sowohl bei Herpes simplex-Virus-Impfstoff als auch Pseudorabies-Virus-Impfstoff starke Immunantworten hervorzurufen. Einige Glycolipide wurden aus langkettigen Alkylaminen und Fettsäuren, die direkt mit den Zuckern durch das anomere Kohlenstoffatom verknüpft sind, synthetisiert, um die Funktionen der natürlich vorkommenden Lipidreste nachzuahmen.

[0109] Das U.S.-Patent Nr. 4,258,029 (ehrt, dass Octadecyltyrosinhydrochlorid (OTH) als ein Hilfsmittel fungiert, wenn es mit Tetanustoxoid und formalininaktiviertem Typ I, II und III Poliomyelitisvirus-Impfstoff komplexiert wird. Ebenfalls berichteten Nixon-George et al. 1990 (Ref. 19), dass Octadecylester von aromatischen Aminosäuren, die mit einem rekombinanten Hepatitis B-Oberflächenantigen komplexiert wurden, die Immunantworten eines Wirts auf Hepatitis B-Virus verstärkten.

Immuntests

[0110] Das rHia-Protein eines nicht typisierbaren Stammes von *Haemophilus*, Analoge und Fragmente von diesem sind nützlich als Immunogene, als Antigene in Immuntests, einschließlich enzymverknüpften Immunsorbenztests (ELISA), RIAs und anderen nicht enzymverknüpften Antikörperbindungstests oder Prozeduren, die im Stand der Technik zum Nachweis von antibakteriellen, *Haemophilus*- und/oder Hia-Antikörpern bekannt sind. In ELISA-Tests wird das Hia-Protein, Analoge und Fragmente, auf einer ausgewählten Oberfläche immobilisiert, beispielsweise einer Oberfläche, die in der Lage ist, Proteine oder Peptide zu binden, wie beispielsweise den Vertiefungen einer Polystyrol-Mikrotiterplatte. Nach einem Waschen, um unvollständig adsorbiertes Hia-Protein, Analoge und/oder Fragmente zu entfernen, kann ein nicht spezifisches Protein, wie z. B. eine Lösung von Rinderserumalbumin (BSA) oder Casein, von dem bekannt ist, dass es in Bezug auf die Testprobe als Antigen neutral ist, an die ausgewählte Oberfläche gebunden werden. Dieses ermöglicht eine Blockierung der nicht spezifischen Adsorptionsstellen auf der immobilisierenden Oberfläche und verringert somit den Hintergrund, der durch nicht spezifische Bindungen von Antiseren an die Oberfläche verursacht wird.

[0111] Die immobilisierende Oberfläche wird dann mit einer zu testenden Probe, wie z. B. klinischen oder biologischen Materialien, in einer Weise in Kontakt gebracht, die für eine Immunkomplex(Antigen/Antikörper)-Bildung förderlich ist. Dieses kann das Verdünnen der Probe mit Verdünnungsmitteln wie z. B. BSA, Rindergammaglobulin (BGG) und/oder phosphatgepufferter Salzlösung (PBS)/Tween einschließen. Die Probe wird dann für ca. 2 bis ca. 4 Stunden bei Temperaturen wie beispielsweise in der Größenordnung von ca. 25 bis ca. 37°C inkubieren gelassen. Nach der Inkubation wird die mit der Probe in Kontakt gebrachte Oberfläche gewaschen, um nicht immunkomplexiertes Material zu entfernen. Der Waschvorgang kann ein Waschen mit einer Lösung wie z. B. PBS/Tween oder einem Boratpuffer einschließen.

[0112] Nach der Bildung von spezifischen Immunkomplexen zwischen der Testprobe und dem gebundenen Hia-Protein, Analogon und/oder Fragmenten, und einer anschließenden Wäsche kann das Auftreten und sogar die Menge der Immunkomplexbildung bestimmt werden, indem der Immunkomplex einem zweiten Antikörper mit Spezifität für den ersten Antikörper ausgesetzt wird. Wenn die Testprobe menschlichen Ursprungs ist, ist der zweite Antikörper ein Antikörper mit Spezifität für menschliche Immunglobuline und im Allgemeinen IgG. Um Mittel zum Nachweis bereitzustellen, kann der zweite Antikörper eine assoziierte Aktivität wie beispielsweise eine enzymatische Aktivität aufweisen, die z. B. eine Farbentwicklung bei Inkubation mit einem geeigneten chromogenen Substrat erzeugen wird. Eine Quantifizierung kann dann erreicht werden, indem das Ausmaß der Farberzeugung gemessen wird, indem beispielsweise ein Spektrophotometer für sichtbare Spektren verwendet wird.

Verwendung von Sequenzen als Hybridisierungssonden

[0113] Die hierin beschriebenen Nukleotidsequenzen, welche die neu isolierten und charakterisierten Sequenzen der *hia*-Gene umfassen, erlauben die Identifizierung und Klonierung der *hia*-Gene aus anderen nicht typisierbaren Stämmen von *Haemophilus*.

[0114] Die Nukleotidsequenzen, welche die Sequenz der hierin beschriebenen *hia*-Gene umfassen, sind aufgrund ihrer Fähigkeit, selektiv Doppelstrangmoleküle mit komplementären Strängen anderer *hia*-Gene zu bilden, nützlich. Abhängig von der Anwendung kann eine Vielzahl von Hybridisierungsbedingungen eingesetzt werden, um variierende Grade an Selektivität der Sonde gegenüber den anderen *hia*-Genen in anderen Stämmen eines nicht typisierbaren *Haemophilus* zu erreichen. Für einen hohen Grad an Selektivität werden relativ stringente Bedingungen verwendet, um die Doppelstränge zu bilden, wie z. B. Bedingungen mit wenig Salz und/oder hoher Temperatur, wie sie beispielsweise durch 0,02 M bis 0,15 M NaCl bei Temperaturen zwischen ca. 50°C bis 70°C bereitgestellt werden. Für einige Anwendungen werden weniger stringente Hybridisierungsbedingungen benötigt, wie z. B. 0,15 M bis 0,9 M Salz, bei Temperaturen, die von zwischen 20°C bis 55°C reichen. Die Hybridisierungsbedingungen können ebenfalls durch die Zugabe von steigenden Mengen an Formamid, um den Hybriddoppelstrang zu destabilisieren, stringenter gemacht werden. So können die besonderen Hybridisierungsbedingungen leicht manipuliert werden und werden im Allgemeinen ein Verfahren der Wahl sein, das von den gewünschten Ergebnissen abhängt. Im Allgemeinen sind geeignete Hybridisierungstemperaturen in der Gegenwart von 50% Formamid und 0,15 M NaCl: 42°C für ein *hia*-Gen, welches zu ca. 95 bis 100% homolog zu dem Zielnukleinsäurefragment ist, 37°C für ca. 90 bis 95% Homologie und 32°C für ca. 85 bis 90% Homologie.

[0115] In einem klinischen diagnostischen Beispiel können die Nukleinsäuresequenzen der *hia*-Gene in Kombination mit einem geeigneten Mittel, wie z. B. einer Markierung, zur Bestimmung der Hybridisierung verwendet

werden. Eine breite Vielzahl von geeigneten Indikatormitteln ist im Stand der Technik bekannt, einschließlich radioaktiver, enzymatischer oder anderer Liganden, wie z. B. Avidin/Biotin, welche in der Lage sind, ein nachweisbares Signal zu liefern. Bei einigen diagnostischen Ausführungsformen kann eine Enzymmarkierung wie z. B. Urease, alkalische Phosphatase oder Peroxidase an Stelle einer radioaktiven Markierung verwendet werden. In dem Fall von Enzymmarkierungen sind kolorimetrische Indikatorsubstrate bekannt, welche eingesetzt werden können, um ein Mittel bereitzustellen, das für das menschliche Auge oder spektrophotometrisch sichtbar ist, um eine spezifische Hybridisierung mit Proben, die his-Gensequenzen enthalten, zu identifizieren.

[0116] Die Nukleinsäuresequenzen von his-Genen sind als Hybridisierungssonden bei Lösungshybridisierungen und bei Beispielen, die Festphasenprozeduren einsetzen, nützlich. Bei Beispielen, an denen Festphasenprozeduren beteiligt sind, wird die Test-DNA (oder RNA) aus Proben, wie z. B. klinischen Proben, einschließlich Exudaten, Körperflüssigkeiten (z. B. Serum, Amnionflüssigkeit, Mittelohreffusion, Sputum, Bronchioalveolar-spülungsflüssigkeit) oder sogar Geweben, auf einer ausgewählten Matrix oder Oberfläche adsorbiert oder in anderer Weise befestigt. Die befestigte, einzelsträngige Nukleinsäure wird dann einer spezifischen Hybridisierung mit ausgewählten Sonden, welche die Nukleinsäuresequenzen der hia-Gene oder von Fragmenten von diesen umfassen, unter gewünschten Bedingungen unterzogen. Die ausgewählten Bedingungen werden von den besonderen Umständen abhängen, die auf den benötigten bestimmten Kriterien basieren, die z. B. von dem G+C-Gehalt, dem Typ der Zielnukleinsäure, der Quelle der Nukleinsäure, der Größe der Hybridisierungs-sonde usw. abhängen. Nach dem Waschen der Hybridisierungsoberfläche, um so nicht spezifisch gebundene Sondenmoleküle zu entfernen, wird die spezifische Hybridisierung mit Hilfe der Markierung nachgewiesen oder sogar quantifiziert. Ein Beispiel wählt Nukleinsäuresequenzteile aus, welche bei den Spezies von *Haemophilus* konserviert sind. Die ausgewählte Sonde kann wenigstens 18 bp in der Länge aufweisen und kann in dem Bereich von 30 bis 90 bp Länge liegen.

Expression der *Haemophilus influenzae* Adhäsion-Gene

[0117] Plasmidvektoren, welche Replikon- und Kontrollsequenzen enthalten, welche von Spezies abgeleitet sind, die mit der Wirtszelle kompatibel sind, können zur Expression der hia-Gene in Expressionssystemen verwendet werden. Der Vektor trägt gewöhnlich eine Replikationsstelle wie auch markierende Sequenzen, welche in der Lage sind, in transformierten Zellen eine phänotypische Selektion bereitzustellen. Beispielsweise kann *E. coli* unter Verwendung von pBR322 transformiert werden, welches Gene für eine Ampicillin- und Tetracyclinresistenz enthält und somit ein leichtes Mittel zur Identifizierung transformierter Zellen bereitstellt. Das pBR322-Plasmid, oder ein anderes mikrobielles Plasmid oder ein Phage, muss ebenfalls Promotoren enthalten oder modifiziert werden, um diese zu enthalten, welche von der Wirtszelle zur Expression ihrer eigenen Proteine verwendet werden können.

[0118] Zusätzlich können Phagenvektoren, die Replikon- und Kontrollsequenzen enthalten, die mit dem Wirt kompatibel sind, als ein transformierender Vektor in Verbindung mit diesen Wirten verwendet werden. Beispielsweise kann der Phage in Lambda GEMTM-11 benutzt werden, um rekombinante Phagenvektoren zu erzeugen, welche verwendet werden können, um Wirtszellen wie z. B. *E. coli* LE392 zu transformieren.

[0119] Promotoren, die üblicherweise bei der Konstruktion rekombinanter DNA verwendet werden, beinhalten das β -Lactamase-(Penicillinase) und Lactosepromotorsystem und andere mikrobielle Promotoren wie z. B. das T7-Promotorsystem, das hierin in bevorzugten Ausführungsformen eingesetzt wird (U.S.-Patent 4,952,496). Details bezüglich der Nukleotidsequenzen von Promotoren sind bekannt, was einer Fachkraft ermöglicht, diese funktional mit Genen zu verbinden. Der verwendete besondere Promotor wird im Allgemeinen eine Frage der Wahl sein, die von den gewünschten Ergebnissen abhängt. Wirte, die zur Expression des Hia-Proteins und immunologischer Fragmente oder Analoga von diesem geeignet sind, beinhalten *E. coli*, *Bordetella*-Spezies, *Bacillus*-Spezies, *Haemophilus*, Pilze, Hefe, oder es kann das Baculovirus-Expressionssystem verwendet werden. *E. coli* ist der bevorzugte Wirt, welcher hierin verwendet wird.

[0120] Hia-Proteine können durch rekombinante Verfahren erzeugt werden, insbesondere wenn das natürlich vorkommende Hia-Protein, wie es aus einer Kultur einer Spezies von *Haemophilus* gereinigt wird, Spurenmen-gen von toxischen Materialien oder andere Verunreinigungen beinhalten könnte. Dieses Problem kann vermieden werden, indem rekombinant erzeugtes Hia-Protein in heterologen Systemen verwendet wird, welches aus dem Wirt in einer Weise isoliert werden kann, dass Verunreinigungen in den gereinigten Materialien minimiert werden, indem speziell die hierin beschriebenen Konstrukte eingesetzt werden.

BIOLOGISCHE HINTERLEGUNGEN

[0121] Ein Vektor, welcher eine Nukleinsäure enthält, die für ein Protein mit hohem Molekulargewicht eines nicht typisierbaren Stammes von *Haemophilus codiert*, welches hierin beschrieben und auf welches hierin Bezug genommen wird, wurde bei der American Type Culture Collection (ATCC), welche an dem University Boulevard, 10801 Manassas, Virginia 20110-2209, USA gelegen ist, gemäß dem Budapester Vertrag und vor der Einreichung dieser Anmeldung hinterlegt. Proben des hinterlegten Vektors werden der Öffentlichkeit zugänglich werden, und alle Beschränkungen, die für den Zugang zu den Hinterlegungen auferlegt wurden, werden bei Erteilung eines Patents auf der Basis dieser U.S.-Patentanmeldung entfernt werden. Zusätzlich wird die Hinterlegung ausgetauscht werden, wenn lebensfähige Proben von der Hinterlegungsstelle nicht abgegeben werden können. Die hierin beschriebene und beanspruchte Erfindung soll im Umfang nicht durch die hinterlegten biologischen Materialien eingeschränkt werden, da die hinterlegte Ausführungsform nur als eine Veranschaulichung der Erfindung gedacht ist.

Zusammenfassung der Hinterlegung

Plasmid	ATCC	Hinterlegungsdatum
BK-96-2-11	203771	11. Februar 1999

BEISPIELE

[0122] Die obige Offenbarung beschreibt die vorliegende Erfindung allgemein. Ein vollständigeres Verständnis kann durch Bezug auf die folgenden spezifischen Beispiele erhalten werden. Diese Beispiele werden lediglich zu Zwecken der Veranschaulichung beschrieben und sind nicht gedacht, um den Umfang der Erfindung zu beschränken. Änderungen in der Form und Austausch durch Äquivalente sind beabsichtigt, wie es die Umstände nahe legen können oder als zweckdienlich erscheinen lassen. Auch wenn hierin spezifische Ausdrücke eingesetzt wurden, sind solche Ausdrücke in einem beschreibenden Sinn und nicht zu Zwecken der Beschränkung gedacht.

[0123] Von Verfahren der molekularen Genetik, Proteinbiochemie, Immunologie und Fermentationstechnologie, die in dieser Offenbarung und diesen Beispielen verwendet werden, aber nicht explizit beschrieben werden, wird in der wissenschaftlichen Literatur reichlich berichtet, und diese liegen ganz innerhalb der Fähigkeiten der Fachleute.

Beispiel 1

[0124] Dieses Beispiel beschreibt die Konstruktion des Plasmids DS-2008-2-3, welches rHia-Proteine in voller Länge aus dem NTHi-Stamm 11 exprimiert.

[0125] Chromosomale DNA wurde aus dem NTHi-Stamm 11 gereinigt, und das hia-Gen in voller Länge wurde unter Verwendung der Oligonukleotide (5038.SL und 5039.SL), die in [Fig. 1B](#) beschrieben werden, mit PCR amplifiziert. Eine NdeI-Stelle wurde an dem 5'-Ende des Gens gentechnisch eingeführt, und eine BamHI-Stelle wurde an dem 3'-Ende zur Klonierung in den pT7-7-Expressionsvektor (Ref. 21) gentechnisch eingeführt. Das amplifizierte Fragment wurde mit NdeI/BamHI verdaut und in pT7-7 kloniert, welcher mit denselben Enzymen verdaut worden war. Das Plasmid DS-2008-2-3 enthält ein hia-Gen aus Stamm 11 mit 3,4 kb stromabwärts des T7-Promotors ([Fig. 1A](#)). Das Plasmid wurde verwendet, um rekombinantes Hia zu exprimieren (Beispiel 9 unten).

Beispiel 2

[0126] Dieses Beispiel veranschaulicht die Erkennung von rHia durch einen anti-natives Adhäsins mit hohem Molekulargewicht aus *Moraxella catarrhalis*-Antikörper.

[0127] Es gibt eine gewisse Sequenzkonservierung, die zwischen den *Haemophilus influenzae*-Hia-Proteinen und einem Adhäsins mit hohem Molekulargewicht aus *Moraxella catarrhalis*, das in dem vorstehend erwähnten U.S.-Patent Nr. 5,808,024 als das *M. catarrhalis* 200 kDa-Protein identifiziert wurde, beobachtet wurde ([Fig. 28](#)). Das native *M. catarrhalis* 200 kDa-Protein wurde wie in dem U.S.-Patent Nr. 5,808,024 beschrieben gelgereinigt, und anti-natives 200 kDa-Antikörper aus Meerschweinchen wurde erzeugt. Das T7-hia-Gen wurde aus dem Plasmid DS-2008-2-3 exprimiert, und die Zellkultur, welche das rHia-Protein enthielt, wurde

auf eine Nitrocellulosemembran elektrogeblottet. Eine Immunoblotanalyse unter Verwendung des anti-natives 200 kDa-Antikörpers zeigte, dass der Antikörper das rHia-Protein erkannte, wie man in [Fig. 2](#) sieht.

Beispiel 3

[0128] Dieses Beispiel beschreibt die Konstruktion der Plasmide DS-2092-1 und DS-2092-40, welche Tandemkopien der T7-hia(11)-Genkassetten enthalten.

[0129] Um die Produktion von rekombinantem Hia-Protein in voller Länge zu verbessern, wurden Tandemkopien der T7-hia-Genkassette, welche das hia-Gen von Stamm 11 enthielt (Beispiel 1), in einen einzigen Vektor eingeführt. Das Plasmid DS-2008-2-3 wurde mit BglII linearisiert und dephosphoryliert. Das Plasmid DS-2008-2-3 wurde ebenfalls mit BglII und BamHI verdaut, um die T7-hia-Genkassette auszuschneiden. Das T7-hia-Fragment wurde in den linearisierten Vektor ligiert, um das Plasmid DS-2092-1, welches zwei Kopien des T7-hia-Gens in einer Gegenuhrzeigersinn-Orientierung (a, a) enthält, und das Plasmid DS-2092-40, welches Tandemkopien in entgegengesetzten Orientierungen (a, c) enthält, zu erzeugen ([Fig. 3](#)). Es gab keine offensichtliche Verbesserung bei der Expression von rHia bei irgendeinem Konstrukt (siehe Beispiel 9 unten).

Beispiel 4

[0130] Dieses Beispiel beschreibt die Konstruktion von Plasmiden, welche trunkierte hia-Gene aus Stamm 11 exprimieren.

[0131] Die Produktion des rHia-Proteins aus einzelnen oder Tandemkopien der T7-hia-Genkassette war sehr gering, und das Protein schien für *E. coli* toxisch zu sein (wie unten in Beispiel 9 beschrieben wird). Da *H. influenzae* Hia ein an der Oberfläche exponiertes Adhäsins-Molekül ist, muss es entweder eine Signalsequenz oder ein akzessorisches Protein(e) zur Sekretion benutzen, es sind aber keine bekannten akzessorischen Gene daran beteiligt. Wenn die Signalsequenz zur Expression des rekombinanten Proteins in *E. coli* entfernt würde, könnte das rHia in Form von Einschlusskörpern exprimiert werden und die toxische Wirkung verringert werden. Eine mutmaßliche Signalsequenz und Spaltungsstellen wurden identifiziert, und vier Konstrukte, die N-terminal trunkierte Hia-Proteine exprimierten, wurden entworfen ([Fig. 4](#)). Es gibt eine einmalige Styl-Stelle in dem hia-Gen des Stamms 11, ca. 500 bp von dem Startcodon entfernt. Das Plasmid DS-2008-2-3 wurde mit NdeI und Styl verdaut und das Vektorfragment mit 5,7 kb gereinigt ([Fig. 5A](#)). PCR-Primer wurden entworfen, um ausgehend von der Trunkierungsstelle bis zu der Styl-Stelle zu amplifizieren, und eine einmalige NheI-Stelle wurde in den Gegensinn-Primer eingeführt, um auf trunkierte Klone zu screenen ([Fig. 5B](#)). Die amplifizierten Fragmente wurden zur leichteren Handhabung in pCRII subkloniert, was die Plasmide DS-2153R-1-2 (E21), DS-2165-4-8 (T33), DS-2153-3-5 (V38) und DS-2153-4-4 (N52) erzeugte. Die pCRII-hia-Plasmide wurden mit NdeI und Styl verdaut und die Fragmente mit dem Vektorstück aus DS-2008-2-3 ligiert. Die Plasmide DS-2186-1-1 (E21), DS-2201-1 (T33), DS-2186-2-1 (V38) und DS-2168-2-6 (N52) wurden erzeugt, welche den T7-Promotor und trunkierte hia-Gene enthielten, wie in Klammern angegeben ist. Diese Plasmide wurden verwendet, um rekombinantes Hia zu exprimieren (siehe Beispiel 9 unten).

Beispiel 5

[0132] Dieses Beispiel beschreibt die Konstruktion des Plasmids BK-96-2-11, welches die T7-V38-hia(11)-Kassette, das *E. coli* cer-Gen und das Kanamycinresistenzgen enthält.

[0133] Das Plasmid DS-1843-2 ist ein auf pBR328 basierendes Plasmid, in welches eine multiple Klonierungsstelle und zwei Transkriptionsterminatoren auf Oligonukleotiden zwischen den EcoRI- und PstI-Stellen eingeführt wurden, wodurch sowohl das Chloramphenicol- als auch das Ampicillinresistenzgen zerstört wurde ([Fig. 6B](#)). Das Kanamycinresistenzgen aus pUC-4K wurde an der Sall-Stelle eingeführt, um das Plasmid DS-2147-1 zu erzeugen, welches kanamycinresistent und tetracyclinempfindlich ist. Das Plasmid DS-2224-1-4 ist ein pUC-Plasmid, welches ein synthetisches *E. coli* cer-Gen enthält (Ref. 15), das aus Oligonukleotiden konstruiert wurde und durch BamHI-Stellen flankiert ist. Das BamHI-Fragment mit 290 bp des cer-Gens wurde in die BamHI-Stelle von DS 2147-1 eingefügt, was das Plasmid BK-2-1-2 erzeugte. Dieses auf pBR basierende Plasmid enthält somit eine multiple Klonierungsstelle, das Kanamycinresistenzgen und das cer-Gen. Das Plasmid BK-2-1-2 wurde mit BglII linearisiert und dephosphoryliert. Das Plasmid DS-2186-2-1 wurde mit BglII und BamHI verdaut, und das T7-V38-hia-Fragment mit 3,6 kb wurde in BK-2-1-2 eingefügt, was das Plasmid BK 96-2-11 erzeugte ([Fig. 6A](#)).

Beispiel 6

[0134] Dieses Beispiel beschreibt die Konstruktion der Plasmide DS-2242-1 und DS-2242-2, welche das hia-Gen des NTHi-Stamms 33 in voller Länge in der Gegenwart des E. coli cer-Gens exprimieren.

[0135] Chromosomale DNA wurde aus dem NTHi-Stamm 33 gereinigt, und eine PCR-Amplifikation wurde durchgeführt, indem die Oligonukleotide 5039.SL und 5040.SL verwendet wurden ([Fig. 17](#)). Der Sinn-Primer (5040.SL) wurde auf der Basis der 5'-flankierenden Sequenz von hia aus Stamm 11 und der konservierten aminoterminalen Sequenzen der NTHi-Hia- und Hib-Hsf-Proteine entworfen. Der Gegensinn-Primer (5039.SL) war derselbe wie der, der in Beispiel 1 beschrieben wird, und dieser basierte auf den konservierten carboxyterminalen Sequenzen der Hia- und Hsf-Proteine. Das hia-PCR-Fragment aus Stamm 33 mit 3 kb wurde in pCRII kloniert, was das Plasmid DS-1917-3-8 erzeugte.

[0136] Um das hia-Gen des Stamms 33 in voller Länge zu exprimieren, wurden ungefähr 106 bp des 5'-Endes des Gens ausgehend von dem Startcodon bis zu einer AlwNI-Stelle aus Oligonukleotiden synthetisiert ([Fig. 7B](#)). Das Plasmid DS-1917-3-8 wurde mit AlwNI und BamHI verdaut, und das Fragment mit ungefähr 2,9 kb, welches das hia-Gen enthielt, wurde gereinigt. Das Plasmid pT7-7 wurde mit NdeI und BamHI verdaut. Die NdeI-AlwNI-Oligonukleotide und das AlwNI-BamHI-hia-Fragment wurden in den pT7-7-Vektor ligiert, was das Plasmid DS-2103-4 erzeugte.

[0137] Um das E. coli cer-Gen aufzunehmen und eine Kanamycinselektion zu benutzen, wurde das BglII-BamHI-Fragment, welches die T7-hia(33)-Genkassette enthielt, aus DS-2103-4 ausgeschnitten und in BK-2-1-1 kloniert, welches mit BglII verdaut und dephosphoryliert worden war. Die Plasmide DS-2242-1 und DS-2242-2 enthalten einzelne Kopien der T7-hia(33)-Genkassette in entgegengesetzten Orientierungen, das E. coli cer-Gen und das Kanamycinresistenzgen ([Fig. 7A](#)).

Beispiel 7

[0138] Dieses Beispiel beschreibt die Konstruktion des Plasmids DS-2340-2-3, welches eine T7-hia-Genkassette mit einem trunkierten V38-hia-Gen aus Stamm 33, das E. coli cer-Gen und das Kanamycinresistenzgen enthält.

[0139] PCR-Primer wurden entworfen, um ein 250 bp-Fragment des 5'-Endes des hia-Gens aus NTHi-Stamm 33 ausgehend von einem V38-Startcodon bis zu einer inneren SnaBI-Stelle zu amplifizieren. Eine NdeI-Stelle wurde an dem 5'-Ende zu Klonierungszwecken angefügt, und das Fragment wurde amplifiziert, indem das Plasmid DS-2242-1 als Matrize verwendet wurde. Das Konstruktionsschema ist in [Fig. 8A](#) gezeigt, und die PCR-Primer sind in [Fig. 8B](#) gezeigt. Das Fragment wurde in pCRII kloniert, was das Plasmid DS-2328-1-1 erzeugte. DS-2242-1 wurde mit NdeI und SnaBI verdaut, und das Vektorfragment mit 8,5 kb gereinigt. DS-2328-1-1 wurde mit NdeI und SnaBI verdaut, und das 5'-hia-Fragment mit 0,25 kb wurde mit dem Vektorfragment mit 8,5 kb aus DS-2242-1 ligiert, um das Plasmid DS-2340-2-3 zu erzeugen.

Beispiel 8

[0140] Dieses Beispiel stellt die Konstruktion der Plasmide DS-2447-2 und DS-2448-17 dar, welche Tandemkopien der T7-V38-hia(11)- bzw. T7-V38-hia(33)-Genkassetten, das E. coli cer-Gen und ein Kanamycinresistenzgen enthalten.

[0141] Das Plasmid BK-96-2-11, welches eine T7-V38-hia(11)-Genkassette enthält, wurde mit BglII linearisiert und dephosphoryliert. Die BglII-BamHI-T7-V38-hia(11)-Genkassette aus DS-2186-2-1 wurde in BK-96-2-11 ligiert, was das Plasmid DS-2447-2 erzeugte, welches Tandemkopien des T7-V38-hia(11)-Gens in derselben Orientierung enthält ([Fig. 9A](#)).

[0142] Das Plasmid DS-2340-2-3 wurde mit EcoRI verdaut, und die T7-V38-hia(33)-Genkassette wurde in pUC-BgXb subkloniert, welches mit EcoRI verdaut und dephosphoryliert worden war. Das resultierende Plasmid, DS-2440-2, wurde mit BglII und BamHI verdaut, um die T7-V38-hia(33)-Kassette freizusetzen, die mit DS-2340-2-3 ligiert wurde, welches mit BglII linearisiert und dephosphoryliert worden war. Das Plasmid DS-2448-17 enthält Tandem-T7-V38-hia(33)-Gene in derselben Orientierung ([Fig. 9B](#)).

Beispiel 9

[0143] Dieses Beispiel veranschaulicht die Expression von rekombinanten *hia*-Genen in voller Länge und trunziert.

[0144] DNA aus Expressionsplasmiden, die wie in den vorhergehenden Beispielen beschrieben hergestellt wurden, wurde in elektrokompente *E. coli* BL21 (DE3)-Zellen unter Verwendung eines BioRad Elektroporators eingeführt. Die Zellen wurden bei 37°C in NZCYM-Medium unter Verwendung der geeigneten Antibiotikaselektion bis auf ein A_{578} von 0,3 kultiviert, bevor Lactose für 4 Stunden bis auf 1,0% zugegeben wurde. Proben wurden mit SDS-PAGE-Lyse + Beladungspuffer auf 0,2 OD/μl eingestellt, und dieselbe Menge jeder Proteinprobe wurde auf SDS-PAGE-Gele geladen (Ref. 22). [Fig. 10](#) stellt die relative Produktion von rHia(11)-Proteinen aus verschiedenen Konstrukten dar. Wie man in Feld A sieht, gibt es eine Zunahme bei der Produktion bei einer verringerten Größe von rHia. V38-(Bahn 5) und N52-trunkiertes rHia (Bahn 6) weisen signifikant höhere Expressionsspiegel auf als ihre längeren Gegenstücke (Bahnen 2, 3, 4). Zusätzlich zeigt Feld B, dass die Produktion von V38-rHia in der Gegenwart des *cer*-Gens offensichtlich zunimmt.

Beispiel 10

[0145] Dieses Beispiel stellt die Reinigung von rHia-Proteinen dar.

[0146] All die rekombinanten Hia-Proteine wurden als Einschlusskörper in *E. coli* exprimiert und wurden durch denselben Vorgang gereinigt ([Fig. 11](#)). *E. coli*-Zellpellets aus 500 ml Kultur wurden in 50 ml 50 mM Tris-HCl, pH 8,0, enthaltend 0,1 M NaCl, resuspendiert und durch Beschallung aufgebrochen. Der Extrakt wurde für 30 min bei 20.000 g zentrifugiert, und der resultierende Überstand wurde verworfen. Das Pellet (PPT₁) wurde weiter in 50 ml 50 mM Tris-HCl, pH 8,0, enthaltend 0,5% Triton X-100 und 10 mM EDTA, extrahiert, dann für 30 min bei 20.000 g zentrifugiert, und der Überstand wurde verworfen. Das Pellet (PPT₂) wurde weiter in 50 ml 50 mM Tris-HCl, pH 8,0, enthaltend 1% Octylglucosid, extrahiert, dann für 30 min bei 20.000 g zentrifugiert, und der Überstand wurde verworfen.

[0147] Das resultierende Pellet (PPT₃), das nach den obigen Extraktionen enthalten wurde, enthält die Einschlusskörper. Das Pellet wurde in 6 ml 50 mM Tris-HCl, pH 8,0, enthaltend 6 M Guanidin und 5 mM DTT, solubilisiert. Zwölf ml 50 mM Tris-HCl, pH 8,0, wurden zu dieser Lösung zugegeben, und die Mischung wurde für 30 min bei 20.000 g zentrifugiert. Der Überstand (SUP₄) wurde mit Polyethylenglycol (PEG) 4000 bei einer Endkonzentration von 7% ausgefällt. Das resultierende Pellet (PPT₅) wurde durch Zentrifugation für 30 min bei 20.000 g entfernt, und der Überstand wurde durch (NH₄)₂SO₄ bei 50% Sättigung ausgefällt. Der (NH₄)₂SO₄-Niederschlag wurde durch Zentrifugation für 30 min bei 20.000 g gesammelt. Das resultierende Pellet (PPT₆) wurde in 2 ml 50 mM Tris-HCl, pH 8,0, enthaltend 6 M Guanidin-HCl und 5 mM DTT, gelöst, und die klare Lösung wurde auf einer Superdex 200 Gelfiltrationssäule, welche in 50 mM Tris-HCl, pH 8,0, enthaltend 2 M Guanidin-HCl, äquilibriert war, gereinigt. Die Fraktionen wurden durch SDS-PAGE analysiert und jene, welche das gereinigte rHia enthielten, wurden gepoolt und über Nacht bei 4°C gegen PBS dialysiert, dann für 30 min bei 20.000 g zentrifugiert. Das Protein blieb unter diesen Bedingungen löslich, und Glycerol wurde zu der rHia-Präparation bei einer Endkonzentration von 20% zur Lagerung bei -20°C zugegeben. Eine SDS-PAGE-Analyse von gereinigtem V38-rHia(11) und V38-rHia(33) ist in [Fig. 12](#) dargestellt. Die durchschnittliche Ausbeute der gereinigten V38-rHia-Proteine beträgt ca. 10 mg L⁻¹ Kultur.

[0148] Um die Stabilität von rHia zu untersuchen, wurde das gereinigte V38-rHia(11)-Protein bei 4°C mit oder ohne Glycerol und bei -20°C mit Glycerol gelagert. Es wurde gefunden, dass das Protein unter allen drei Bedingungen stabil war und bei wiederholtem Einfrieren und Auftauen für wenigstens acht Wochen intakt blieb ([Fig. 13](#)).

Beispiel 11

[0149] Dieses Beispiel stellt die Immunogenität von V38-rHia(11)- und V38-rHia(33)-Proteinen dar.

[0150] Hyperimmunantiseren gegen rHia-Proteine wurden erzeugt, indem zwei Meerschweinchen (Charles River) intramuskulär (i.m.) mit 5 μg-Dosen von Antigen, das in Freund's vollständigem Adjuvans (CFA, Difco) emulgiert war, am Tag 1 immunisiert wurden. Die Tiere erhielten an den Tagen 14 und 28 5 μg-Dosen an Protein in Freund's unvollständigem Adjuvans (IFA) als Booster, und die Seren wurden am Tag 42 gesammelt. Anti-Hib-Stamm MinnA- und anti-Haemophilus Typ a-Stamm ATCC 9006-Antiseren wurden erzeugt, indem dieselbe Vorschrift verwendet wurde, außer dass eine hitzeinaktivierte Bakterienpräparation als das Immunogen

(1×10^8 cfu pro Dosis) verwendet wurde.

[0151] Um die Immunogenität der V38-rHia-Proteine zu untersuchen, wurden Gruppen von fünf CD-1-Mäusen (Charles River, Quebec) an den Tagen 1 und 28 mit 0,3, 1, 3 und 10 µg Antigen in der Gegenwart von AlPO_4 (Alaun) (1,5 mg pro Dosis) s.c. immunisiert. Blutproben wurden an den Tagen 1, 28 und 42 gesammelt. Mäuse erzeugten signifikante anti-V38-rHia-Antikörperantworten, selbst bei einer einzigen Injektion von 0,3 µg Antigen (**Fig. 14**, Feld A), was nahe legt, dass beide Proteine die Immunogenität nach Extraktion der Einschlusskörper und Solubilisierung beibehielten. Es wurde kein statistisch signifikanter Unterschied bei den Antikörpertitern gefunden, welche durch die V38-rHia-Proteine, die von den Stämmen 11 oder 33 abgeleitet waren, induziert wurden.

[0152] Um die Immunogenität des V38-rHia(11)-Proteins in BALG/c-Mäusen zu untersuchen, wurden Gruppen von fünf Tieren (Charles River, Quebec) an den Tagen 1, 28 und 42 mit 0,3, 1, 3 und 10 µg Antigen in der Gegenwart von AlPO_4 (1,5 mg per Dosis) s.c. immunisiert. Blutproben wurden an den Tagen 1, 14, 28, 42 und 56 gesammelt. Hohe Antikörpertiter wurden in allen Gruppen beobachtet, was zeigt, dass das Protein selbst bei 0,3 µg pro Dosis sehr immunogen ist (**Fig. 15**, Feld A).

[0153] Um die Immunogenität des V38-rHia(11)-Proteins in Meerschweinchen zu untersuchen, wurden Gruppen von fünf Tieren (Charles River, Quebec) an den Tagen 1, 28 und 42 mit 0,3, 1, 3, und 10 µg Antigen in der Gegenwart von AlPO_4 (1,5 mg pro Dosis) s.c. immunisiert. Blutproben wurden an den Tagen 1, 14, 28, 42 und 56 gesammelt. Es wurden bei allen Gruppen hohe Antikörpertiter beobachtet, was zeigt, dass das Protein ebenfalls in Meerschweinchen sehr immunogen ist (**Fig. 15**, Feld B).

Beispiel 12

[0154] Dieses Beispiel stellt die Analyse des Schutzes dar, der durch anti-rHia-Antikörper in passiven Babyrattenmodellen der Bakteriämie geboten wird.

[0155] Schwangere Wistar-Ratten wurden bei Charles River gekauft. Bei dem H. influenzae Typ b-Bakteriämiemodell wurde Gruppen von 6 bis 10 fünf Tage alten Babyratten s.c. im Dorsalbereich 0,1 ml Meerschweinchen anti-rHia- oder anti-Stamm MinnA-Antiserum injiziert. Die Kontrolltiere erhielten Injektionen nur mit Präimmunserum. 20 Stunden später wurden die Tiere intraperitoneal (i.p.) mit 200 bis 240 koloniebildenden Einheiten (cfu) von frisch kultiviertem Hib-Stamm MinnA (0,1 ml) herausgefordert. Blutproben wurden 20 h nach der Herausforderung über eine Herzpunktion unter Betäubung mit Isofluran gesammelt und auf Schokoladen-Agar-Platten plattiert. Die Kolonien wurden nach einem Tag gezählt, und die Ergebnisse wurden statistisch durch den Exact-Test von Fischer analysiert.

[0156] Bei dem H. influenzae Typ a-Bakteriämiemodell (Ref. 23) wurde Gruppen von 9 bis 10 fünf Tage alten Babyratten im Dorsalbereich 0,1 ml Meerschweinchen anti-rHia- oder anti-Stamm ATCC 9006-Antiserum injiziert. Den Tieren in der Kontrollgruppe wurde Meerschweinchen-Präimmunserum injiziert. Zwanzig Stunden später wurden die Tiere i.p. mit 100.000 cfu von frisch kultiviertem H. influenzae Typ a-Stamm ATCC 9006 (0,1 ml) herausgefordert. Blutproben wurden 20 h nach der Herausforderung gesammelt und wie oben beschrieben analysiert.

[0157] Wie in den Tabellen 1 und 2 unten gezeigt ist, waren die Babyratten, die passiv mit entweder Meerschweinchen anti-rHia(11)- oder anti-V38-rHia(11)-Antiseren immunisiert wurden, alle signifikant gegen eine von H. influenzae Typ a oder Typ b verursachte Bakteriämie geschützt. Diese Ergebnisse zeigen, dass Antikörper, die gegen leicht trunkiertes Hia-Protein (V38-rHia) hervorgerufen wurden, ebenso wirksam sind wie jene, die gegen das Protein in voller Länge hervorgerufen wurden, um Tiere gegen eine Bakteriämie zu schützen, die durch H. influenzae Typ a oder Typ b hervorgerufen wird. Ein solcher Schutz, der von einem von NTHi abgeleiteten rekombinanten Protein vor einer invasiven Erkrankung, die durch eingekapselte Bakterien verursacht wird, geboten wird, veranschaulicht den Nutzen der rHia-Proteine als Impfstoffantigene.

Beispiel 13

[0158] Dieses Beispiel stellt den Schutz dar, der durch Immunisierung mit V38-rHia-Protein in einem Chinchillamodell der nasopharyngealen Besiedlung geboten wird.

[0159] Ein nasopharyngeales Besiedlungsmodell wurde von Yang et al. (Ref. 20) beschrieben. Das Modell arbeitet gut für jene NTHi-Stämme, die HMW-Adhäsine produzieren, aber eine reproduzierbare Besiedlung

konnte mit Hia-produzierenden Stämmen unter denselben Bedingungen nicht erreicht werden. Wiederholte Versuche, mit dem Prototyp des Hia-produzierenden NTHi-Stamms 11 zu besiedeln, waren nicht erfolgreich. Eine Besiedlung wurde mit dem NTHi-Stamm 33 bei 5×10^8 cfu pro Inokulum erreicht, im Vergleich zu nur 10^8 cfu, die für den Prototyp des HMW-produzierenden NTHi-Stamms 12 benötigt wurden. Unter diesen Bedingungen wurde ein partieller Schutz bei Tieren beobachtet, die mit 100 µg V38-rHia(33) immunisiert wurden und mit dem homologen NTHi-Stamm 33 herausgefordert wurden.

Beispiel 14

[0160] Dieses Beispiel stellt die Klonierung und Sequenzanalyse von zusätzlichen hia-Genen aus *H. influenzae*-Stämmen dar.

[0161] Oligonukleotide (5040.SL und 5039.SL) zur PCR-Amplifikation wurden auf der Basis des konservierten Promotors, der N-terminalen und C-terminalen Sequenzen der hia- und hsf-Gene und Proteine entworfen ([Fig. 17](#)). Die Stämme, die zur PCR-Amplifikation ausgewählt wurden, wurden auf der Basis ihrer Reaktivität mit anti-rHia(11)-Antisera ausgewählt.

[0162] Chromosomale DNA wurde aus den NTHi-Stämmen 12, 29, 32, M4071, K9 und K22 und dem Haemophilus Typ c-Stamm API hergestellt. Eine PCR-Amplifikation wurde wie folgt durchgeführt: Jede Reaktionsmischung enthielt 5 bis 100 ng DNA, 1 µg jedes Primers, 5 Einheiten taq+ oder tsg+ (Sangon) oder taq plus Jong (Stratagene), 2 mM dNTPs, 20 mM Tris-HCl (pH 8,8), 10 mM KCl, 10 mM $(\text{NH}_4)_2\text{SO}_4$, 2 mM MgSO_4 , 0,1% Triton X-100, BSA. Die Zyklusbedingungen waren: 95°C für 1 min, gefolgt von 25 Zyklen von 95°C für 30 Sek., 45°C für 1 min, 72°C für 2 min; dann 72°C für 10 min.

[0163] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm 33 sind in **Fig. 18** gezeigt. Das vorhergesagte Hia-Protein aus dem Stamm 33 weist ein Molekulargewicht von 103,6 kDa und einen pI von 9,47 auf. Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm 32 sind in **Fig. 19** gezeigt. Das vorhergesagte Hia-Protein aus dem Stamm 32 weist ein Molekulargewicht von 70,4 kDa und einen pI von 5,67 auf. Es gibt eine KDEL-Sequenz, die zwischen den Resten 493 und 496 vorliegt. Solche Sequenzen wurden mit der Verankerung von Proteinen an dem endoplasmatischen Retikulum in Verbindung gebracht. Das abgeleitete Hia-Protein des Stamms 32 ist signifikant kleiner und weist einen signifikant verschiedenen pI auf, jedoch enthält es viele der Motive, die in anderen Hia-Molekülen vorliegen.

[0164] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm 29 sind in **Fig. 20** gezeigt. Das vorhergesagte Hia-Protein aus dem Stamm 29 weist ein Molekulargewicht von 114,4 kDa und einen pI von 7,58 auf. Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm K22 sind in **Fig. 23** gezeigt. Das vorhergesagte Hia-Protein aus dem Stamm K22 weist ein Molekulargewicht von 114,4 kDa und einen pI von 7,58 auf. Es wurde gefunden, dass die abgeleiteten Hia-Sequenzen aus den NTHi-Stämmen 29 und K22 identisch waren. Der Stamm 29 wurde aus einem 7 Monate alten Kind mit Otitis media in Cleveland, Ohio, isoliert, während der Stamm K22 aus einem Aborigine nahe Kimberly, Australien, isoliert wurde.

[0165] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm 4071 sind in **Fig. 21** gezeigt. Das vorhergesagte Hia-Protein aus dem Stamm M4071 weist ein Molekulargewicht von 103,4 kDa und einen pI von 9,49 auf. Es gibt eine KDEL-Sequenz, die zwischen den Resten 534 und 537 vorliegt.

[0166] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Stamm K9 sind in **Fig. 22** gezeigt. Das vorhergesagte Hia-Protein aus K9 weist ein Molekulargewicht von 113,8 kDa und einen pI von 6,45 auf.

[0167] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Gens aus dem Typ c Haemophilus-Stamm API sind in **Fig. 24** gezeigt. Das vorhergesagte Hia-Protein aus API weist ein Molekulargewicht von 249,4 kDa und einen pI von 5,34 auf. Die abgeleitete Hia/Hsf-Sequenz aus dem Typ c-Stamm API ist nahezu identisch mit der veröffentlichten Typ b-Hsf-Sequenz, außer für ein Insert mit 60 Resten. Da das auf NTHi basierende Hia-Protein, das hierin bereitgestellt wird, in passiven Modellen einer Typ a- und Typ b-Infektion schützt, ist es wahrscheinlich, dass es aufgrund der Sequenzähnlichkeit zwischen den Proteinen vom Typ b und Typ c ebenso gegen eine Typ c-Erkrankung schützen wird.

[0168] Die Nukleotid- und die abgeleitete Aminosäuresequenz des hia-Locus aus dem Stamm 12 sind in

Fig. 25 gezeigt. Der NTHi-Stamm 12 produziert kein Hia. Jedoch kann ein Teil des hia-Gens PCR-amplifiziert werden, es gibt eine inkonsistente positive Reaktivität von SB12-Zelllysaten mit dem anti-rHia-Antikörper, und es gibt eine Reaktivität mit einer DNA-Sonde, die von dem 3'-Ende des hia-Gens des Stamms 11 abgeleitet ist, auf Southern Blots. Eine Analyse der PCR-amplifizierten DNA zeigte ein 1,8 kb-Fragment, welches 1 kb des 3'-Endes des stromaufwärts liegenden HI1732-verwandten Gens und 0,8 kb des 3'-Endes des hia-Gens enthält.

[0169] Eine PCR-Amplifikation, die Primer verwendete, die über die mutmaßliche Verknüpfung dieser zwei Gene in dem Stamm 12 hinweg amplifizieren würde, bestätigte die genetische Zusammensetzung des Locus. Somit würde es scheinen, dass der Stamm 12 kein Hia produziert, da er eine Deletion des 5'-Endes des hia-Gens erfahren hat. **Fig. 27** zeigt einen Sequenzvergleich zwischen dem stromaufwärts liegenden orf von Stamm 12 und dem aus dem Rd-Genom abgeleiteten HI1733-Protein. Über den Homologiebereich hinweg sind die zwei Proteine zu 95% identisch.

[0170] Eine Ausrichtung der abgeleiteten Hia-Sequenzen aus den NTHi-Stämmen 33, 32, 29, K22, M4071, 11 und K9 und dem Typ c-Stamm API im Vergleich mit *H. influenzae* Typ b-Hsf, dem aidA-ähnlichen (Hsf/Hia) HI1732-Gen aus dem Rd-Genom und dem *M. catarrhalis* 200 kDa-Protein aus den Stämmen 4223 und LES-1 ist in **Fig. 28** gezeigt. Hier gibt es eine Rasterverschiebung bei der Rd-Genomsequenz, welche zu einer vorzeitigen Trunkierung des HI1732-Proteins führt. Eine zusätzliche stromabwärts liegende Sequenz, die mit hia verwandt ist, ist hier umfasst. Die Sterne unter der Sequenz zeigen konservierte Reste. Die N-terminalen (ungefähr 50 Reste) und C-terminalen Sequenzen (ungefähr 150 Reste) sind bei den *Haemophilus*-Stämmen in hohem Maße konserviert, während eine gewisse Ähnlichkeit mit dem Gegenpart von *M. catarrhalis* ersichtlich ist. Eine Sequenzanalyse zeigt, dass es zwei potentielle Genfamilien von Hia-Proteinen gibt, von denen eine mit dem Prototypstamm 11 verwandt ist und die andere enger mit dem Stamm 33 verwandt ist. Die Proteine der Stämme 11 und K9 scheinen mehr Ähnlichkeit zu den Hsf-Proteinen aus den *Haemophilus*-Stämmen vom Typ b, Typ c oder Typ d aufzuweisen, während die Proteine der Stämme 33, 32, 29, K22 und M4071 eine zweite Familie zu bilden scheinen.

Beispiel 15

[0171] Dieses Beispiel beschreibt die Konstruktion des Plasmids JB-2930-3, welches eine T7-hia-Genkassette mit einem trunkierten S44-hia-Gen des Stamms 11, das *E. coli* cer-Gen und das Kanamycin-Antibiotikaresistenzgen enthält, und die Expression von S44-Hia-Proteinen.

[0172] PCR-Primer wurden entworfen, um den S44-Hia-N-Terminus des hia-Gens des NTHi-Stamms 11 ausgehend von der Aminosäure S44 bis zu einer inneren Styl-Stelle zu amplifizieren (**Fig. 29**). Eine NdeI-Stelle wurde an dem 5'-Ende zu Klonierungszwecken angefügt, und das Fragment wurde unter Verwendung des Plasmids DS-2242-1 als einer Matrize amplifiziert. Das Fragment wurde in pCRII kloniert, was das Plasmid JB-2910-1-1 erzeugte. Das Konstruktionsschema ist in **Fig. 30** gezeigt. Das Plasmid JB-2910-1-1 wurde mit NdeI und Styl verdaut und das 5'-PCR-hia-Fragment isoliert. Das Plasmid IA-46-5, welches das V38-hia-Gen enthielt, wurde mit NdeI und Styl verdaut und das größere Fragment mit ungefähr 8,5 kb gereinigt. Die zwei gereinigten Fragmente wurden zusammen ligiert, um das Plasmid JB-2917-1 zu erzeugen. Dieses Plasmid wurde dann mit NdeI verdaut und mit Kalbsintestinumphosphatase (CAP) behandelt, und in dieses wurde der T7-Promotor aus dem Plasmid IA-46-5 kloniert. Der Promotor wurde herausgeschnitten, indem ein NdeI-Verdau von IA-46-5 verwendet wurde. Das resultierende Plasmid, JB-2925-3, wurde mit BglII und BamHI verdaut, und das hia-Gen wurde isoliert. Dieses Fragment wurde in das mit BglII/CAP behandelte Plasmid BK-2-1-2 ligiert, um das Plasmid JB-2930-3 zu erzeugen. Dieses Plasmid enthält den T7-Promotor, das S44-hia-Gen und das *E. coli* cer-Gen und eine Kanamycinresistenz.

[0173] Der rekombinante S44-hia-Vektor wurde für Expressionsstudien in *E. coli* BL21 (DE3) transformiert. Die Prozedur zur Expression in *E. coli* war wie in Beispiel 9 beschrieben. **Fig. 31**, eine SDS-PAGE-Analyse, zeigt die Expression von rekombinantem S44-hia aus zwei verschiedenen Vektoren, JB-2930-3 (oben beschrieben) und dem pET-Vektor IA-191-3-1. Das Plasmid IA-191-3-1 ist mit JB-2930-3 identisch, außer dass dieses ein pET-Vektor ist, welcher den lacI^q-Repressor enthält, und daher die produzierte Menge an S44-Hia niedriger als die des T7-S44 aus JB-2930-3 ist. Das Plasmid ist zusammen mit dem Plasmid JB-2930-3 gezeigt, **Fig. 32**. **Fig. 31** zeigt das S44-Hia als eine Doubletbande (Bahn 3) bei ungefähr 116 kDa. Bei einer weiteren Analyse unter Verwendung von gereinigtem S44-hia aus JB-2930-3 wurde gefunden, dass die untere Bande des Doublets eine C-terminale Trunkierung von 94 Aminosäuren aufweist, während sie den erwarteten N-Terminus beibehält. Der Reinigungsprozess, der zur Isolation des trunkierten Hia verwendet wurde, war wie in Beispiel 10 beschrieben.

ZUSAMMENFASSUNG DER OFFENBARUNG

[0174] Als Zusammenfassung dieser Offenbarung stellt die vorliegende Erfindung neue isolierte und gereinigte Nukleinsäuremoleküle bereit, welche N-terminal trunke Haemophilus influenzae Adhäsine(Hia)-Proteine aus Haemophilus codieren, welche ermöglichen, dass schützende Hia-Proteine rekombinant produziert werden.

REFERENZEN

1. Barbour, M. L., R. T. Mayon-White, C. Coles, D. W. M. Crook und E. R. Moxon. 1995. The impact of conjugate vaccine on carriage of Haemophilus influenzae type b. J. Infect. Dis. 171: 93–98.
2. Berkowitz et al. 1987. J. Pediatr. 110: 509.
3. Claesson et al. 1989. J. Pediatr. 114: 97.
4. Black, S. B., H. R. Shinefield, B. Fireman, R. Hiatt, M. Polen, E. Vittinghoff, The Northern California Kaiser Permanent Vaccine Study Center Pediatrics Group. Efficacy in infancy of oligosaccharide conjugate Haemophilus influenzae type b (HbOC) vaccine in a United States population of 61,080 children. 1991. Pediatr. Infect. Dis. J. 10: 97–104.
5. Nitta, D. M., M. A. Jackson, V. F. Burry und L. C. Olson. 1995. Invasive Haemophilus influenzae type f disease. Pediatr. Infect. Dis. J. 14: 157–160.
6. Waggoner-Fountain, L. A., J. O. Hendley, E. J. Cody, V. A. Perriello und L. G. Donowitz. 1995. The emergence of Haemophilus influenzae types e and f as significant pathogens. Clin. Infect. Dis. 21: 1322–1324.
7. Madore, D. V. 1996. Impact of immunization on Haemophilus influenzae type b disease. Infectious Agents and Disease 5: 8–20.
8. Bluestone, C. D. 1982. Current concepts in otolaryngology. Otitis media in children: to treat or not to treat? N. Engl. J. Med. 306: 1399–1404.
9. Barenkamp, S. J. und E. Leininger. 1992. Cloning, expression, and DNA sequence analysis of genes encoding nontypeable Haemophilus influenzae high-molecular-weight surface-exposed proteins related to filamentous hemagglutinin of Bordetella pertussis. Infect. Immun. 60: 1302–1313.
10. St. Geme III, J. W., S. Falkow und S. J. Barenkamp. 1993. High-molecularweight proteins of nontypeable Haemophilus influenzae mediate attachment to human epithelial cells. Proc. Natl. Acad. Sci. USA 90: 2875–2879.
11. Barenkamp, S. J. 1996. Immunization with high-molecular-weight adhesion proteins of nontypeable Haemophilus influenzae modifies experimental otitis media in chinchillas. Infect. Immun. 64: 1246–1251.
12. St. Geme, J. W. und D. Cutter. 1995. Evidence that surface fibrils expressed by Haemophilus influenzae type b promote attachment to human epithelial cells. Molec. Microbiol. 15: 77–85.
13. Barenkamp, S. J. und J. W. St. Geme. 1996. Identification of a second family of high-molecular-weight adhesion proteins expressed by non-typable Haemophilus influenzae. Molec. Microbiol. 19: 1215–1223.
14. St. Geme, J. W., D. Cutter und S. J. Barenkamp. 1996. Characterization of the genetic locus encoding Haemophilus influenzae type b surface fibrils. J. Bact. 178: 6281–6287.
15. Patient, M. E. und D. K. Summers. 1993. ColE1 multimer formation triggers inhibition of Escherichia coli cell division. Molec. Microbiol. 9: 1089–1095.
16. O'Hagan, D. T. 1992. Oral delivery of vaccines. Formulation and clinical pharmacokinetic considerations. Clin. Pharmacokinet 22(t): 1–10.
17. Ulmer et al. 1993. Curr. Opin. Invest. Drugs 2: 983–989.
18. Lockhoff, O., 1991. Glycolipids as immunomodulators: Synthesis and properties.
19. Nixon-George A., et al., 1990. The adjuvant effect of stearyl tyrosine on a recombinant subunit hepatitis B surface antigen. J. Immunol 144(12): 4798–4802.
20. Yang, Y-P., S. M. Loosmore, B. J. Underdown und M. H. Klein. 1998. Nasopharyngeal colonization with nontypeable Haemophilus influenzae in chinchillas. Infect. Immun. 66: 1973–1980.
21. Tabor, S. und C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. Proc. Natl. Acad. Sci. USA 82: 1074–1078.
22. Laemmli, U.K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680–685.
23. Loosmore, S. M., Y-P. Yang, D. C. Coleman, J. M. Shortreed, D. M. England und M. H. Klein. 1997. Outer membrane protein D15 is conserved among Haemophilus influenzae species and may represent a universal protective antigen against invasive disease. Infect. Immun. 65: 3701–3707.
24. Needleman, S. B. und Wunsch, C. D. 1970, J. Mol. Biol. 48: 443–453.
25. Sellers, P. H. 1974 On the theory and computation of evolutionary distances. J. Appl. Math(Siam) 26: 787–793.
26. Waterman, M. S., Smith, T. F. und Beyer, W. A. 1976. Advan. Math. 20: 367–387.

27. Smith, T. F. und Waterman, M. S. 1981. Identification of common molecular subsequences. J. Mol. Biol. 147: 195–197.
28. Sobel, E. und Martinez, H. M. 1985. A multiple Sequence Alignment Program. Nucleic Acid Res. 14: 363-374.

[0175]

TABELLE 1

Schützende Wirkung von Meerschweinchen-anti-rHia(volle Länge)-Antiserum vor H. influenzae Typ a oder b in dem Babyrattenmodell der Bakteriämie

Gruppe (#)	Meerschweinchen-serum	Anti-rHia-Antikörpertiter	Anzahl bakterämisch/Anzahl herausgefordert	Mittlere cfu/100 µl Blut
1	Anti-Typ a	nd	0/10*	0**
2	Anti-rHia	204.800	1/10*	0**
3	Präimmun	< 100	7/10	88
Gruppe (#)	Meerschweinchen-serum	Anti-rHia-Antikörpertiter	Anzahl bakterämisch/Anzahl herausgefordert	Mittlere cfu/2,5 µl Blut
4	Anti-MinnA	nd	0/10*	0**
5	Anti-rHia	204.800	1/10*	2**
6	Präimmun	< 100	10/10	600

[0176] Fünf Tage alte Babyratten wurden passiv s.c. mit 0,1 ml des angegebenen Meerschweinchen-Antiserums oder -Präimmunserums immunisiert. 20 Stunden später wurden die Babyratten i.p. mit entweder einem frisch kultivierten H. influenzae Typ a-Stamm ATCC 9006 (10^5 cfu, 0,1 ml) für die Gruppen # 1 bis 3; oder mit frisch kultiviertem Hib-Stamm MinnA (240 cfu, 0,1 ml) für die Gruppen # 4 bis 6 herausgefordert. Infizierte Tiere sind definiert als > 20 cfu, die aus 100 µl Blut gewonnen wurden, für die Gruppen # 1 bis 3; oder > 30 cfu, die aus 2,5 µl Blut gewonnen wurden, für die Gruppen # 4 bis 6.

* Exact-Test von Fischer. Es wurde eine statistische Signifikanz im Vergleich zu Tieren in der Gruppe 3 oder 6 gefunden ($P < 0,05$).

** Student's unpaired t-Test. Es wurde eine statistische Signifikanz im Vergleich zu Tieren in der Gruppe 3 oder 6 gefunden ($P < 0,05$).

nd: nicht bestimmt.

TABELLE 2

Schützende Wirkung von Meerschweinchen-anti-V38 rHia(SB11)-Antiserum vor H. influenzae Typ a oder b in dem Babyrattenmodell der Bakteriämie

Gruppe (#)	Meerschweinchen-serum	Anti-rHia-Antikörpertiter	Anzahl bakterämisch/Anzahl herausgefordert	Mittlere cfu/20 µl Blut
1	Anti-Typ a	nd	0/6*	0**
2	Anti-rHia	204.800	1/9*	5**
3	Präimmun	< 100	5/8	165
Gruppe (#)	Meerschweinchen-serum	Anti-rHia-Antikörpertiter	Anzahl bakterämisch/Anzahl herausgefordert	Mittlere cfu/2 µl Blut
4	Anti-MinnA	nd	0/6*	0**
5	Anti-rHia	204.800	1/9*	2**
6	Präimmun	< 100	10/10	820

[0177] Fünf Tage alte Babyratten wurden passiv s.c. mit 0,1 ml des angegebenen Meerschweinchen-Antiserums oder -Präimmunserums immunisiert. 20 Stunden später wurden die Babyratten i.p. mit entweder einem frisch kultivierten H. influenzae-Typ a-Stamm ATCC 9006 (10^5 cfu, 0,1 ml) für die Gruppen # 1 bis 3; oder mit frisch kultiviertem Hib-Stamm MinnA (190 cfu, 0,1 ml) für die Gruppen #4 bis 6 herausgefordert. Infizierte Tiere sind definiert als > 20 cfu, die aus 20 µl Blut gewonnen wurden, für die Gruppen # 1 bis 3; oder > 30 cfu, die aus 2 µl Blut gewonnen wurden, für die Gruppen #4 bis 6.

* Exact-Test von Fischer. Es wurde eine statistische Signifikanz im Vergleich zu Tieren in der Gruppe 3 oder 6 gefunden ($P < 0,05$).

** Student's unpaired t-Test. Es wurde eine statistische Signifikanz im Vergleich zu Tieren in der Gruppe 3 oder 6 gefunden ($P < 0,05$).

nd: nicht bestimmt.

SEQUENZPROTOKOLL

<110> Connaught Laboratories Limited
 Loosmore, Sheena M
 Yang, Yan-Ping
 Klein, Michel H

<120> Rekombinierte Haemophilus influenzae Adhäsion-Proteine

<130> 1038-1025 MIS:jb

<140> PCT/CA00/00289

<141> 2000-03-16

<150> 09/268,347

<151> 1999-03-16

<160> 61

<170> PatentIn Ver. 2.1

<210> 1

<211> 40

<212> DNA

<213> Haemophilus influenzae

<400> 1

gcgaattcat atgaacaaaa tttttaacgt tatttggaat

40

<210> 2

<211> 10

<212> PRT

<213> Haemophilus influenzae

<400> 2

Met Asn Lys Ile Phe Asn Val Ile Trp Asn
 1 5 10

<210> 3

<211> 56

<212> DNA

<213> Haemophilus influenzae

<400> 3

gcggatccgg aattctatta ccactggtaa ccaacacctg ctgcaacgcc tgtttt

56

<210> 4

<211> 42

<212> DNA

<213> Haemophilus influenzae

<400> 4

aaaacaggcg ttgcagcagg tggtgggttac cagtggtaat ag

42

<210> 5

<211> 12

<212> PRT

<213> Haemophilus influenzae

<400> 5

Lys Thr Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
1 5 10

<210> 6

<211> 64

<212> PRT

<213> Haemophilus influenzae

<400> 6

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp
1 5 10 15

Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Cys Ala Ser Ala
20 25 30

Thr Val Ala Val Ala Val Leu Ala Thr Leu Leu Ser Ala Thr Val Glu
35 40 45

Ala Asn Ala Asn Thr Pro Val Thr Asn Lys Leu Lys Ala Tyr Gly Asp
50 55 60

<210> 7

<211> 43

<212> DNA

<213> Haemophilus influenzae

<400> 7

gggaattcat atggaactca ctcgcaccca caccaaattgg gcc 43

<210> 8

<211> 11

<212> PRT

<213> Haemophilus influenzae

<400> 8

Met Glu Leu Thr Arg Thr His Thr Lys Cys Ala
1 5 10

<210> 9

<211> 43

<212> DNA

<213> Haemophilus influenzae

<400> 9

gggaattcat atgaccgtgg cggttgccgt attggcaacc ctg 43

<210> 10

<211> 11

<212> PRT

<213> Haemophilus influenzae

<400> 10

Met Thr Val Ala Val Ala Val Leu Ala Thr Leu
 1 5 10

<210> 11

<211> 40

<212> DNA

<213> Haemophilus influenzae

<400> 11

gggaattcat atggtattgg caaccctggt gtccgcaacg 40

<210> 12

<211> 10

<212> PRT

<213> Haemophilus influenzae

<400> 12

Met Val Leu Ala Thr Leu Leu Ser Ala Thr
 1 5 10

<210> 13

<211> 43

<212> DNA

<213> Haemophilus influenzae

<400> 13

gggaattcat atgaatactc ctgttacgaa taagttgaag gct 43

<210> 14

<211> 11

<212> PRT

<213> Haemophilus influenzae

<400> 14

Met Asn Thr Pro Val Thr Asn Lys Leu Lys Ala
 1 5 10

<210> 15

<211> 45

<212> DNA

<213> Haemophilus influenzae

<400> 15

-cgggatccca ccaaggtctt tcgctagcgc aaaggtaatg gtgtg 45

<210> 16

<211> 38

<212> DNA

<213> Haemophilus influenzae

<400> 16
cacaccatta cctttgcgct agcgaaagac cttggtgg 38

<210> 17
<211> 12
<212> PRT
<213> Haemophilus influenzae

<400> 17
His Thr Ile Thr Phe Ala Leu Ala Lys Asp Leu Gly
1 5 10

<210> 18
<211> 47
<212> DNA
<213> Haemophilus influenzae

<400> 18
cgaagcttca atttacgtaa attgcctacg gttgcggtgg ttctgtc 47

<210> 19
<211> 47
<212> DNA
<213> Haemophilus influenzae

<400> 19
gacgaaacca ccgcaaccgt aggcaattta cgtaaattga agcttcg 47

<210> 20
<211> 13
<212> PRT
<213> Haemophilus influenzae

<400> 20
Asp Glu Thr Thr Ala Thr Val Gly Asn Leu Arg Lys Leu
1 5 10

<210> 21
<211> 42
<212> DNA
<213> Haemophilus influenzae

<400> 21
ttaaatataa ggtaaataaa aatgaacaaa atttttaacg tt 42

<210> 22
<211> 7
<212> PRT
<213> Haemophilus influenzae

<400> 22

Met Asn Lys Ile Phe Asn Val

1

5

<210> 23

<211> 3036

<212> DNA

<213> Haemophilus influenzae

<400> 23

```

gaattcggct taaataaaaa tgaacaaaat ttttaacggt atttggaatg ttatgactca 60
aacttgggct gtcgtatctg aactcactcg cgccacacc aaacgtgcct ccgcaaccgt 120
ggcagccgct gtattggcga ccgtattgtc tgcaacggtt caggcgagtg caggcagtac 180
gacaggtaca aatagtttga atgtttatgg aaagaataat tcgaatttca attcagccaa 240
taattcaata gcagatttaa ataaacaaaa tgatagtgtt tacgatgggt tattaaatct 300
gaatgaaaaa ggtacggata agtcaaaatt cctgggtgct gacgaaacca ccgcaaccgt 360
aggcaattta cgtaaattgg gttgggtagt atcaaccaa aacagtacga aagaagaaag 420
caatcaagtc aaacaggcgg atgaagtgtt gtttgaaggc aaagacgggt taacgggtac 480
ttccaaatct gaaaacggca aacacaccgt tacttttgcc cttgcgaatg accttaatgt 540
aaaaaacgca accggttagcg ataaattatc gcttgggtgca aacggcaaga aagtcgatat 600
taccagtgat gcaaacggct tgaattttgc gaaacagggt acgaatgggt aaaaacggtaa 660
tgttcactta aacgggtattg cttcgacttt agatgatcct cgtgtgggtg gaaaaacagc 720
acaccttaca aaagaaatca gcgatacaga acgtaaccgt gctgcgagcg tgggcgatgt 780
attgaatgcg ggttggataa ttcgtggcgc aaaaacgatt ggcggtacag tggataatgt 840
tgattttgtt tcaacttatg acactgttga atttgccagc ggcgcaaacg caaatgtgag 900
cgttacgact gatgataaca aaaaaacaac cgtccgtgtg gatgtaacag gcttgccggt 960
ccaatatgtt acggaagaca gcaaaaccgt tgtgaaagtg ggcaatgagt attacgaagc 1020
caagcaagac ggttcggcgg atatggataa aaaagtgcga aatggcaagc tggcgaaaac 1080
taaagtgaat ttggtatcgg caaacggtag aaatccggtg aaatccagca atgttgcgga 1140
cggcacggaa tatccgatcg cggtcagctt taagcagttg aaagccttgc aagataaaca 1200
ggttacgtta agtgcgagca atgcttatgc caatggcggg agcgatgccg acggcggcaa 1260
ggcaactcaa actttaggca atgatttgaa ttttaaatat aaatccacag acagcgaggt 1320
gttgaacatc aaagcagcag gtgacacggg tacctttacg ccgaaaaaag gttcgggtgca 1380
ggttggcgat gatggtaagg ctacgattca agacggcgcg aaaacaacta ccggtttggt 1440
tgaggcttct gaattgggtg acagcctgaa caaattgggc tggaaaaggc gcgttggtta 1500
agacggcaca ggagcgacgt atggcacgca taccgacact ttagtgaagt cgggcgataa 1560
agtaactttg aaagccggcg ataatctgaa ggtcaaacaa gagggataca acttcactta 1620
cgtgctcaga gatgaattga cgggcgtaaa gagcgtggag tttaaagaca cggagaatgg 1680
tgcaaacggg gcaagcacga agattaccaa agacggcttg accattacgc cggcaaacga 1740
tgcaaatggt gcggcgcgga ctgatgtgca caagattaaa gtggcttcag acggcattag 1800
tgcggtgaat aaagcagtta aaaacggtgt gagcggactg aagaaatttg gtgatgcgaa 1860
tttcaatccg ctgactagct cagccgacaa cttaacgaaa caatatgaca atgcctataa 1920
aggcttgacc aatctggatg aaaaaagtaa aggcaagcaa actccgaccg ttgctgacaa 1980
taccgctgca accgtggcgg atttgcgcgg tttgggctgg gtcattttctg cagacaaaac 2040
cacaggcgag tcaaaggaat atagcgcgca agtgcgtaac gccaatgaag tgaattcaa 2100
gagcggcaac ggtatcaatg tttccggtaa aacattggat aacggtagcg gcgaaattac 2160
ttttgaattg gctaaagacg aaaatgccat tgctttcggg tctggctcaa aagccttgcg 2220
cgataacacg gtggcgattg gtacgggcaa cgttgtgaat gcggaaaaat ctggtgcatt 2280
cggcgatccg aactacatcg aagataaagc cgggtggcagc tacgctttcg gtaacgataa 2340
ccgtattact tctaaaaaca cttttgtgtt gggtaatgga gttaatgcca aatataaagc 2400
caatggagat gttgatacgg aaaccgtaac tggttaaggac aaagacggta aagagactac 2460
cgttactggt cctaaagcgt taggggtac. ggttgaaaac tccgtttatt tgggtaataa 2520
atcgactcgc acaaaagata agggtaaaaa tctgaaatct gatggtacgg cgggtaaacac 2580
tacaactgct ggtacaacgg gtacggtaaa cggctttgcc ggtgcaacgg cgcacgggtc 2640
ggttttctgtc ggcgcaagcg gcgaagaaag acgtatccaa aacgtttcgg caggcgaaat 2700
ttccgctact tccaccgatg cgattaacgg cagccagttg tatgccgtgg caaaaggggt 2760
aacaacacct gctggacaag tgaataaagt gggcaaacgt gcagatgcag gtacagcaag 2820
tgcattagcg gcttcacagt taccacaagc ctctatgtca ggtaaatcaa tggtttctat 2880

```

tgcgggaagt agttatcaag gtcaaagtgg tttagctatc ggggtatcaa gaatttccga 2940
 taatggcaaa gtgattattc gcttgtcagg cacaaccaat agccaaggta aaacaggcgt 3000
 tgcagcaggt gttggttacc agtggttaata gaattc 3036

<210> 24

<211> 1002

<212> PRT

<213> Haemophilus influenzae

<400> 24

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
 1 5 10 15

Ala Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Arg Ala Ser Ala
 20 25 30

Thr Val Ala Ala Ala Val Leu Ala Thr Val Leu Ser Ala Thr Val Gln
 35 40 45

Ala Ser Ala Gly Ser Thr Thr Gly Thr Asn Ser Leu Asn Val Tyr Gly
 50 55 60

Lys Asn Asn Ser Asn Phe Asn Ser Ala Asn Asn Ser Ile Ala Asp Leu
 65 70 75 80

Asn Lys Gln Asn Asp Ser Val Tyr Asp Gly Leu Leu Asn Leu Asn Glu
 85 90 95

Lys Gly Thr Asp Lys Ser Lys Phe Leu Val Ala Asp Glu Thr Thr Ala
 100 105 110

Thr Val Gly Asn Leu Arg Lys Leu Gly Trp Val Val Ser Thr Lys Asn
 115 120 125

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

Phe Glu Gly Lys Asp Gly Val Thr Val Thr Ser Lys Ser Glu Asn Gly
 145 150 155 160

Lys His Thr Val Thr Phe Ala Leu Ala Asn Asp Leu Asn Val Lys Asn
 165 170 175

Ala Thr Val Ser Asp Lys Leu Ser Leu Gly Ala Asn Gly Lys Lys Val
 180 185 190

Asp Ile Thr Ser Asp Ala Asn Gly Leu Lys Phe Ala Lys Gln Gly Thr
 195 200 205

Asn Gly Gln Asn Gly Asn Val His Leu Asn Gly Ile Ala Ser Thr Leu
 210 215 220

Asp Asp Pro Arg Val Gly Gly Lys Thr Ala His Leu Thr Lys Glu Ile
 225 230 235 240

Ser Asp Thr Glu Arg Asn Arg Ala Ala Ser Val Gly Asp Val Leu Asn
 245 250 255

Ala Gly Trp Asn Ile Arg Gly Ala Lys Thr Ile Gly Gly Thr Val Asp
 260 265 270
 Asn Val Asp Phe Val Ser Thr Tyr Asp Thr Val Glu Phe Ala Ser Gly
 275 280 285
 Ala Asn Ala Asn Val Ser Val Thr Thr Asp Asp Asn Lys Lys Thr Thr
 290 295 300
 Val Arg Val Asp Val Thr Gly Leu Pro Val Gln Tyr Val Thr Glu Asp
 305 310 315 320
 Ser Lys Thr Val Val Lys Val Gly Asn Glu Tyr Tyr Glu Ala Lys Gln
 325 330 335
 Asp Gly Ser Ala Asp Met Asp Lys Lys Val Glu Asn Gly Lys Leu Ala
 340 345 350
 Lys Thr Lys Val Lys Leu Val Ser Ala Asn Gly Thr Asn Pro Val Lys
 355 360 365
 Ile Ser Asn Val Ala Asp Gly Thr Glu Asp Thr Asp Ala Val Ser Phe
 370 375 380
 Lys Gln Leu Lys Ala Leu Gln Asp Lys Gln Val Thr Leu Ser Ala Ser
 385 390 395 400
 Asn Ala Tyr Ala Asn Gly Gly Ser Asp Ala Asp Gly Gly Lys Ala Thr
 405 410 415
 Gln Thr Leu Gly Asn Asp Leu Asn Phe Lys Phe Lys Ser Thr Asp Ser
 420 425 430
 Glu Leu Leu Asn Ile Lys Ala Ala Gly Asp Thr Val Thr Phe Thr Pro
 435 440 445
 Lys Lys Gly Ser Val Gln Val Gly Asp Asp Gly Lys Ala Thr Ile Gln
 450 455 460
 Asp Gly Ala Lys Thr Thr Thr Gly Leu Val Glu Ala Ser Glu Leu Val
 465 470 475 480
 Asp Ser Leu Asn Lys Leu Gly Trp Lys Val Gly Val Gly Lys Asp Gly
 485 490 495
 Thr Gly Ala Thr Asp Gly Thr His Thr Asp Thr Leu Val Lys Ser Gly
 500 505 510
 Asp Lys Val Thr Leu Lys Ala Gly Asp Asn Leu Lys Val Lys Gln Glu
 515 520 525
 Gly Thr Asn Phe Thr Tyr Val Leu Arg Asp Glu Leu Thr Gly Val Lys
 530 535 540
 Ser Val Glu Phe Lys Asp Thr Glu Asn Gly Ala Asn Gly Ala Ser Thr
 545 550 555 560
 Lys Ile Thr Lys Asp Gly Leu Thr Ile Thr Pro Ala Asn Asp Ala Asn
 565 570 575

Gly Ala Ala Thr Asp Ala Asp Lys Ile Lys Val Ala Ser Asp Gly
 580 585 590
 Ile Ser Ala Gly Asn Lys Ala Val Lys Asn Val Val Ser Gly Leu Lys
 595 600 605
 Lys Phe Gly Asp Ala Asn Phe Asn Pro Leu Thr Ser Ser Ala Asp Asn
 610 615 620
 Leu Thr Lys Gln Tyr Asp Asn Ala Tyr Lys Gly Leu Thr Asn Leu Asp
 625 630 635 640
 Glu Lys Ser Lys Gly Lys Gln Thr Pro Thr Val Ala Asp Asn Thr Ala
 645 650 655
 Ala Thr Val Gly Asp Leu Arg Gly Leu Gly Trp Val Ile Ser Ala Asp
 660 665 670
 Lys Thr Thr Gly Glu Ser Lys Glu Tyr Ser Ala Gln Val Arg Asn Ala
 675 680 685
 Asn Glu Val Lys Phe Lys Ser Gly Asn Gly Ile Asn Val Ser Gly Lys
 690 695 700
 Thr Leu Asp Asn Gly Thr Arg Glu Ile Thr Phe Glu Leu Ala Lys Asp
 705 710 715 720
 Glu Asn Ala Ile Ala Phe Gly Ser Gly Ser Lys Ala Leu Arg Asp Asn
 725 730 735
 Thr Val Ala Ile Gly Thr Gly Asn Val Val Asn Ala Glu Lys Ser Gly
 740 745 750
 Ala Phe Gly Asp Pro Asn Tyr Ile Glu Asp Lys Ala Gly Gly Ser Tyr
 755 760 765
 Ala Phe Gly Asn Asp Asn Arg Ile Thr Ser Lys Asn Thr Phe Val Leu
 770 775 780
 Gly Asn Gly Val Asn Ala Lys Tyr Lys Ala Asn Gly Asp Val Asp Thr
 785 790 795 800
 Glu Thr Val Thr Val Lys Asp Lys Asp Gly Lys Glu Thr Thr Val Thr
 805 810 815
 Val Pro Lys Ala Leu Gly Ala Thr Val Glu Asn Ser Val Tyr Leu Gly
 820 825 830
 Asn Lys Ser Thr Ala Thr Lys Asp Lys Gly Lys Asn Leu Lys Ser Asp
 835 840 845
 Gly Thr Ala Gly Asn Thr Thr Thr Ala Gly Thr Thr Gly Thr Val Asn
 850 855 860
 Gly Phe Ala Gly Ala Thr Ala His Gly Ala Val Ser Val Gly Ala Ser
 865 870 875 880

Gly Glu Glu Arg Arg Ile Gln Asn Val Ala Ala Gly Glu Ile Ser Ala
885 890 895

Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln Leu Tyr Ala Val Ala Lys
900 905 910

Gly Val Thr Asn Leu Ala Gly Gln Val Asn Lys Val Gly Lys Arg Ala
915 920 925

Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala
930 935 940

Ser Met Ser Gly Lys Ser Met Val Ser Ile Ala Gly Ser Ser Tyr Gln
945 950 955 960

Gly Gln Ser Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly
965 970 975

Lys Val Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr
980 985 990

Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
995 1000

<210> 25

<211> 2079

<212> DNA

<213> Haemophilus influenzae

<400> 25

```

gaattcggct ttaaatataa ggtaaataaa aatgaacaaa atttttaacg ttatttggaa 60
tggtgtgact caaacttggg ttgtcgtatc tgaactcact cgcacccaca ccaaattggc 120
ctccgccacc gtggcagttg ccgtattggc aaccctgttg tccgcaacgg ttcaggcgaa 180
tgctaccgat gaaaacgaag atgatgaaga agagttagaa cccgtacaac gctctgtttt 240
aagggtggagc ttcaaattccg ctaagggaagg cactggagaa caagagggaa caacagaggt 300
aataaatttg aacacagatt catcaggaaa tgcagtagga agcagcacia tcaccttcaa 360
agccggcgac aacctgaaaa tcaaacaaag cggcaatgac ttcacctact cgctgaaaaa 420
agagctgaaa aacctgacca gtgttgaac tgaaaaatta tcgtttggcg caaacggcaa 480
taaagttgat attaccagt atgcaaatgg cttgaaattg gcgaaaacag gtaacggaaa 540
tggtcaaaac agtaattgtc acttaaacgg tattgcttcg actttgaccg atacgcttgc 600
cggtggcaca acaggacacg ttgacaccaa cattgatgag gttaattatc atcgcgctgc 660
aagcgtacaa gatgtgttaa acagcggttg gaatatccaa ggcaatggaa acaatgtcga 720
ttttgtccgt acttacgaca ccgtggactt tgtcaatggc gcgaatgcca atgtgagcgt 780
tacggctgat acggctcaca aaaagacaac tgtccgtgtg gatgtaacag gcttgccggt 840
tcaatatgtt acggaagacg gcaaaaccgt tgtgaaagtg ggcaatgagt attacaaagc 900
caaagatgac ggttcggcgg atatgaatca aaaagtcgaa aacggcgagc tggcgaaaaa 960
caaagtgaag ttggtatcgg caagcggtag aaatccggtg aaaattagca atgttgaga 1020
cggcacggaa gacaccgatg cggtcagctt taagcaatta aaagccttgc aagacaaaca 1080
ggttacgttg agcagagca atgcttatgc caatggcggg acagataacg acggcggcaa 1140
ggcaactcaa actttaagca atggtttgaa ttttaaattt aaatctagcg atggcgagtt 1200
gttgaaaatt agcgcgaccg gcgatacggg tacttttacg ccgaaaaaag gttcgggtaca 1260
ggttggcgat gatggcaagg cttcaatttc aaaagggtgca aatacaactg aaggtttggt 1320
tgaggcttct gaattggttg aaagcctgaa caaactgggt tggaaagtag gggttgagaa 1380
agtcggcagc ggcgagcttg atggtacatc caaggaaact ttagtgaaat cgggcgataa 1440
agtaactttg aaagccggcg acaatctgaa ggtcaaacia gagggcacia acttcaacta 1500
cgcgctcaaa gatgaattga cgggcgtgaa gagcgtggag tttaaagaca cggcgaatgg 1560
tgcaaacggg gcaagcacga agattaccaa agacggcttg accattacgc tggcaaacgg 1620
tgcaaatggg gcgacgggtg atgatgccga caagattaaa gttgcttcgg acggcattag 1680

```

```

cgcgggtaat aaagcagtta aaaacgtcgc ggcaggcgaa atttctgcc cttccaccga 1740
tgcgattaac ggaagccagt tgtatgccgt ggcaaaaggg gtaacaaacc ttgctggaca 1800
agtgaataat cttgagggca aagtgaataa agtgggcaaa cgtgcagatg caggtactgc 1860
aagtgcatta gcggcttcac agttaccaca agccactatg ccaggtaa at caatgggttc 1920
tattgcggga agtagttatc aagggtcaaaa tggtttagct atcgggggtat caagaatttc 1980
cgataatggc aaagtgatta ttcgcttgtc aggcacaacc aatagtcaag gtaaaacagg 2040
cgttgcagca ggtgttggtt accagtggtg atagaattc 2079

```

<210> 26

<211> 679

<212> PRT

<213> Haemophilus influenzae

<400> 26

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp
 1             5             10             15

Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Cys Ala Ser Ala
      20             25             30

Thr Val Ala Val Ala Val Leu Ala Thr Leu Leu Ser Ala Thr Val Gln
      35             40             45

Ala Asn Ala Thr Asp Glu Asn Glu Asp Asp Glu Glu Glu Leu Glu Pro
      50             55             60

Val Gln Arg Ser Val Leu Arg Trp Ser Phe Lys Ser Ala Lys Glu Gly
      65             70             75             80

Thr Gly Glu Gln Glu Gly Thr Thr Glu Val Ile Asn Leu Asn Thr Asp
      85             90             95

Ser Ser Gly Asn Ala Val Gly Ser Ser Thr Ile Thr Phe Lys Ala Gly
      100            105            110

Asp Asn Leu Lys Ile Lys Gln Ser Gly Asn Asp Phe Thr Tyr Ser Leu
      115            120            125

Lys Lys Glu Leu Lys Asn Leu Thr Ser Val Glu Thr Glu Lys Leu Ser
      130            135            140

Phe Gly Ala Asn Gly Asn Lys Val Asp Ile Thr Ser Asp Ala Asn Gly
      145            150            155            160

Leu Lys Leu Ala Lys Thr Gly Asn Gly Asn Gly Gln Asn Ser Asn Val
      165            170            175

His Leu Asn Gly Ile Ala Ser Thr Leu Thr Asp Thr Leu Ala Gly Gly
      180            185            190

Thr Thr Gly His Val Asp Thr Asn Ile Asp Ala Val Asn Tyr His Arg
      195            200            205

Ala Ala Ser Val Gln Asp Val Leu Asn Ser Gly Trp Asn Ile Gln Gly
      210            215            220

Asn Gly Asn Asn Val Asp Phe Val Arg Thr Tyr Asp Thr Val Asp Phe
      225            230            235            240

```

Val Asn Gly Ala Asn Ala Asn Val Ser Val Thr Ala Asp Thr Ala His
 245 250 255
 Lys Lys Thr Thr Val Arg Val Asp Val Thr Gly Leu Pro Val Gln Tyr
 260 265 270
 Val Thr Glu Asp Gly Lys Thr Val Val Lys Val Gly Asn Glu Tyr Tyr
 275 280 285
 Lys Ala Lys Asp Asp Gly Ser Ala Asp Met Asn Gln Lys Val Glu Asn
 290 295 300
 Gly Glu Leu Ala Lys Thr Lys Val Lys Leu Val Ser Ala Ser Gly Thr
 305 310 315 320
 Asn Pro Val Lys Ile Ser Asn Val Ala Asp Gly Thr Glu Asp Thr Asp
 325 330 335
 Ala Val Ser Phe Lys Gln Leu Lys Ala Leu Gln Asp Lys Gln Val Thr
 340 345 350
 Leu Ser Thr Ser Asn Ala Tyr Ala Asn Gly Gly Thr Asp Asn Asp Gly
 355 360 365
 Gly Lys Ala Thr Gln Thr Leu Ser Asn Gly Leu Asn Phe Lys Phe Lys
 370 375 380
 Ser Ser Asp Gly Glu Leu Leu Lys Ile Ser Ala Thr Gly Asp Thr Val
 385 390 395 400
 Thr Phe Thr Pro Lys Lys Gly Ser Val Gln Val Gly Asp Asp Gly Lys
 405 410 415
 Ala Ser Ile Ser Lys Gly Ala Asn Thr Thr Glu Gly Leu Val Glu Ala
 420 425 430
 Ser Glu Leu Val Glu Ser Leu Asn Lys Leu Gly Trp Lys Val Gly Val
 435 440 445
 Glu Lys Val Gly Ser Gly Glu Leu Asp Gly Thr Ser Lys Glu Thr Leu
 450 455 460
 Val Lys Ser Gly Asp Lys Val Thr Leu Lys Ala Gly Asp Asn Leu Lys
 465 470 475 480
 Val Lys Gln Glu Gly Thr Asn Phe Thr Tyr Ala Leu Lys Asp Glu Leu
 485 490 495
 Thr Gly Val Lys Ser Val Glu Phe Lys Asp Thr Ala Asn Gly Ala Asn
 500 505 510
 Gly Ala Ser Thr Lys Ile Thr Lys Asp Gly Leu Thr Ile Thr Leu Ala
 515 520 525
 Asn Gly Ala Asn Gly Ala Thr Val Thr Asp Ala Asp Lys Ile Lys Val
 530 535 540

Ala Ser Asp Gly Ile Ser Ala Gly Asn Lys Ala Val Lys Asn Val Ala
545 550 555 560

Ala Gly Glu Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln
565 570 575

Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly Gln Val Asn
580 585 590

Asn Leu Glu Gly Lys Val Asn Lys Val Gly Lys Arg Ala Asp Ala Gly
595 600 605

Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala Thr Met Pro
610 615 620

Gly Lys Ser Met Val Ser Ile Ala Gly Ser Ser Tyr Gln Gly Gln Asn
625 630 635 640

Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly Lys Val Ile
645 650 655

Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr Gly Val Ala
660 665 670

Ala Gly Val Gly Tyr Gln Trp
675

<210> 27

<211> 3353

<212> DNA

<213> Haemophilus influenzae

<400> 27

ttaaataataa ggtaaataaaa aatgaacaaa atttttaacg ttatttgga tgttgtgact 60
caaacttggg ttgtcgtatc tgaactcact cgcgcccaca ccaaatgcgc ctccgccacc 120
gtggcggttg cgtatttggc aactgcgttg tctgcaacgg ctgaagcgaa caacaatact 180
tctgttacga atgggttgaa tgcctatggc gatactaatt ttaatacaac caataattcg 240
atagcagatt tggaaaaaca cgttcaagat gcttataaag gcttattaaa tctgaatgaa 300
aaagatacaa ataagtcaag tttcttgggt gccgacaata ccgccgcaac cgtaggcaat 360
ttgcgtaaat tgggctgggt attgtctagc aaaaacggca caaggaacga gaaaagctat 420
caagtaaaac aagctgatga agttctcttt actggatctg gtgctgcaac ggtagttcc 480
agctctaaag acggtaaaca taccattacc atttctgtta ccaaaggtag ttttgctgag 540
gtaaaaactg atgcaactac tggaggtcaa gttaacgccg accgtggtaa agtgaaagct 600
gaggacgaga atggagctga tgttgataag aaagttgcaa ctgtaaaaga tgttgctaag 660
gcgattaacg atgccgcaac tttcgtgaaa gtggaaagca cagatgatga cattgaaaat 720
ggtgctgcag gcaaaaatga aactacagac caagctctca aagcaggcga caccctaacc 780
ttaaagcgg gtaaaaactt aaaagctaag ttagaccaaa atggtaaatc agtaaccttt 840
gcttttagcga aagaccttga tgtgacctct gcgaaagtga gtgataagtt gtctatttgt 900
aaagatacga ataaagttga tattaccagt gatgcaaag gcttgaaatt ggcgaaaaca 960
ggtaacggaa atggtcaaaa cggtaatgtc cacttaaatg gtattgtctc gactttgacc 1020
gataccatta caggtatgac aacacaagca agcaatggcg tggctgtgca gaatcataat 1080
cgtgctgcga gtgtggctga tgtattaaat gcaggctgga atattcaagg caacggagcg 1140
agcgttgatt ttgtcaatgc ttacgacaca gtagattttg tcaatggtag aaacaccaat 1200
gtgaacgtta cgactgatac ggctcacaaa aagacaaccg tccgtgtgga tgaacaggc 1260
ttgccggttc aatatgttac ggaagacggc aaaaccgttg tgaagtgga caataagtat 1320
tacgaagcta agcaagacgg ttccggcgat atggataaaa aagtcgaaaa tggcgagctg 1380
gcgaaaacca aagtgaattt ggtgtcggca agcggtcaaa atccggtgaa aatcagcaat 1440
gttgcggaag gcacggaaga aaacgatgcg gtcagcttta agcaattgaa agccttgcaa 1500

```

gagaaacagg ttactttaac tgcgagcaat gcttatgcc aatgggtggtaa cgatgccgac 1560
ggcggaacagg caactcaaac tttaacaat gggttgaatt ttaaatttaa atccacagac 1620
ggcgagttgt tgaacatcaa agtagaaaat gacacagtta cctttacgcc gaaaaaagg 1680
tcggtacagg ttggcgaaga cggtaaggct acgattcaaa atggtagcaa aacaaccgac 1740
gggttgggtt aagcttccga attggttgaa agcctgaaca aactgggctg gaaagtggg 1800
gttgataaag acggcagcgg cgagcttgat ggtgcatcca atgaaacttt agtgaagtcg 1860
ggcgataaag taactttgaa agccggcgag aatctgaagg tcaacaaga cggcacaac 1920
ttcacttacg cgctcaaaga tgaattgacg ggcgtgaaga gcgtggagtt taaagacacg 1980
gcgaatggtt caaacgggtg aagcacgaag attaccaaag acggcttgac cattacgtcg 2040
gcaaacggtg cgaatggtg ggccggcgact gatgcggaca agattaaagt ggcttcagac 2100
ggcatcagtg cgggtaataa agcgggttaa aacgttgtga gcggactgaa gaaatttggt 2160
gatgcgaatt tcaatccact gaccagttcc gccgacaact taacgaaaca atatgacgat 2220
gcctataaag gcttgacca tttggatgaa aaagggtgcg acaagcaaac tctgactgtt 2280
gccgacaata ctgccgaac cgtgggcgat ttgcgcggct tgggctgggt catttctgcg 2340
gacaaaacca caggcgaact caataaggaa tacaacgcgc aagtgcgtaa cgccaatgaa 2400
gtgaaattca agagcggcaa cggtatccat gtttccggta aaacggtcaa cggtaggcgc 2460
gaaattactt ttgaattggc taaagacgaa aatgccattg ctttcggtta tggctcaaaa 2520
gccttgcgcg ataacacggt ggcaattggt acgggcaacg ttgtgaatgc ggaaaaatct 2580
gggtgcattcg gcgatccgaa ctacatcgaa gataaagccg gtggcagcta cgctttcgg 2640
aacgataacc gtattacttc taaaaacact tttgtgttg gtaatggagt taatgcgaaa 2700
tataaagcca atggagatgt tgatacggaa accgtaaccg ttaaggacaa agacggtaaa 2760
gagactaccg ttactgttcc taaagcgtta ggggctacgg ttgaaaactc cgtttatttg 2820
ggtaataaat cgactgcgac aaaagataag ggtaaaaacc tgaaatctga tggtagcgcg 2880
ggtaacacta caactgctgg cacaacgggt acggtaaacg gctttgcgg tgcaacggcg 2940
cacgggtgcgg tttctgtcgg cgcaacggc gaagaaagac gtatccaaa cgtcgcggca 3000
ggcgaaattt ccgccacttc caccgatgcg attaacggca gccagttgta tgctgtggca 3060
aaaggggtaa caaatcttgc tggacaagtg aataaagtgg gcaaacgtgc agatgcaggt 3120
acagcaagtg cattagcagc ttcacagtta ccacaagcct ctatgccagg taaatcaatg 3180
gtttctattg cgggaagtag ttatcaaggt caaatgggt tagctatcgg ggtatcacga 3240
atttcgata atggcaagt gattattcgc ttgtcaggca caaccaatag ccaaggtaaa 3300
acaggcgttg cagcaggtgt tggttaccag tggtaataga attccggatc cgc 3353

```

<210> 28

<211> 1104

<212> PRT

<213> Haemophilus influenzae

<400> 28

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp
1 5 10 15

Val Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Cys Ala Ser Ala
20 25 30

Thr Val Ala Val Ala Val Leu Ala Thr Ala Leu Ser Ala Thr Ala Glu
35 40 45

Ala Asn Asn Asn Thr Ser Val Thr Asn Gly Leu Asn Ala Tyr Gly Asp
50 55 60

Thr Asn Phe Asn Thr Thr Asn Asn Ser Ile Ala Asp Leu Glu Lys His
65 70 75 80

Val Gln Asp Ala Tyr Lys Gly Leu Leu Asn Leu Asn Glu Lys Asp Thr
85 90 95

Asn Lys Ser Ser Phe Leu Val Ala Asp Asn Thr Ala Ala Thr Val Gly
100 105 110

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

Gln Tyr Val Thr Glu Asp Gly Lys Thr Val Val Lys Val Asp Asn Lys
 420 425 430
 Tyr Tyr Glu Ala Lys Gln Asp Gly Ser Ala Asp Met Asp Lys Lys Val
 435 440 445
 Glu Asn Gly Glu Leu Ala Lys Thr Lys Val Lys Leu Val Ser Ala Ser
 450 455 460
 Gly Gln Asn Pro Val Lys Ile Ser Asn Val Ala Glu Gly Thr Glu Glu
 465 470 475 480
 Asn Asp Ala Val Ser Phe Lys Gln Leu Lys Ala Leu Gln Glu Lys Gln
 485 490 495
 Val Thr Leu Thr Ala Ser Asn Ala Tyr Ala Asn Gly Gly Asn Asp Ala
 500 505 510
 Asp Gly Gly Lys Ala Thr Gln Thr Leu Asn Asn Gly Leu Asn Phe Lys
 515 520 525
 Phe Lys Ser Thr Asp Gly Glu Leu Leu Asn Ile Lys Val Glu Asn Asp
 530 535 540
 Thr Val Thr Phe Thr Pro Lys Lys Gly Ser Val Gln Val Gly Glu Asp
 545 550 555 560
 Gly Lys Ala Thr Ile Gln Asn Gly Thr Lys Thr Thr Asp Gly Leu Val
 565 570 575
 Glu Ala Ser Glu Leu Val Glu Ser Leu Asn Lys Leu Gly Trp Lys Val
 580 585 590
 Gly Val Asp Lys Asp Gly Ser Gly Glu Leu Asp Gly Ala Ser Asn Glu
 595 600 605
 Thr Leu Val Lys Ser Gly Asp Lys Val Thr Leu Lys Ala Gly Glu Asn
 610 615 620
 Leu Lys Val Lys Gln Asp Gly Thr Asn Phe Thr Tyr Ala Leu Lys Asp
 625 630 635 640
 Glu Leu Thr Gly Val Lys Ser Val Glu Phe Lys Asp Thr Ala Asn Gly
 645 650 655
 Ser Asn Gly Ala Ser Thr Lys Ile Thr Lys Asp Gly Leu Thr Ile Thr
 660 665 670
 Ser Ala Asn Gly Ala Asn Gly Ala Ala Ala Thr Asp Ala Asp Lys Ile
 675 680 685
 Lys Val Ala Ser Asp Gly Ile Ser Ala Gly Asn Lys Ala Val Lys Asn
 690 695 700
 Val Val Ser Gly Leu Lys Lys Phe Gly Asp Ala Asn Phe Asn Pro Leu
 705 710 715 720
 Thr Ser Ser Ala Asp Asn Leu Thr Lys Gln Tyr Asp Asp Ala Tyr Lys
 725 730 735

Gly Leu Thr Asn Leu Asp Glu Lys Gly Ala Asp Lys Gln Thr Leu Thr
 740 745 750
 Val Ala Asp Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly
 755 760 765
 Trp Val Ile Ser Ala Asp Lys Thr Thr Gly Glu Leu Asn Lys Glu Tyr
 770 775 780
 Asn Ala Gln Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn
 785 790 795 800
 Gly Ile His Val Ser Gly Lys Thr Val Asn Gly Arg Arg Glu Ile Thr
 805 810 815
 Phe Glu Leu Ala Lys Asp Glu Asn Ala Ile Ala Phe Gly Tyr Gly Ser
 820 825 830
 Lys Ala Leu Arg Asp Asn Thr Val Ala Ile Gly Thr Gly Asn Val Val
 835 840 845
 Asn Ala Glu Lys Ser Gly Ala Phe Gly Asp Pro Asn Tyr Ile Glu Asp
 850 855 860
 Lys Ala Gly Gly Ser Tyr Ala Phe Gly Asn Asp Asn Arg Ile Thr Ser
 865 870 875 880
 Lys Asn Thr Phe Val Leu Gly Asn Gly Val Asn Ala Lys Tyr Lys Ala
 885 890 895
 Asn Gly Asp Val Asp Thr Glu Thr Val Thr Val Lys Asp Lys Asp Gly
 900 905 910
 Lys Glu Thr Thr Val Thr Val Pro Lys Ala Leu Gly Ala Thr Val Glu
 915 920 925
 Asn Ser Val Tyr Leu Gly Asn Lys Ser Thr Ala Thr Lys Asp Lys Gly
 930 935 940
 Lys Asn Leu Lys Ser Asp Gly Thr Ala Gly Asn Thr Thr Thr Ala Gly
 945 950 955 960
 Thr Thr Gly Thr Val Asn Gly Phe Ala Gly Ala Thr Ala His Gly Ala
 965 970 975
 Val Ser Val Gly Ala Ser Gly Glu Glu Arg Arg Ile Gln Asn Val Ala
 980 985 990
 Ala Gly Glu Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln
 995 1000 1005
 Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly Gln Val Asn
 1010 1015 1020
 Lys Val Gly Lys Arg Ala Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala
 1025 1030 1035 1040

Ser Gln Leu Pro Gln Ala Ser Met Pro Gly Lys Ser Met Val Ser Ile
 1045 1050 1055

Ala Gly Ser Ser Tyr Gln Gly Gln Asn Gly Leu Ala Ile Gly Val Ser
 1060 1065 1070

Arg Ile Ser Asp Asn Gly Lys Val Ile Ile Arg Leu Ser Gly Thr Thr
 1075 1080 1085

Asn Ser Gln Gly Lys Thr Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
 1090 1095 1100

<210> 29

<211> 3030

<212> DNA

<213> Haemophilus influenzae

<400> 29

```

gcgaattcat atgaacaaaa tttttaacgt tatttggaaat gttatgactc aaacttgggc 60
tgtcgtatct gaactcactc gcgcccacac caaacgtgcc tccgcaaccg tggcaaccgc 120
cgtattggcg acgttggtgt ctacaacagt tcaggcgaca actactggcg gtacgacaag 180
tacaaacggg ttgaaagctt atggaagtac gaataatccg aatttcaatg ctgcaggtaa 240
ctctgcaact gatttagcta gacagtgtga tgggtgcttat gacggtttat taaatctaaa 300
tgaaaaagat gcgaataaaa atctgttggg gactgatgat aaggcggcga ccgtaggcaa 360
tttgcgtaaa ttgggttggg tattgtctag taaaaacggc acaagggaacg agaaaagcca 420
acaagtcaaa cacgcggatg aagtgttgtt tgaaggcaaa gacggtgtaa cggttacttc 480
caaatctgaa aacggtaaac acaccgttac ttttaccctt gagaaagacc ttaatgtaaa 540
aaacgcaacc gttagcgata aattatcgct tgggtgcaaac ggcaataaag tcgatattac 600
cagtgataca aacggttga aatttgcgaa accaagtacg aatggtcaaa acggtaatgt 660
tcacttaaac ggtattgcct ctaccttaac tgacacaatt acaggtacaa caaatctgc 720
aactaatggt gtagatgtgc agaatacataa tcgtgtctgc agtgtagctg atgtattgaa 780
tgagggtgg aatattcaag gcaacggagc gagcgttgat tttgtcaata cttacgacac 840
agtagatttt gtcaatgggt taaataccaa tgtgaacgtt acgactgata cggctcaca 900
caaaaagaca accgtccgtg tggatgtaac gggcttgccg gtccaatatg ttacggaaga 960
cggcgaaacc gttgtgaaag tgggcaatga gtattacgaa gccaaagcaag acggttcggc 1020
ggatattgat aaaaaagtcg aaaatggcaa gctggcgaaa actaaagtta aattggtatc 1080
ggcaaacggg acaaatccgg tgaaaatcag caatgttgcg gacggcacgg aaaataccga 1140
tgcggtcagc ttttaagcag tgaaagcctt gcaagacaaa caggttacgt taagtgcgag 1200
caatgcttat gccaatggcg gtagcgatgc cgacggcggc aagggaattc aaactttaag 1260
caatgggttg aattttaaat ttaaatccac agacggcgag ttgttgaata tcaaagcaga 1320
aatgacacg gttaccttta cgccgaaaaa aggttcggtg caggttggcg atgatggtaa 1380
ggctacgatt caagacggcg caaaaacaac taccggtttg gttgaggctt ctgaattggt 1440
tgacagcctg aacaaattgg gttggaaagt gggcaccggc actgacggca caggagtgc 1500
cgatggcacg cataccgaca ctttagtgaa gtcggcgat aaagtaactt tgaaagccgg 1560
cgacaatctg aaggtcaaac aagagggtag aaacttcaat tatgcgctca aagatgaatt 1620
gacggacgtg aagagcgtgg agtttaaaag cacggcgaat ggtgcaaacg gtgcaagcac 1680
gaagattacc aaagacggct tgaccattac gccggcaaac ggtgcgggtg cggcagggtc 1740
aaacactgca aacaccatta gcgtaacca agacggcatt agcgcgggta ataaagcagt 1800
taaaaacggt gtgagcggac tgaagaaatt tgggtgatgc aatttcgatc cgctgactag 1860
ctcagccgac aacttaacga aacaatatga caatgcctat aaaggcttga ccaatctgga 1920
tgaaaaaagt aaaggcaagc aaactccgac cgttgcgtgc aataccgctg caaccgtggg 1980
cgatttgcg ggcttgggct gggtcatttc tgcagacaaa accaaaggcg aactcaataa 2040
ggaatacaac gcacaagtgc gtaacgctaa tgaagtgaat ttcaagagcg gcaacgggat 2100
caatgtttcc ggtaaaacat tggataacgg tacgcgcgaa attacttttg aattggctaa 2160

```

```

agacgaaaat gccattgctt tccggttctgg ctcaaaaagcc ttgcgcgata acacgggtggc 2220
aattgggtacg ggcaacgttg tgaatgcgga aaaatctggt gcattcggcg atccgaacta 2280
catcgaagat aaagccgggtg gcagctacgc ttccggtaac gataaccgta ttacttctaa 2340
aaacactttt gtgttgggta atagtgttaa tgcgaaacgt gatgcaaag gcaatgtact 2400
gaccgaagaa aaagaagtgg ttggaaaaga cgggtgcgaag acgaaagtaa ccgtgccgca 2460
agccttaggc gaaaccgtag aaaattctgt ttatctcggg aatgcttcaa ctgcgacaaa 2520
agataagggt aaaaaacctga aatctgatgg tacggcgggt aacactacaa ctgctggcgc 2580
aacgggtacg gtaaacggct ttgccgggtgc aacggcgcac ggtgcgggtt ctgtcggcgc 2640
aagtggcgaa gaaagacgta tccaaaacgt cgcggcaggc gaaatttccg ctacttccac 2700
agatgcgatt aacggtagcc agttgtatgc tgtggcaaaa ggggtaacaa accttgctgg 2760
acaagtgaat aaagtgggca aacgtgcaga tgcaggtaca gcaagtgcac tagcggcttc 2820
acagttacca caagcctcta tgcaggttaa atcaatggtt tctattgcgg gaagtagtta 2880
tcaagggtcaa agtgggttag ctatcggggg atcaagaatt tccgataatg gcaaagtgat 2940
tattcgttg tcaggcacia ccaatagcca aggtaaaaca ggcgttcgag cagggtgttg 3000
ttaccagtgg taatagaatt ccgatccgc 3030

```

<210> 30

<211> 1004

<212> PRT

<213> Haemophilus influenzae

<400> 30

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
 1             5             10             15

Ala Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Arg Ala Ser Ala
      20             25             30

Thr Val Ala Thr Ala Val Leu Ala Thr Leu Leu Ser Thr Thr Val Gln
      35             40             45

Ala Thr Thr Thr Gly Gly Thr Thr Ser Thr Asn Gly Leu Lys Ala Tyr
      50             55             60

Gly Ser Thr Asn Asn Pro Asn Phe Asn Ala Ala Gly Asn Ser Ala Thr
      65             70             75             80

Asp Leu Ala Arg Gln Phe Asp Gly Ala Tyr Asp Gly Leu Leu Asn Leu
      85             90             95

Asn Glu Lys Asp Ala Asn Lys Asn Leu Leu Val Thr Asp Asp Lys Ala
      100            105            110

Ala Thr Val Gly Asn Leu Arg Lys Leu Gly Trp Val Leu Ser Ser Lys
      115            120            125

Asn Gly Thr Arg Asn Glu Lys Ser Gln Gln Val Lys His Ala Asp Glu
      130            135            140

Val Leu Phe Glu Gly Lys Asp Gly Val Thr Val Thr Ser Lys Ser Glu
      145            150            155            160

Asn Gly Lys His Thr Val Thr Phe Thr Leu Glu Lys Asp Leu Asn Val
      165            170            175

Lys Asn Ala Thr Val Ser Asp Lys Leu Ser Leu Gly Ala Asn Gly Asn
      180            185            190

```

Lys Val Asp Ile Thr Ser Asp Thr Asn Gly Leu Lys Phe Ala Lys Pro
 195 200 205
 Ser Thr Asn Gly Gln Asn Gly Asn Val His Leu Asn Gly Ile Ala Ser
 210 215 220
 Thr Leu Thr Asp Thr Ile Thr Gly Thr Thr Lys Ser Ala Thr Asn Gly
 225 230 235 240
 Val Asp Val Gln Asn His Asn Arg Ala Ala Ser Val Ala Asp Val Leu
 245 250 255
 Asn Ala Gly Trp Asn Ile Gln Gly Asn Gly Ala Ser Val Asp Phe Val
 260 265 270
 Asn Thr Tyr Asp Thr Val Asp Phe Val Asn Gly Leu Asn Thr Asn Val
 275 280 285
 Asn Val Thr Thr Asp Thr Ala His Asn Lys Lys Thr Thr Val Arg Val
 290 295 300
 Asp Val Thr Gly Leu Pro Val Gln Tyr Val Thr Glu Asp Gly Glu Thr
 305 310 315 320
 Val Val Lys Val Gly Asn Glu Tyr Tyr Glu Ala Lys Gln Asp Gly Ser
 325 330 335
 Ala Asp Met Asp Lys Lys Val Glu Asn Gly Lys Leu Ala Lys Thr Lys
 340 345 350
 Val Lys Leu Val Ser Ala Asn Gly Thr Asn Pro Val Lys Ile Ser Asn
 355 360 365
 Val Ala Asp Gly Thr Glu Asn Thr Asp Ala Val Ser Phe Lys Gln Leu
 370 375 380
 Lys Ala Leu Gln Asp Lys Gln Val Thr Leu Ser Ala Ser Asn Ala Tyr
 385 390 395 400
 Ala Asn Gly Gly Ser Asp Ala Asp Gly Gly Lys Gly Ile Gln Thr Leu
 405 410 415
 Ser Asn Gly Leu Asn Phe Lys Phe Lys Ser Thr Asp Gly Glu Leu Leu
 420 425 430
 Asn Ile Lys Ala Glu Asn Asp Thr Val Thr Phe Thr Pro Lys Lys Gly
 435 440 445
 Ser Val Gln Val Gly Asp Asp Gly Lys Ala Thr Ile Gln Asp Gly Ala
 450 455 460
 Lys Thr Thr Thr Gly Leu Val Glu Ala Ser Glu Leu Val Asp Ser Leu
 465 470 475 480
 Asn Lys Leu Gly Trp Lys Val Gly Thr Gly Thr Asp Gly Thr Gly Val
 485 490 495
 Thr Asp Gly Thr His Thr Asp Thr Leu Val Lys Ser Gly Asp Lys Val
 500 505 510

Thr Leu Lys Ala Gly Asp Asn Leu Lys Val Lys Gln Glu Gly Thr Asn
 515 520 525
 Phe Thr Tyr Ala Leu Lys Asp Glu Leu Thr Asp Val Lys Ser Val Glu
 530 535 540
 Phe Lys Asp Thr Ala Asn Gly Ala Asn Gly Ala Ser Thr Lys Ile Thr
 545 550 555 560
 Lys Asp Gly Leu Thr Ile Thr Pro Ala Asn Gly Ala Gly Ala Ala Gly
 565 570 575
 Ala Asn Thr Ala Asn Thr Ile Ser Val Thr Lys Asp Gly Ile Ser Ala
 580 585 590
 Gly Asn Lys Ala Val Lys Asn Val Val Ser Gly Leu Lys Lys Phe Gly
 595 600 605
 Asp Ala Asn Phe Asp Pro Leu Thr Ser Ser Ala Asp Asn Leu Thr Lys
 610 615 620
 Gln Tyr Asp Asn Ala Tyr Lys Gly Leu Thr Asn Leu Asp Glu Lys Ser
 625 630 635 640
 Lys Gly Lys Gln Thr Pro Thr Val Ala Asp Asn Thr Ala Ala Thr Val
 645 650 655
 Gly Asp Leu Arg Gly Leu Gly Trp Val Ile Ser Ala Asp Lys Thr Lys
 660 665 670
 Gly Glu Leu Asn Lys Glu Tyr Asn Ala Gln Val Arg Asn Ala Asn Glu
 675 680 685
 Val Lys Phe Lys Ser Gly Asn Gly Ile Asn Val Ser Gly Lys Thr Leu
 690 695 700
 Asp Asn Gly Thr Arg Glu Ile Thr Phe Glu Leu Ala Lys Asp Glu Asn
 705 710 715 720
 Ala Ile Ala Phe Gly Ser Gly Ser Lys Ala Leu Arg Asp Asn Thr Val
 725 730 735
 Ala Ile Gly Thr Gly Asn Val Val Asn Ala Glu Lys Ser Gly Ala Phe
 740 745 750
 Gly Asp Pro Asn Tyr Ile Glu Asp Lys Ala Gly Gly Ser Tyr Ala Phe
 755 760 765
 Gly Asn Asp Asn Arg Ile Thr Ser Lys Asn Thr Phe Val Leu Gly Asn
 770 775 780
 Ser Val Asn Ala Lys Arg Asp Ala Asn Gly Asn Val Leu Thr Glu Glu
 785 790 795 800
 Lys Glu Val Val Gly Lys Asp Gly Ala Lys Thr Lys Val Thr Val Pro
 805 810 815

Gln Ala Leu Gly Glu Thr Val Glu Asn Ser Val Tyr Leu Gly Asn Ala
 820 825 830
 Ser Thr Ala Thr Lys Asp Lys Gly Lys Asn Leu Lys Ser Asp Gly Thr
 835 840 845
 Ala Gly Asn Thr Thr Thr Ala Gly Ala Thr Gly Thr Val Asn Gly Phe
 850 855 860
 Ala Gly Ala Thr Ala His Gly Ala Val Ser Val Gly Ala Ser Gly Glu
 865 870 875 880
 Glu Arg Arg Ile Gln Asn Val Ala Ala Gly Glu Ile Ser Ala Thr Ser
 885 890 895
 Thr Asp Ala Ile Asn Gly Ser Gln Leu Tyr Ala Val Ala Lys Gly Val
 900 905 910
 Thr Asn Leu Ala Gly Gln Val Asn Lys Val Gly Lys Arg Ala Asp Ala
 915 920 925
 Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala Ser Met
 930 935 940
 Pro Gly Lys Ser Met Val Ser Ile Ala Gly Ser Ser Tyr Gln Gly Gln
 945 950 955 960
 Ser Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly Lys Val
 965 970 975
 Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr Gly Val
 980 985 990
 Ala Ala Gly Val Gly Tyr Gln Trp Asn Ser Gly Ser
 995 1000

<210> 31

<211> 3300

<212> DNA

<213> Haemophilus influenzae

<400> 31

atgaacaaaa tttttaacgt tatttggaat gttatgactc aaacttgggc tgtcgtatct 60
 gaactcactc gcgccacac caaacgtgcc tccgcaaccg tggcgaccgc cgtattggcg 120
 acgcagttgt ctgcaacggc tgaagcgaac agtagtgctt ctgttacgag taggttgaat 180
 gtttatggcg atacgaatac taaattcaat gcagccaata attcaatagc agattttaat 240
 aaacaaaaatg atggtgttca cgatgggtta ttaaattctga atgaaaacgg tgcgaataaa 300
 aagctgttgg tggatgacaa tactgcggcg accgtaggcg atttacgtaa attgggctgg 360
 gtcgtatcaa ccaaaaatgg caaggaaaat gagaaaagcc aacaagtcaa acaggcggat 420
 gaagtgttgt ttaaaggcag caaaggcggg gtgcagggtta cttccacctc tgaaaacggc 480
 aaacacgcca ttacctttgc tttagcgaaa gaccttgata tgagaactgc gactgtgagt 540
 gataccttaa cgattggcgg tagttactact acaggtagtg caacaacacc aaaagtgaat 600
 gtgactagca cggcaagcgg cttgaacttt gcgaaaaggc ctacagggtgc taatggcgat 660
 actacggttc acttgactaa tattgcttca actttgcaag atactctatt gaatactggg 720
 gttgtgagta aattagatgg taatggtatt actgctgacg agaaaaaacg tgcggcaagc 780
 gttcaagatg ttttaaatag tgggttggaa atcaagggtg ttaaaacagg tgcgacgact 840
 tctgataacg ttgattttgt ccgtacttac gacacagttg agtttttgag cggaagtga 900
 gaaactacac tggttacagt ggatagtgaag agtaatggaa aatctactaa agttaaatac 960

```

ggtgcgaaga cctctgttat caaagaaaaa gacggtaagt tatttactgg aaaagctaata 1020
aaagacacaa atcaagtcgc aagtaataat gcagctgatg atacggatga gggcaaaggc 1080
ttagtcactg cagagactgt tatcaatgca gtaacaagg ctggttggag aattaaaaca 1140
acgggtgcta ataataaagc tgggtcagttt gaaactgtca catcaggcac aaatgtaacc 1200
tttgctgatg gcaatggtag aactgcagtc gtaactggcg atgctaccaaa tggattact 1260
gttaaatatg aagcgaaaagt tggcgacggc ttgaagattg gtaacgacca aaaaatcact 1320
gcagatacga ccgcacttac tgtgacgggc ggtaaagtta ctgcccctga tgcaaccaat 1380
ggtaagaaac ttgttaatgc aagtggttta gctgatgcgt taaacaaatt aagtggact 1440
gcaaaagctg aagcagatac tgctaattgc ggcgagcttg atggaactgc agatgaaaaa 1500
gaagttaaag caggcgaaac ggtaaccttt aaagcgggca agaacttaaa agtgaaacaa 1560
gatggtgcca actttactta ttcactgcaa gatgctttaa caggcttaac gagcattact 1620
ttaggtacag gaaataatgg tgcgaaaact gaaatcaaca aagacggctt aaccatcaca 1680
ccagcaaatg gtgcggtgac aaataatgca aacaccatca gcgtaaccaaa agacggcatt 1740
agtgcggggc gtcagtcggt taaaaacggt gtgagcggac tgaagaaatt tgggtgatgcg 1800
aatttcgatc cgctgactag ctccgccgac aacttaacga aacaatatga cgatgcctat 1860
aaaggcttga ccaatttggg tgaaaaaggc gcggacaagc aaactctgac tgttgccgac 1920
aatactgccg caaccgtggg cgatttgccg ggcttgggct gggctatttc tgcggacaaa 1980
accacaggcg aactcgataa ggaatacaac gcgcaagtgc gtaacgcaa tgaaagtga 2040
ttcaaaagcg gcaacggtat caatgtttcc ggtaaaactg tcaacggtag gcgtgaaatt 2100
acttttgaat tggctaaaag cgaagtgggt aaatcgaatg aatttactgt caaagaaacc 2160
aatggcaagg aaacgagcct gggttaaagt ggcgataaat attacagcaa agaggatatt 2220
gaccagcaa ccggtaaaac gaaagttaaa aatggcaatg cagttgctgc gaaatatcaa 2280
gataaagatg gcaaagtcgt ttctgctgac ggcagcagca ataccgctgt taccctaacc 2340
aacaagggtt atggctatgt aacaggtaac caagtggcag atgcgattgc gaaatcaggc 2400
tttgagcttg gtttggtgta tgcagaaaaa gcgaaagctg cgtttggcga tgaaacaaaa 2460
gccttgtctt ctgataaatt ggaaaccgta aatgccaacg acaaagtcg ttttgcta 2520
ggtttaataa ccaaagttag cgcggcaacg gtggaaagca tcgatgcaaa cggcgataaa 2580
gtgactacaa cctttgtgaa aaccgatgtg gaattgcctt taacgcaaat ctacaatacc 2640
gatgcaaacg gtaagaaaat cggtaaaaat ggcgataaat ggtattacac gaaagatgac 2700
ggctcaactg atatgactaa agaagttacc cttggcaatg tggattcaga cggcaagaaa 2760
gttggtgaaag aagacaacaa gtggtatcac gtraaatctg atggttctac ggataaaaca 2820
cagggtggtc aagaagctaa agtttctacc gatgaaaaac acgttgtcag cttgatcca 2880
aatgatcaat caaaaggtaa aggcgtggtc attacaata tggctaattg cgaaatttct 2940
gccacttcca ccgatgcgat taacggaagt cagttgtatg ccgtggcaaa aggggtaaca 3000
aaccttgctg gacaagtga taatcttgag ggcaaagtga ataaagtggg caaacgtgca 3060
gatgcaggta ctgcaagtgc attagcggct tcacagttac cacaagccac tatgccagg 3120
aaatcaatgg tttctattgc gggaagtagt tatcaaggtc aaaatgggtt agctatcggg 3180
gtatcaagaa tttccgataa tggcaaaagt attattcgct tgtcaggcac aaccaatagt 3240
caaggtaaaa caggcggtgc agcaggtgtt gggtaccagt ggtaatagaa ttccggatcc 3300

```

<210> 32

<211> 1094

<212> PRT

<213> Haemophilus influenzae

<400> 32

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
  1             5             10             15

```

```

Ala Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Arg Ala Ser Ala
      20             25             30

```

```

Thr Val Ala Thr Ala Val Leu Ala Thr Gln Leu Ser Ala Thr Ala Glu
      35             40             45

```

```

Ala Asn Ser Ser Ala Ser Val Thr Ser Arg Leu Asn Val Tyr Gly Asp
      50             55             60

```

Thr Asn Thr Lys Phe Asn Ala Ala Asn Asn Ser Ile Ala Asp Leu Asn
 65 70 75 80
 Lys Gln Asn Asp Gly Val His Asp Gly Leu Leu Asn Leu Asn Glu Asn
 85 90 95
 Gly Ala Asn Lys Lys Leu Leu Val Asp Asp Asn Thr Ala Ala Thr Val
 100 105 110
 Gly Asp Leu Arg Lys Leu Gly Trp Val Val Ser Thr Lys Asn Gly Lys
 115 120 125
 Glu Asn Glu Lys Ser Gln Gln Val Lys Gln Ala Asp Glu Val Leu Phe
 130 135 140
 Lys Gly Ser Lys Gly Gly Val Gln Val Thr Ser Thr Ser Glu Asn Gly
 145 150 155 160
 Lys His Ala Ile Thr Phe Ala Leu Ala Lys Asp Leu Asp Met Arg Thr
 165 170 175
 Ala Thr Val Ser Asp Thr Leu Thr Ile Gly Gly Ser Thr Thr Thr Gly
 180 185 190
 Ser Ala Thr Thr Pro Lys Val Asn Val Thr Ser Thr Ala Ser Gly Leu
 195 200 205
 Asn Phe Ala Lys Gly Ala Thr Gly Ala Asn Gly Asp Thr Thr Val His
 210 215 220
 Leu Thr Asn Ile Ala Ser Thr Leu Gln Asp Thr Leu Leu Asn Thr Gly
 225 230 235 240
 Val Val Ser Lys Leu Asp Gly Asn Gly Ile Thr Ala Asp Glu Lys Lys
 245 250 255
 Arg Ala Ala Ser Val Gln Asp Val Leu Asn Ser Gly Trp Asn Ile Lys
 260 265 270
 Gly Val Lys Thr Gly Ala Thr Thr Ser Asp Asn Val Asp Phe Val Arg
 275 280 285
 Thr Tyr Asp Thr Val Glu Phe Leu Ser Gly Ser Glu Glu Thr Thr Leu
 290 295 300
 Val Thr Val Asp Ser Glu Ser Asn Gly Lys Ser Thr Lys Val Lys Ile
 305 310 315 320
 Gly Ala Lys Thr Ser Val Ile Lys Glu Lys Asp Gly Lys Leu Phe Thr
 325 330 335
 Gly Lys Ala Asn Lys Asp Thr Asn Gln Val Ala Ser Asn Asn Ala Ala
 340 345 350
 Asp Asp Thr Asp Glu Gly Lys Gly Leu Val Thr Ala Glu Thr Val Ile
 355 360 365
 Asn Ala Val Asn Lys Ala Gly Trp Arg Ile Lys Thr Thr Gly Ala Asn
 370 375 380

Asn Gln Ala Gly Gln Phe Glu Thr Val Thr Ser Gly Thr Asn Val Thr
 385 390 395 400
 Phe Ala Asp Gly Asn Gly Thr Thr Ala Val Val Thr Gly Asp Ala Thr
 405 410 415
 Asn Gly Ile Thr Val Lys Tyr Glu Ala Lys Val Gly Asp Gly Leu Lys
 420 425 430
 Ile Gly Asn Asp Gln Lys Ile Thr Ala Asp Thr Thr Ala Leu Thr Val
 435 440 445
 Thr Gly Gly Lys Val Thr Ala Pro Asp Ala Thr Asn Gly Lys Lys Leu
 450 455 460
 Val Asn Ala Ser Gly Leu Ala Asp Ala Leu Asn Lys Leu Ser Trp Thr
 465 470 475 480
 Ala Lys Ala Glu Ala Asp Thr Ala Asn Gly Gly Glu Leu Asp Gly Thr
 485 490 495
 Ala Asp Glu Lys Glu Val Lys Ala Gly Glu Thr Val Thr Phe Lys Ala
 500 505 510
 Gly Lys Asn Leu Lys Val Lys Gln Asp Gly Ala Asn Phe Thr Tyr Ser
 515 520 525
 Leu Gln Asp Ala Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Thr Gly
 530 535 540
 Asn Asn Gly Ala Lys Thr Glu Ile Asn Lys Asp Gly Leu Thr Ile Thr
 545 550 555 560
 Pro Ala Asn Gly Ala Gly Ala Asn Asn Ala Asn Thr Ile Ser Val Thr
 565 570 575
 Lys Asp Gly Ile Ser Ala Gly Gly Gln Ser Val Lys Asn Val Val Ser
 580 585 590
 Gly Leu Lys Lys Phe Gly Asp Ala Asn Phe Asp Pro Leu Thr Ser Ser
 595 600 605
 Ala Asp Asn Leu Thr Lys Gln Tyr Asp Asp Ala Tyr Lys Gly Leu Thr
 610 615 620
 Asn Leu Asp Glu Lys Gly Ala Asp Lys Gln Thr Leu Thr Val Ala Asp
 625 630 635 640
 Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly Trp Val Ile
 645 650 655
 Ser Ala Asp Lys Thr Thr Gly Glu Leu Asp Lys Glu Tyr Asn Ala Gln
 660 665 670
 Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn Gly Ile Asn
 675 680 685

Val Ser Gly Lys Thr Val Asn Gly Arg Arg Glu Ile Thr Phe Glu Leu
 690 695 700
 Ala Lys Gly Glu Val Val Lys Ser Asn Glu Phe Thr Val Lys Glu Thr
 705 710 715 720
 Asn Gly Lys Glu Thr Ser Leu Val Lys Val Gly Asp Lys Tyr Tyr Ser
 725 730 735
 Lys Glu Asp Ile Asp Pro Ala Thr Gly Lys Pro Lys Val Thr Asn Gly
 740 745 750
 Asn Ala Val Ala Ala Lys Tyr Gln Asp Lys Asp Gly Lys Val Val Ser
 755 760 765
 Ala Asp Gly Ser Ser Asn Thr Ala Val Thr Leu Thr Asn Lys Gly Tyr
 770 775 780
 Gly Tyr Val Thr Gly Asn Gln Val Ala Asp Ala Ile Ala Lys Ser Gly
 785 790 795 800
 Phe Glu Leu Gly Leu Ala Asp Ala Glu Lys Ala Lys Ala Ala Phe Gly
 805 810 815
 Asp Glu Thr Lys Ala Leu Ser Ser Asp Lys Leu Glu Thr Val Asn Ala
 820 825 830
 Asn Asp Lys Val Arg Phe Ala Asn Gly Leu Asn Thr Lys Val Ser Ala
 835 840 845
 Ala Thr Val Glu Ser Ile Asp Ala Asn Gly Asp Lys Val Thr Thr Thr
 850 855 860
 Phe Val Lys Thr Asp Val Glu Leu Pro Leu Thr Gln Ile Tyr Asn Thr
 865 870 875 880
 Asp Ala Asn Gly Lys Lys Ile Val Lys Asn Gly Asp Lys Trp Tyr Tyr
 885 890 895
 Thr Lys Asp Asp Gly Ser Thr Asp Met Thr Lys Glu Val Thr Leu Gly
 900 905 910
 Asn Val Asp Ser Asp Gly Lys Lys Val Val Lys Glu Asp Asn Lys Trp
 915 920 925
 Tyr His Val Lys Ser Asp Gly Ser Thr Asp Lys Thr Gln Val Val Glu
 930 935 940
 Glu Ala Lys Val Ser Thr Asp Glu Lys His Val Val Ser Leu Asp Pro
 945 950 955 960
 Asn Asp Gln Ser Lys Gly Lys Gly Val Val Ile Asn Asn Met Ala Asn
 965 970 975
 Gly Glu Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln Leu
 980 985 990
 Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly Gln Val Asn Asn
 995 1000 1005

Leu Glu Gly Lys Val Asn Lys Val Gly Lys Arg Ala Asp Ala Gly Thr
1010 1015 1020

Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala Thr Met Pro Gly
1025 1030 1035 1040

Lys Ser Met Val Ser Ile Ala Gly Ser Ser Tyr Gln Gly Gln Asn Gly
1045 1050 1055

Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly Lys Val Ile Ile
1060 1065 1070

Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr Gly Val Ala Ala
1075 1080 1085

Gly Val Gly Tyr Gln Trp
1090

<210> 33

<211> 3339

<212> DNA

<213> Haemophilus influenzae

<400> 33

```
gcgaattcat atgaacaaaa tttttaacgt tatttggaaat gttgtgactc aaacttgggt 60
tgctcgatct gaactcactc gcgcccacac caaatgcgcc tccgccaccg tggcggttgc 120
cgtattggca actgcgttgt ctgcaacggc tgaagcgaac aacaatactt ctgttacgaa 180
tgggttgaat gcttatggcg atactaattt taatacaacc aataattcga tagcagattt 240
ggaaaaacac gttcaagatg cttataaagg cttattaaat ctgaatgaaa aagatacaaa 300
taagtcaagt ttcttggttg cgcacaatac cgccgcaacc gtaggcaatt tgcgtaaaat 360
gggctgggta ttgtctagca aaaacggcac aagggaacgag aaaagctatc aagtaaaaca 420
agctgatgaa gttctcttta ctggatctgg tgctgcaacg gttagtcca gctctaaaga 480
cggtaaacat accattacca tttctgttac caaaggtagt tttgtgagg taaaaactga 540
tgcaactact ggaggtaag taaacgccga ccgtggtaaa gtgaaagctg aggacgagaa 600
tggagctgat gttgataaga aagttgcaac tgtaaaagat gttgctaagg cgattaacga 660
tgccgcaact ttctgtaaaag tggaaagcac agatgatgac attgaaaatg gtgctgcagg 720
caaaaatgaa actacagacc aagctctcaa agcaggcgac accttaacct taaaagcggg 780
taaaaactta aaagctaagt tagaccaaaa gtataaatca gtaacctttg ctttagcgaa 840
agaccttgat gtgacctctg cgaagtgag tgataagttg tctattggta aagatacgaa 900
taaagttgat attaccagt atgcaaatgg cttgaaattg gcgaaaacag gtaacggaaa 960
tggtaaaaac ggtaatgtcc acttaaatgg tattgcttcg actttgaccg ataccattac 1020
aggatgaca acacaagcaa gcaatggcgt ggctgtgcag aatcataatc gtgctgcgag 1080
tgtggctgat gtattaaatg caggctggaa tattcaaggc aacggagcga gcgttgattt 1140
tgtcaatgct tacgacacag tagattttgt caatggtaca aacaccaatg tgaacgttac 1200
gactgatacg gctcacaaaa agacaaccgt ccgtgtggat gtaacaggct tgccggttca 1260
atatgttacg gaagacggca aaaccgttgt gaaagtggac aataagtatt acgaagctaa 1320
gcaagacggg tcggcgggata tggataaaaa agtcgaaaaat ggcgagctgg cgaaaaccaa 1380
agtgaattg gtgtcggcaa gcggtcaaaa tccggtgaaa atcagcaatg ttgcggaagg 1440
cacggaagaa aacgatggcg tcagctttaa gcaattgaaa gccttgcaag agaaacagg 1500
tactttaact gcgagcaatg cttatgccaa tgggtgtaac gatgccgacg gcggcaaggc 1560
aactcaaact ttaacaatg gtttgaattt taaatttaaa tccacagacg gcgagttgtt 1620
gaacatcaaa gtagaaaatg acacagttac ctttacgccg aaaaaagggt cggtacaggt 1680
tggcgaagac ggtaaggcta cgattcaaaa tggtaacgaa acaaccgacg gtttggttga 1740
agcttccgaa ttggttgaaa gcctgaacaa actgggctgg aaagtgggcg ttgataaaga 1800
cggcagcggc gagcttgatg gtgcattcaa tgaaacttta gtgaagtcgg gcgataaagt 1860
aactttgaaa gccggcgaga atctgaagg caaacaagac ggcacaaact tcacttacgc 1920
gctcaaagat gaattgacgg gcgtgaagag cgtggagttt aaagacacgg cgaatggttc 1980
```

```

aaacgggtgca agcacgaaga ttaccaaaga cggcttgacc attacgtcgg caaacgggtgc 2040
gaatgggtgcg gcggcgactg atgcggacaa gattaaagtg gcttcagacg gcatcagtgc 2100
gggtaataaaa gcgggttaaaa acgttggtgag cggactgaag aaatttgggtg atgcgaattt 2160
caatccactg accagttccg cgcacaactt aacgaaacaa tatgacgatg cctataaagg 2220
cttgaccaat ttggatgaaa aagggtgcgga caagcaaact ctgactgttg ccgacaatac 2280
tgccgcaacc gtgggcgatt tgcgcggctt gggctgggtc atttctgcgg acaaaaccac 2340
aggcgaaactc aataaggaat acaacgcgca agtgcgtaac gccaatgaag tgaaattcaa 2400
gagcggcaac ggtatccatg tttccggtaa aacgggtcaac ggtaggcgcg aaattacttt 2460
tgaattggct aaagacgaaa atgccattgc tttcggttat ggctcaaaag ccttgcgcga 2520
taacacgggtg gcaattggta cgggcaacgt tgtgaatgcg gaaaaatctg gtgcattcgg 2580
cgatccgaac tacatcgaag ataaagccgg tggcagctac gctttcggta acgataaccg 2640
tattacttct aaaaacactt ttgtgttggg taatggagtt aatgcgaaat ataaagccaa 2700
tgagatggtt gatacggaaa ccgtaaccgt taaggacaaa gacggtaaag agactaccgt 2760
tactgttcct aaagcggttag gggctacggt tgaaaactcc gtttatttgg gtaataaatc 2820
gactgcgaca aaagataagg gtaaaaacct gaaatctgat ggtacggcgg gtaacactac 2880
aactgtctggc acaacgggta cggtaaacgg ctttgcgggt gcaacggcgc acggtgcggt 2940
ttctgtcggc gcaagcggcg aagaaagacg tatccaaaac gtcgcggcag gcgaaatttc 3000
cgccacttcc accgatgcga ttaacggcag ccagttgtat gctgtggcaa aaggggtaac 3060
aaatcttgcg ggacaagtga ataaagtggg caaacgtgca gatgcaggtg cagcaagtgc 3120
attagcagct tcacagttac cacaagcctc tatgccaggt aaateaatgg tttctattgc 3180
gggaagtagt tatcaaggtc aaaatggttt agctatcggg gtatcacgaa tttccgataa 3240
tggcaagtg attattecgt tgtcaggcac aaccaatagc caaggtaaaa caggcggtgc 3300
agcaggtggt ggttaccagt ggtaatagaa ttgatccgc 3339

```

<210> 34

<211> 1104

<212> PRT

<213> Haemophilus influenzae

<400> 34

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp
  1             5             10             15

```

```

Val Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Cys Ala Ser Ala
      20             25             30

```

```

Thr Val Ala Val Ala Val Leu Ala Thr Ala Leu Ser Ala Thr Ala Glu
      35             40             45

```

```

Ala Asn Asn Asn Thr Ser Val Thr Asn Gly Leu Asn Ala Tyr Gly Asp
      50             55             60

```

```

Thr Asn Phe Asn Thr Thr Asn Asn Ser Ile Ala Asp Leu Glu Lys His
      65             70             75             80

```

```

Val Gln Asp Ala Tyr Lys Gly Leu Leu Asn Leu Asn Glu Lys Asp Thr
      85             90             95

```

```

Asn Lys Ser Ser Phe Leu Val Ala Asp Asn Thr Ala Ala Thr Val Gly
      100            105            110

```

```

Asn Leu Arg Lys Leu Gly Trp Val Leu Ser Ser Lys Asn Gly Thr Arg
      115            120            125

```

```

Asn Glu Lys Ser Tyr Gln Val Lys Gln Ala Asp Glu Val Leu Phe Thr
      130            135            140

```

Gly Ser Gly Ala Ala Thr Val Ser Ser Ser Ser Lys Asp Gly Lys His
 145 150 155 160
 Thr Ile Thr Ile Ser Val Thr Lys Gly Ser Phe Ala Glu Val Lys Thr
 165 170 175
 Asp Ala Thr Thr Gly Gly Gln Val Asn Ala Asp Arg Gly Lys Val Lys
 180 185 190
 Ala Glu Asp Glu Asn Gly Ala Asp Val Asp Lys Lys Val Ala Thr Val
 195 200 205
 Lys Asp Val Ala Lys Ala Ile Asn Asp Ala Ala Thr Phe Val Lys Val
 210 215 220
 Glu Ser Thr Asp Asp Asp Ile Glu Asn Gly Ala Ala Gly Lys Asn Glu
 225 230 235 240
 Thr Thr Asp Gln Ala Leu Lys Ala Gly Asp Thr Leu Thr Leu Lys Ala
 245 250 255
 Gly Lys Asn Leu Lys Ala Lys Leu Asp Gln Asn Gly Lys Ser Val Thr
 260 265 270
 Phe Ala Leu Ala Lys Asp Leu Asp Val Thr Ser Ala Lys Val Ser Asp
 275 280 285
 Lys Leu Ser Ile Gly Lys Asp Thr Asn Lys Val Asp Ile Thr Ser Asp
 290 295 300
 Ala Asn Gly Leu Lys Leu Ala Lys Thr Gly Asn Gly Asn Gly Gln Asn
 305 310 315 320
 Gly Asn Val His Leu Asn Gly Ile Ala Ser Thr Leu Thr Asp Thr Ile
 325 330 335
 Thr Gly Met Thr Thr Gln Ala Ser Asn Gly Val Ala Val Gln Asn His
 340 345 350
 Asn Arg Ala Ala Ser Val Ala Asp Val Leu Asn Ala Gly Trp Asn Ile
 355 360 365
 Gln Gly Asn Gly Ala Ser Val Asp Phe Val Asn Ala Tyr Asp Thr Val
 370 375 380
 Asp Phe Val Asn Gly Thr Asn Thr Asn Val Asn Val Thr Thr Asp Thr
 385 390 395 400
 Ala His Lys Lys Thr Thr Val Arg Val Asp Val Thr Gly Leu Pro Val
 405 410 415
 Gln Tyr Val Thr Glu Asp Gly Lys Thr Val Val Lys Val Asp Asn Lys
 420 425 430
 Tyr Tyr Glu Ala Lys Gln Asp Gly Ser Ala Asp Met Asp Lys Lys Val
 435 440 445
 Glu Asn Gly Glu Leu Ala Lys Thr Lys Val Lys Leu Val Ser Ala Ser
 450 455 460

Gly Gln Asn Pro Val Lys Ile Ser Asn Val Ala Glu Gly Thr Glu Glu
 465 470 475 480
 Asn Asp Ala Val Ser Phe Lys Gln Leu Lys Ala Leu Gln Glu Lys Gln
 485 490 495
 Val Thr Leu Thr Ala Ser Asn Ala Tyr Ala Asn Gly Gly Asn Asp Ala
 500 505 510
 Asp Gly Gly Lys Ala Thr Gln Thr Leu Asn Asn Gly Leu Asn Phe Lys
 515 520 525
 Phe Lys Ser Thr Asp Gly Glu Leu Leu Asn Ile Lys Val Glu Asn Asp
 530 535 540
 Thr Val Thr Phe Thr Pro Lys Lys Gly Ser Val Gln Val Gly Glu Asp
 545 550 555 560
 Gly Lys Ala Thr Ile Gln Asn Gly Thr Lys Thr Thr Asp Gly Leu Val
 565 570 575
 Glu Ala Ser Glu Leu Val Glu Ser Leu Asn Lys Leu Gly Trp Lys Val
 580 585 590
 Gly Val Asp Lys Asp Gly Ser Gly Glu Leu Asp Gly Ala Ser Asn Glu
 595 600 605
 Thr Leu Val Lys Ser Gly Asp Lys Val Thr Leu Lys Ala Gly Glu Asn
 610 615 620
 Leu Lys Val Lys Gln Asp Gly Thr Asn Phe Thr Tyr Ala Leu Lys Asp
 625 630 635 640
 Glu Leu Thr Gly Val Lys Ser Val Glu Phe Lys Asp Thr Ala Asn Gly
 645 650 655
 Ser Asn Gly Ala Ser Thr Lys Ile Thr Lys Asp Gly Leu Thr Ile Thr
 660 665 670
 Ser Ala Asn Gly Ala Asn Gly Ala Ala Ala Thr Asp Ala Asp Lys Ile
 675 680 685
 Lys Val Ala Ser Asp Gly Ile Ser Ala Gly Asn Lys Ala Val Lys Asn
 690 695 700
 Val Val Ser Gly Leu Lys Lys Phe Gly Asp Ala Asn Phe Asn Pro Leu
 705 710 715 720
 Thr Ser Ser Ala Asp Asn Leu Thr Lys Gln Tyr Asp Asp Ala Tyr Lys
 725 730 735
 Gly Leu Thr Asn Leu Asp-Glu Lys Gly Ala Asp Lys Gln Thr Leu Thr
 740 745 750
 Val Ala Asp Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly
 755 760 765

Trp Val Ile Ser Ala Asp Lys Thr Thr Gly Glu Leu Asn Lys Glu Tyr
 770 775 780
 Asn Ala Gln Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn
 785 790 795 800
 Gly Ile His Val Ser Gly Lys Thr Val Asn Gly Arg Arg Glu Ile Thr
 805 810 815
 Phe Glu Leu Ala Lys Asp Glu Asn Ala Ile Ala Phe Gly Tyr Gly Ser
 820 825 830
 Lys Ala Leu Arg Asp Asn Thr Val Ala Ile Gly Thr Gly Asn Val Val
 835 840 845
 Asn Ala Glu Lys Ser Gly Ala Phe Gly Asp Pro Asn Tyr Ile Glu Asp
 850 855 860
 Lys Ala Gly Gly Ser Tyr Ala Phe Gly Asn Asp Asn Arg Ile Thr Ser
 865 870 875 880
 Lys Asn Thr Phe Val Leu Gly Asn Gly Val Asn Ala Lys Tyr Lys Ala
 885 890 895
 Asn Gly Asp Val Asp Thr Glu Thr Val Thr Val Lys Asp Lys Asp Gly
 900 905 910
 Lys Glu Thr Thr Val Thr Val Pro Lys Ala Leu Gly Ala Thr Val Glu
 915 920 925
 Asn Ser Val Tyr Leu Gly Asn Lys Ser Thr Ala Thr Lys Asp Lys Gly
 930 935 940
 Lys Asn Leu Lys Ser Asp Gly Thr Ala Gly Asn Thr Thr Thr Ala Gly
 945 950 955 960
 Thr Thr Gly Thr Val Asn Gly Phe Ala Gly Ala Thr Ala His Gly Ala
 965 970 975
 Val Ser Val Gly Ala Ser Gly Glu Glu Arg Arg Ile Gln Asn Val Ala
 980 985 990
 Ala Gly Glu Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln
 995 1000 1005
 Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly Gln Val Asn
 1010 1015 1020
 Lys Val Gly Lys Arg Ala Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala
 1025 1030 1035 1040
 Ser Gln Leu Pro Gln Ala Ser Met Pro Gly Lys Ser Met Val Ser Ile
 1045 1050 1055
 Ala Gly Ser Ser Tyr Gln Gly Gln Asn Gly Leu Ala Ile Gly Val Ser
 1060 1065 1070
 Arg Ile Ser Asp Asn Gly Lys Val Ile Ile Arg Leu Ser Gly Thr Thr
 1075 1080 1085

Asn Ser Gln Gly Lys Thr Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
 1090 1095 1100

<210> 35
 <211> 7253
 <212> DNA
 <213> Haemophilus influenzae

<400> 35
 atgaacaaaa tttttaacgt tatttggaaat gttatgactc aaacttgggt tgtcgtatct 60
 gaactcactc gcacccacac caaacgcgcc tccgcaaccg tggagaccgc cgtattggcg 120
 acactgttgt ttgcaacgggt tcaggcgaat gctaccgatg aagatgaaga gtttagacccc 180
 gtagtacgca ctgctcccggt gttgagcttc cattccgata aagaaggcac gggagaaaaa 240
 gaagttacag aaaattcaaa ttggggaata tatttccaca ataaaggagt actaaaagcc 300
 ggagcaatca ccctcaaagc cggcgacaac ctgaaaatca aacaaagcac caatgccagt 360
 agcttcacct actcgtcgaa aaaagacctc acagatctga ccagtgttgc aactgaaaaa 420
 ttatcgtttg gcgcaaacgg cgataaagtt gatattacca gtgatgcaaa tggcttgaaa 480
 ttggcgaaaa caggtaacgg aaatgttcat ttgaatgggt tggattcaac tttgcctgat 540
 gcggtaacga atacaggtgt gttaagttca tcaagtttta cacctaataa tgttgaaaaa 600
 acaagagctg caactgttaa agatgtttta aatgcagggt ggaacattaa aggtgctaaa 660
 actgctggag gtaatgttga gagtgttgat ttagtgtccg cttataataa tgttgaattt 720
 attacaggcg ataaaaacac gcttgatgtt gtattaacag ctaaagaaaa cggtaaaaca 780
 accgaagtga aattcacacc gaaaacctct gttatcaaag aaaaagacgg taagttattt 840
 actggaagag agaataacga cacaataaaa gttacaagta acacggcgac tgataatata 900
 gatgagggtg atggccttagt cactgcaaaa gctgtgattg atgctgtgaa caaggctggc 960
 tggagagtta aaacaactac tgctaattgggt caaatggcg acttcgcaac tgttgcgtca 1020
 ggcacaaatg taacctttga agtggcgat ggtacaacag cgtcagtaac taaagatact 1080
 aacggcaatg gcatcactgt taagtacgac gcgaaagttg gcgacggctt gaaatttgat 1140
 agcgataaaa aaatcgttgc agatcagacc gcacttactg tgacaggtgg taaggtagct 1200
 gaaattgcta aagaagatga caagaaaaaa cttgttaatg caggcgattt ggtaacagct 1260
 ttaggtaatc taagttggaa agcaaaagct gaggtgata ctgatactga tgggtgcgtt 1320
 gaggggattt caaaagacca agaagtcaaa gcaggcgaaa cggtaacctt taaagcgggc 1380
 aagaacttaa aagtgaaca ggtgggtgcg aactttactt attcactgca agatgcttta 1440
 acgggtttta cgaagcttac tttagggtgggt acaactaatg gcggaatga tgcgaaaacc 1500
 gtcatacaaa aagacgggtt aaccatcacg ccagcaggta atggcggtac gacaggtaca 1560
 aacacatca gcgtaaccaa agatggcatt aaagcaggta ataaagctat tactaatgtt 1620
 gcgagtggtt taagagctta tgacgatgag aattttgatg ttttaataaa ctctgcaact 1680
 gattttaaata gacacgttga agatgcttat aaaggtttat taaatctaaa tgaaaaaaat 1740
 gcaaataaac aaccgtttgt gactgacagc acggcggcga ctgtaggcga tttacgtaaa 1800
 ttgggttggg tagtatcaac caaaaacgggt acgaaagaag aaagcaatca agttaaacaa 1860
 gctgatgaag tcctctttac cggagccgggt gctgctacgg ttacttccaa atctgaaaac 1920
 ggtaaacata cgattaccgt tagtgtggct gaaactaaag cggatagcgg tcttgaaaaa 1980
 gatggcgata ctattaagct caaagtggat aatcaaaaca ctgataatgt tttactgtt 2040
 ggtaataatg gtactgctgt cactaaaggt ggctttgaaa ctgttaaaac tggagcgact 2100
 gatgcagatc gcggtaaagt aactgtaaaa gatgctactg ctaatgacgc tgataagaaa 2160
 gtcgcaactg taaaagatgt tgcaaccgca attaatagtg cggcgacttt tgtgaaaaca 2220
 gagaatttaa ctacctctat tgatgaagat aatcctacag ataacggcaa agatgacgca 2280
 cttaaagcgg gcgatacctt aacctttaaa gcaggtaaaa acctgaaagt taaacgtgat 2340
 ggaaaaaata ttacttttga cttggcgaaa aaccttgagg tgaaaactgc gaaagtgagt 2400
 gatactttta cgattggcgg gaatacacct acaggtggca ctactgcgac gccaaaagt 2460
 aatattacta gcacggctga tgggttgaat ttgcaaaaag aaacagccga tgcctcgggt 2520
 tctaagaatg tttatttgaa aggtattgag acaactttaa ctgagccaag cgcgggagcg 2580
 aagtcttcac acgttgattt aaatgtggat gcgacgaaaa aatccaatgc agcaagtatt 2640
 gaagatgtat tgcgcgcagg ttggaatatt caaggtaatg gtaataatgt tgattatgta 2700

```

gcgacgtatg acacagtaaa ctttaccgat gacagcacag gtacaacaac ggtaaccgta 2760
acccaaaaag cagatggcaa aggtgctgac gttaaaatcg gtgcgaaaac ttctgttatc 2820
aaagaccaca acggcaaaact gtttacaggc aaagacctga aagatgcgaa taatggtgca 2880
accgttagtg aagatgatgg caaagacacc ggcacaggct tagttactgc aaaaactgtg 2940
attgatgcag taaataaaaag cgggtggagg gtaaccgggtg agggcgcgac tgcgaaaacc 3000
ggtgcaaccg ccgtgaatgc gggtaacgct gaaaccgtta catcaggcac gagcgtgaac 3060
ttcaaaaacg gcaatgcgac cacagcgacc gtaagcaaaag ataatggcaa catcaatgtc 3120
aaatacgatg taaatgttgg tgacggccttg aagattggcg atgacaaaaa aatcgttgca 3180
gacacgacca cacttactgt aacagggtggg aagggtgtctg ttctgtctgg tgctaatagt 3240
gttaataaca ataagaaact tgttaatgca gaggggttag cgactgcttt aaacaacct 3300
agctggacgg caaaagccga taaatatgca gatggcgagt cagagggcga aaccgaccaa 3360
gaagtcaaaag caggcgacaa agtaaccttt aaagcaggca agaacttaa agtgaaacag 3420
tctgaaaaag actttactta ttactgcaa gacacttta caggcttaac gagcattact 3480
ttaggtggta cagctaattg cagaaatgat acgggaaccg tcatcaacaa agacggctta 3540
accatcacgc tggcaaatgg tgctgcgcca ggcacagatg cgtctaaccg aaacaccatc 3600
agtgtaaacca aagacggcat tagtgcggtt aataaagaaa ttaccaatgt taagagtgt 3660
ttaaaaacct ataaagatac tcaaaacact gcagggtgcaa ctcaacctgc ggctaataca 3720
actgaagtag ccaaacaga cttgggtgat ttaactaaac ctgcgacagg tgcagctgga 3780
aatgggtgac atgcaaaagc tcccgatacc acagctgcaa ccgtaggcga cttgctgggt 3840
ttgggctggg tgctttcagc taagaaaact gcagatgaaa cacaagataa agagttccac 3900
gccgccgtta aaaacgcaaa tgaagttgag ttcgtgggtg aaaacgggtg aaccgtgtct 3960
gcaaaaactg ataacaacgg aaaacatact gtaacgattg atgttgaga agccaaagt 4020
ggtgatggtc ttgaaaaaga tactgacggc aagattaaac tcaagtaga taatacagat 4080
gggaataatc tattaaccgt tgatgcaaca aaagggtgat caaggcaca atgccaatga gcgcggtaaa 4200
aatgccgtaa caacagatgc aactacagcc tgggtgcaact gctaccgaaa ctgacaagaa aaaagtggca 4260
gtggttgtca aggtttcaaa agcttgetaa agcgattaac gatcgagcaa ctttcgtgaa agtggaatat 4320
actgttggtg ctagcattga tgatagccca acagatgatg gcgcaaatga tgctctcaa 4380
gacgacagtg ccttgacctt aaaagcgggt aaaaacttaa aagttaaacg tgatggtaaa 4440
gcaggcgaca ccttgacctt gaacgacctt agtgtaaaaa gcgcaaccgt tagcgataaa 4500
aatattactt ttgcccttgc caataaagtc aatatcaca gcgacacca aggtctgaac 4560
ttatcgcttg gtacaaacgg atagtaagac aggcgatgat gctaattatc acttaaatgg cattgcttca 4620
tctgctaaag atacattggt atacattggt cgcacaacca atttaggtgg taatggtatt 4680
actttaactg agaaaaaacg cgcggcgagc gttaaagatg tcttgaatgc gggttggaat 4740
actgataacg agaaaaaacg atctgcaaat aatcaagtgg agaatatcga ctttgtagca 4800
gttctggttg ttaaaccggc tgtagtgga gataaagaca ccacgagtgt aactgttgaa 4860
acctacgaca cagtggactt atggcaagag aaccgaagtt cgaagacttc tgttatcaa 4920
agtaaagata atggcaagag tacaggcaaa gagctgaagg atgctaacaa taatggcgta 4980
gaccacaacg gcaaaactgt caaagacgag ggtaatggtt tagtgactgc aaaagctgtg 5040
actgttaccg aaaccgacgg tggttggaga ttgaaaacaa caggtgctaa tggtcagaat 5100
attgatgccg tgaataaggc gtcaaggcaca aatgtaacct ttgctgatgg taatggcaca 5160
gatgacttcg caactgttgc aaacgacggc aggcgataaa atcgttgca acacgaccgt acttactgtg 5280
actgccgaag taactaaagc cggcgataaa tccgaataat ggcgatggta agaaatttgt tgatgcaagt 5340
gctgatggct taaaactaga tccgaataat tggacggcaa ctgctggtaa agaaggcact 5400
gcagatggta aagttacagc taaattaagc caagaagtca aagcggcgca caaagtaacc 5460
gggttagcgg atgcgttaa atcagcaggg gaaaatcaaa aagactttac ctactcgctg 5520
ggtgaagttg atcctgcaaa gacgacacct gaccagcgta gagttcaaag acgcaaacgg cggtaacagg 5580
tttaagccg gcgacaacct caaagacggc ttgaccatta cgccggcaaa cgggtgcgggt 5640
aaaaagagc tgaaagacct aacaccatt agcgtaacca aagatggcat tagcgcggt 5700
agtgaagca ccaagattac tgtgagcgga ctgaagaaat ttggtgatgg tcatcgttg 5760
gcgacagtg caaactgc ttacaaacgt ttttgaagag cattatgaca atgcctataa agacttgacc 5820
aataaagcag ctgttgctga ggataataat ccgactgttg ccgacaatac cgctgcaacc 5880
gcaaatggca aaaaaggcgc gggctgggtc atttctgcgg acaaaaccac aggcgaaccc 5940
aatcaggaat acaacgcgca agtgcgtaac gccaatgaag tgaaattcaa gagcggcaac 6000
gggtatcaatg tttccggtaa aacattgaac ggtacgcgag tgattacctt tgaattggct 6060
aaaggcgag tgggttaaat gaattgaatt accgttaaga atgccgatgg ttcggaaacg 6120
aacttggtta aggttgcgga tatgtattac agcaagagg atattgaccc ggcaaccagt 6180
aaaccgatga caggtaaaac tgaaaaatat aagggtgaaa acggcaaggt cgtttctgct 6240

```



```

aacggcagca agaccgaagt taccctaacc aacaaagggt cgggctatgt aacaggtaac 6300
caagtggctg atgcgattgc gaaatcaggc tttgagcttg gtttggctga tgcggcagaa 6360
gctgaaaaag cctttgcaga aagcgcaaaa gacaagcaat tgtctaaaga taaagcggaa 6420
actgtaaatg cccacgataa agtccggttt gctaattggt taaataccaa agtgagcgcg 6480
gcaacggtgg aaagcactga tgcaaacggc gataaagtga ccacaacctt tgtgaaaacc 6540
gatgtggaat tgcctttaac gcaaatctac aataccgatg caaacggtaa taagatcggt 6600
aaaaaagctg acggaaaatg gtatgaactg aatgctgatg gtacggcgag taacaaagaa 6660
gtgacacttg gtaacgtgga tgcaaacggg aagaaagttg tgaaagtaac cgaaaatggt 6720
gcgataagt ggtattacac caatgctgac ggtgctgctg ataaaaccaa aggcgaagt 6780
agcaatgata aagtttctac cgatgaaaaa caggttgctc gccttgatcc gaacaatcaa 6840
tcgaacggca aaggcggtgt cattgacaat gtggctaata gcgaaatttc tgccacttcc 6900
accgatgcga ttaacggaag tcagttgtat gccgtggcaa aaggggtaac aaaccttgct 6960
ggacaagtga ataactctga gggcaaatg aataaagtgg gcaaacgtgc agatgcaggt 7020
acagcaagtg cattagcggc ttcacagtta ccacaagcca ctatgccagg taaatcaatg 7080
gttgctattg cgggaagttag ttatcaaggt caaatgggt tagctatcgg ggtatcaaga 7140
atttccgata atggcaaatg gattattcgc ttgtcaggca caaccaatag tcaaggtaaa 7200
acaggcggtg cagcaggtgt tggttaccag tggtaataga attccggatc cgc 7253

```

<210> 36

<211> 2411

<212> PRT

<213> Haemophilus influenzae

<400> 36

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
  1             5             10             15

```

```

Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Arg Ala Ser Ala
      20             25             30

```

```

Thr Val Glu Thr Ala Val Leu Ala Thr Leu Leu Phe Ala Thr Val Gln
    35             40             45

```

```

Ala Asn Ala Thr Asp Glu Asp Glu Glu Leu Asp Pro Val Val Arg Thr
    50             55             60

```

```

Ala Pro Val Leu Ser Phe His Ser Asp Lys Glu Gly Thr Gly Glu Lys
    65             70             75             80

```

```

Glu Val Thr Glu Asn Ser Asn Trp Gly Ile Tyr Phe His Asn Lys Gly
      85             90             95

```

```

Val Leu Lys Ala Gly Ala Ile Thr Leu Lys Ala Gly Asp Asn Leu Lys
    100            105            110

```

```

Ile Lys Gln Ser Thr Asn Ala Ser Ser Phe Thr Tyr Ser Leu Lys Lys
    115            120            125

```

```

Asp Leu Thr Asp Leu Thr Ser Val Ala Thr Glu Lys Leu Ser Phe Gly
    130            135            140

```

```

Ala Asn Gly Asp Lys Val Asp Ile Thr Ser Asp Ala Asn Gly Leu Lys
    145            150            155            160

```

```

Leu Ala Lys Thr Gly Asn Gly Asn Val His Leu Asn Gly Leu Asp Ser
    165            170            175

```

Thr Leu Pro Asp Ala Val Thr Asn Thr Gly Val Leu Ser Ser Ser Ser
 180 185 190
 Phe Thr Pro Asn Asp Val Glu Lys Thr Arg Ala Ala Thr Val Lys Asp
 195 200 205
 Val Leu Asn Ala Gly Trp Asn Ile Lys Gly Ala Lys Thr Ala Gly Gly
 210 215 220
 Asn Val Glu Ser Val Asp Leu Val Ser Ala Tyr Asn Asn Val Glu Phe
 225 230 235 240
 Ile Thr Gly Asp Lys Asn Thr Leu Asp Val Val Leu Thr Ala Lys Glu
 245 250 255
 Asn Gly Lys Thr Thr Glu Val Lys Phe Thr Pro Lys Thr Ser Val Ile
 260 265 270
 Lys Glu Lys Asp Gly Lys Leu Phe Thr Gly Lys Glu Asn Asn Asp Thr
 275 280 285
 Asn Lys Val Thr Ser Asn Thr Ala Thr Asp Asn Thr Asp Glu Gly Asn
 290 295 300
 Gly Leu Val Thr Ala Lys Ala Val Ile Asp Ala Val Asn Lys Ala Gly
 305 310 315 320
 Trp Arg Val Lys Thr Thr Thr Ala Asn Gly Gln Asn Gly Asp Phe Ala
 325 330 335
 Thr Val Ala Ser Gly Thr Asn Val Thr Phe Glu Ser Gly Asp Gly Thr
 340 345 350
 Thr Ala Ser Val Thr Lys Asp Thr Asn Gly Asn Gly Ile Thr Val Lys
 355 360 365
 Tyr Asp Ala Lys Val Gly Asp Gly Leu Lys Phe Asp Ser Asp Lys Lys
 370 375 380
 Ile Val Ala Asp Thr Thr Ala Leu Thr Val Thr Gly Gly Lys Val Ala
 385 390 395 400
 Glu Ile Ala Lys Glu Asp Asp Lys Lys Lys Leu Val Asn Ala Gly Asp
 405 410 415
 Leu Val Thr Ala Leu Gly Asn Leu Ser Trp Lys Ala Lys Ala Glu Ala
 420 425 430
 Asp Thr Asp Thr Asp Gly Ala Leu Glu Gly Ile Ser Lys Asp Gln Glu
 435 440 445
 Val Lys Ala Gly Glu Thr Val Thr Phe Lys Ala Gly Lys Asn Leu Lys
 450 455 460
 Val Lys Gln Asp Gly Ala Asn Phe Thr Tyr Ser Leu Gln Asp Ala Leu
 465 470 475 480
 Thr Gly Leu Thr Ser Ile Thr Leu Gly Gly Thr Thr Asn Gly Gly Asn
 485 490 495

Asp Ala Lys Thr Val Ile Asn Lys Asp Gly Leu Thr Ile Thr Pro Ala
 500 505 510
 Gly Asn Gly Gly Thr Thr Gly Thr Asn Thr Ile Ser Val Thr Lys Asp
 515 520 525
 Gly Ile Lys Ala Gly Asn Lys Ala Ile Thr Asn Val Ala Ser Gly Leu
 530 535 540
 Arg Ala Tyr Asp Asp Ala Asn Phe Asp Val Leu Asn Asn Ser Ala Thr
 545 550 555 560
 Asp Leu Asn Arg His Val Glu Asp Ala Tyr Lys Gly Leu Leu Asn Leu
 565 570 575
 Asn Glu Lys Asn Ala Asn Lys Gln Pro Leu Val Thr Asp Ser Thr Ala
 580 585 590
 Ala Thr Val Gly Asp Leu Arg Lys Leu Gly Trp Val Val Ser Thr Lys
 595 600 605
 Asn Gly Thr Lys Glu Glu Ser Asn Gln Val Lys Gln Ala Asp Glu Val
 610 615 620
 Leu Phe Thr Gly Ala Gly Ala Ala Thr Val Thr Ser Lys Ser Glu Asn
 625 630 635 640
 Gly Lys His Thr Ile Thr Val Ser Val Ala Glu Thr Lys Ala Asp Ser
 645 650 655
 Gly Leu Glu Lys Asp Gly Asp Thr Ile Lys Leu Lys Val Asp Asn Gln
 660 665 670
 Asn Thr Asp Asn Val Leu Thr Val Gly Asn Asn Gly Thr Ala Val Thr
 675 680 685
 Lys Gly Gly Phe Glu Thr Val Lys Thr Gly Ala Thr Asp Ala Asp Arg
 690 695 700
 Gly Lys Val Thr Val Lys Asp Ala Thr Ala Asn Asp Ala Asp Lys Lys
 705 710 715 720
 Val Ala Thr Val Lys Asp Val Ala Thr Ala Ile Asn Ser Ala Ala Thr
 725 730 735
 Phe Val Lys Thr Glu Asn Leu Thr Thr Ser Ile Asp Glu Asp Asn Pro
 740 745 750
 Thr Asp Asn Gly Lys Asp Asp Ala Leu Lys Ala Gly Asp Thr Leu Thr
 755 760 765
 Phe Lys Ala Gly Lys Asn Leu Lys Val Lys Arg Asp Gly Lys Asn Ile
 770 775 780
 Thr Phe Asp Leu Ala Lys Asn Leu Glu Val Lys Thr Ala Lys Val Ser
 785 790 795 800

Asp Thr Leu Thr Ile Gly Gly Asn Thr Pro Thr Gly Gly Thr Thr Ala
 805 810 815
 Thr Pro Lys Val Asn Ile Thr Ser Thr Ala Asp Gly Leu Asn Phe Ala
 820 825 830
 Lys Glu Thr Ala Asp Ala Ser Gly Ser Lys Asn Val Tyr Leu Lys Gly
 835 840 845
 Ile Ala Thr Thr Leu Thr Glu Pro Ser Ala Gly Ala Lys Ser Ser His
 850 855 860
 Val Asp Leu Asn Val Asp Ala Thr Lys Lys Ser Asn Ala Ala Ser Ile
 865 870 875 880
 Glu Asp Val Leu Arg Ala Gly Trp Asn Ile Gln Gly Asn Gly Asn Asn
 885 890 895
 Val Asp Tyr Val Ala Thr Tyr Asp Thr Val Asn Phe Thr Asp Asp Ser
 900 905 910
 Thr Gly Thr Thr Thr Val Thr Val Thr Gln Lys Ala Asp Gly Lys Gly
 915 920 925
 Ala Asp Val Lys Ile Gly Ala Lys Thr Ser Val Ile Lys Asp His Asn
 930 935 940
 Gly Lys Leu Phe Thr Gly Lys Asp Leu Lys Asp Ala Asn Asn Gly Ala
 945 950 955 960
 Thr Val Ser Glu Asp Asp Gly Lys Asp Thr Gly Thr Gly Leu Val Thr
 965 970 975
 Ala Lys Thr Val Ile Asp Ala Val Asn Lys Ser Gly Trp Arg Val Thr
 980 985 990
 Gly Glu Gly Ala Thr Ala Glu Thr Gly Ala Thr Ala Val Asn Ala Gly
 995 1000 1005
 Asn Ala Glu Thr Val Thr Ser Gly Thr Ser Val Asn Phe Lys Asn Gly
 1010 1015 1020
 Asn Ala Thr Thr Ala Thr Val Ser Lys Asp Asn Gly Asn Ile Asn Val
 1025 1030 1035 1040
 Lys Tyr Asp Val Asn Val Gly Asp Gly Leu Lys Ile Gly Asp Asp Lys
 1045 1050 1055
 Lys Ile Val Ala Asp Thr Thr Thr Leu Thr Val Thr Gly Gly Lys Val
 1060 1065 1070
 Ser Val Pro Ala Gly Ala Asn Ser Val Asn Asn Asn Lys Lys Leu Val
 1075 1080 1085
 Asn Ala Glu Gly Leu Ala Thr Ala Leu Asn Asn Leu Ser Trp Thr Ala
 1090 1095 1100
 Lys Ala Asp Lys Tyr Ala Asp Gly Glu Ser Glu Gly Glu Thr Asp Gln
 1105 1110 1115 1120

Glu Val Lys Ala Gly Asp Lys Val Thr Phe Lys Ala Gly Lys Asn Leu
 1125 1130 1135
 Lys Val Lys Gln Ser Glu Lys Asp Phe Thr Tyr Ser Leu Gln Asp Thr
 1140 1145 1150
 Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Gly Thr Ala Asn Gly Arg
 1155 1160 1165
 Asn Asp Thr Gly Thr Val Ile Asn Lys Asp Gly Leu Thr Ile Thr Leu
 1170 1175 1180
 Ala Asn Gly Ala Ala Ala Gly Thr Asp Ala Ser Asn Gly Asn Thr Ile
 1185 1190 1195 1200
 Ser Val Thr Lys Asp Gly Ile Ser Ala Gly Asn Lys Glu Ile Thr Asn
 1205 1210 1215
 Val Lys Ser Ala Leu Lys Thr Tyr Lys Asp Thr Gln Asn Thr Ala Gly
 1220 1225 1230
 Ala Thr Gln Pro Ala Ala Asn Thr Ala Glu Val Ala Lys Gln Asp Leu
 1235 1240 1245
 Val Asp Leu Thr Lys Pro Ala Thr Gly Ala Ala Gly Asn Gly Ala Asp
 1250 1255 1260
 Ala Lys Ala Pro Asp Thr Thr Ala Ala Thr Val Gly Asp Leu Arg Gly
 1265 1270 1275 1280
 Leu Gly Trp Val Leu Ser Ala Lys Lys Thr Ala Asp Glu Thr Gln Asp
 1285 1290 1295
 Lys Glu Phe His Ala Ala Val Lys Asn Ala Asn Glu Val Glu Phe Val
 1300 1305 1310
 Gly Lys Asn Gly Ala Thr Val Ser Ala Lys Thr Asp Asn Asn Gly Lys
 1315 1320 1325
 His Thr Val Thr Ile Asp Val Ala Glu Ala Lys Val Gly Asp Gly Leu
 1330 1335 1340
 Glu Lys Asp Thr Asp Gly Lys Ile Lys Leu Lys Val Asp Asn Thr Asp
 1345 1350 1355 1360
 Gly Asn Asn Leu Leu Thr Val Asp Ala Thr Lys Gly Ala Ser Val Ala
 1365 1370 1375
 Lys Gly Glu Phe Asn Ala Val Thr Thr Asp Ala Thr Thr Ala Gln Gly
 1380 1385 1390
 Thr Asn Ala Asn Glu Arg Gly Lys Val Val Val Lys Gly Ser Asn Gly
 1395 1400 1405
 Ala Thr Ala Thr Glu Thr Asp Lys Lys Lys Val Ala Thr Val Gly Asp
 1410 1415 1420

Val Ala Lys Ala Ile Asn Asp Ala Ala Thr Phe Val Lys Val Glu Asn
 1425 1430 1435 1440
 Asp Asp Ser Ala Thr Ile Asp Asp Ser Pro Thr Asp Asp Gly Ala Asn
 1445 1450 1455
 Asp Ala Leu Lys Ala Gly Asp Thr Leu Thr Leu Lys Ala Gly Lys Asn
 1460 1465 1470
 Leu Lys Val Lys Arg Asp Gly Lys Asn Ile Thr Phe Ala Leu Ala Asn
 1475 1480 1485
 Asp Leu Ser Val Lys Ser Ala Thr Val Ser Asp Lys Leu Ser Leu Gly
 1490 1495 1500
 Thr Asn Gly Asn Lys Val Asn Ile Thr Ser Asp Thr Lys Gly Leu Asn
 1505 1510 1515 1520
 Phe Ala Lys Asp Ser Lys Thr Gly Asp Asp Ala Asn Ile His Leu Asn
 1525 1530 1535
 Gly Ile Ala Ser Thr Leu Thr Asp Thr Leu Leu Asn Ser Gly Ala Thr
 1540 1545 1550
 Thr Asn Leu Gly Gly Asn Gly Ile Thr Asp Asn Glu Lys Lys Arg Ala
 1555 1560 1565
 Ala Ser Val Lys Asp Val Leu Asn Ala Gly Trp Asn Val Arg Gly Val
 1570 1575 1580
 Lys Pro Ala Ser Ala Asn Asn Gln Val Glu Asn Ile Asp Phe Val Ala
 1585 1590 1595 1600
 Thr Tyr Asp Thr Val Asp Phe Val Ser Gly Asp Lys Asp Thr Thr Ser
 1605 1610 1615
 Val Thr Val Glu Ser Lys Asp Asn Gly Lys Arg Thr Glu Val Lys Ile
 1620 1625 1630
 Gly Ala Lys Thr Ser Val Ile Lys Asp His Asn Gly Lys Leu Phe Thr
 1635 1640 1645
 Gly Lys Glu Leu Lys Asp Ala Asn Asn Asn Gly Val Thr Val Thr Glu
 1650 1655 1660
 Thr Asp Gly Lys Asp Glu Gly Asn Gly Leu Val Thr Ala Lys Ala Val
 1665 1670 1675 1680
 Ile Asp Ala Val Asn Lys Ala Gly Trp Arg Val Lys Thr Thr Gly Ala
 1685 1690 1695
 Asn Gly Gln Asn Asp Asp Phe Ala Thr Val Ala Ser Gly Thr Asn Val
 1700 1705 1710
 Thr Phe Ala Asp Gly Asn Gly Thr Thr Ala Glu Val Thr Lys Ala Asn
 1715 1720 1725

Asp Gly Ser Ile Thr Val Lys Tyr Asn Val Lys Val Ala Asp Gly Leu
 1730 1735 1740
 Lys Leu Asp Gly Asp Lys Ile Val Ala Asp Thr Thr Val Leu Thr Val
 1745 1750 1755 1760
 Ala Asp Gly Lys Val Thr Ala Pro Asn Asn Gly Asp Gly Lys Lys Phe
 1765 1770 1775
 Val Asp Ala Ser Gly Leu Ala Asp Ala Leu Asn Lys Leu Ser Trp Thr
 1780 1785 1790
 Ala Thr Ala Gly Lys Glu Gly Thr Gly Glu Val Asp Pro Ala Asn Ser
 1795 1800 1805
 Ala Gly Gln Glu Val Lys Ala Gly Asp Lys Val Thr Phe Lys Ala Gly
 1810 1815 1820
 Asp Asn Leu Lys Ile Lys Gln Ser Gly Lys Asp Phe Thr Tyr Ser Leu
 1825 1830 1835 1840
 Lys Lys Glu Leu Lys Asp Leu Thr Ser Val Glu Phe Lys Asp Ala Asn
 1845 1850 1855
 Gly Gly Thr Gly Ser Glu Ser Thr Lys Ile Thr Lys Asp Gly Leu Thr
 1860 1865 1870
 Ile Thr Pro Ala Asn Gly Ala Gly Ala Ala Gly Ala Asn Thr Ala Asn
 1875 1880 1885
 Thr Ile Ser Val Thr Lys Asp Gly Ile Ser Ala Gly Asn Lys Ala Val
 1890 1895 1900
 Thr Asn Val Val Ser Gly Leu Lys Lys Phe Gly Asp Gly His Thr Leu
 1905 1910 1915 1920
 Ala Asn Gly Thr Val Ala Asp Phe Glu Lys His Tyr Asp Asn Ala Tyr
 1925 1930 1935
 Lys Asp Leu Thr Asn Leu Asp Glu Lys Gly Ala Asp Asn Asn Pro Thr
 1940 1945 1950
 Val Ala Asp Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly
 1955 1960 1965
 Trp Val Ile Ser Ala Asp Lys Thr Thr Gly Glu Pro Asn Gln Glu Tyr
 1970 1975 1980
 Asn Ala Gln Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn
 1985 1990 1995 2000
 Gly Ile Asn Val Ser Gly Lys Thr Leu Asn Gly Thr Arg Val Ile Thr
 2005 2010 2015
 Phe Glu Leu Ala Lys Gly Glu Val Val Lys Ser Asn Glu Phe Thr Val
 2020 2025 2030
 Lys Asn Ala Asp Gly Ser Glu Thr Asn Leu Val Lys Val Gly Asp Met
 2035 2040 2045

Tyr Tyr Ser Lys Glu Asp Ile Asp Pro Ala Thr Ser Lys Pro Met Thr
 2050 2055 2060
 Gly Lys Thr Glu Lys Tyr Lys Val Glu Asn Gly Lys Val Val Ser Ala
 2065 2070 2075 2080
 Asn Gly Ser Lys Thr Glu Val Thr Leu Thr Asn Lys Gly Ser Gly Tyr
 2085 2090 2095
 Val Thr Gly Asn Gln Val Ala Asp Ala Ile Ala Lys Ser Gly Phe Glu
 2100 2105 2110
 Leu Gly Leu Ala Asp Ala Ala Glu Ala Glu Lys Ala Phe Ala Glu Ser
 2115 2120 2125
 Ala Lys Asp Lys Gln Leu Ser Lys Asp Lys Ala Glu Thr Val Asn Ala
 2130 2135 2140
 His Asp Lys Val Arg Phe Ala Asn Gly Leu Asn Thr Lys Val Ser Ala
 2145 2150 2155 2160
 Ala Thr Val Glu Ser Thr Asp Ala Asn Gly Asp Lys Val Thr Thr Thr
 2165 2170 2175
 Phe Val Lys Thr Asp Val Glu Leu Pro Leu Thr Gln Ile Tyr Asn Thr
 2180 2185 2190
 Asp Ala Asn Gly Asn Lys Ile Val Lys Lys Ala Asp Gly Lys Trp Tyr
 2195 2200 2205
 Glu Leu Asn Ala Asp Gly Thr Ala Ser Asn Lys Glu Val Thr Leu Gly
 2210 2215 2220
 Asn Val Asp Ala Asn Gly Lys Lys Val Val Lys Val Thr Glu Asn Gly
 2225 2230 2235 2240
 Ala Asp Lys Trp Tyr Tyr Thr Asn Ala Asp Gly Ala Ala Asp Lys Thr
 2245 2250 2255
 Lys Gly Glu Val Ser Asn Asp Lys Val Ser Thr Asp Glu Lys His Val
 2260 2265 2270
 Val Arg Leu Asp Pro Asn Asn Gln Ser Asn Gly Lys Gly Val Val Ile
 2275 2280 2285
 Asp Asn Val Ala Asn Gly Glu Ile Ser Ala Thr Ser Thr Asp Ala Ile
 2290 2295 2300
 Asn Gly Ser Gln Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala
 2305 2310 2315 2320
 Gly Gln Val Asn Asn Leu Glu Gly Lys Val Asn Lys Val Gly Lys Arg
 2325 2330 2335
 Ala Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln
 2340 2345 2350

Ala Thr Met Pro Gly Lys Ser Met Val Ala Ile Ala Gly Ser Ser Tyr
2355 2360 2365

Gln Gly Gln Asn Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn
2370 2375 2380

Gly Lys Val Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys
2385 2390 2395 2400

Thr Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
2405 2410

<210> 37

<211> 1872

<212> DNA

<213> Haemophilus influenzae

<400> 37

```

gaattctatt accactggta accaacacct gctgcaacgc cagaaacagc acaacaaatt 60
cactggctac atcaattttac caaagctcgc attcaatggc gcaaaaccca ttccttattc 120
tttaaagaaa aaccccgatta tgcttttggt ctggcagaaa acggcgaagt gcaagaaatc 180
aaagcagaat atcgccgcat tgccaatcaa attgtggaag aagcaatgat tattgccaac 240
atctgcgccg cccaattttt acacgaacag gcaaaaacag gcattttcaa cgcccacagc 300
ggttttgata aaaaaatactt agaaaatgcg caccatttct taatggcaaa tttagccaat 360
gaacaaaatc aaactgaact ggcagaacgt tattcagtag aaaacttagc aaccttaaac 420
ggctattgcc aaatgcgtca cgtatttgaa cccatcgaaa gcgattattt agaactgcgt 480
ttacgccggt atttaacttt cgccgaattt aaatcagaat tagcaccgca ctttgggtctt 540
ggtttagaag gctatgccac ttggacatcg cccatccgca aatattcaga tatggttaat 600
catcgcttaa tcaaagccgt gctggcaaaa cagccttatg aaaaaccaca aaatgacgtg 660
ttggcacggt tgcaagagtc tcgcccgcga aatcgcttag tggaacgtga tattgccgat 720
tggttatatt gccgttatct tgctgacaaa gtggctgaaa atgtggaatt taatgcagaa 780
gtgcaagatg taatgcgtgc aggccttacgc gtacaactgc tcgaaaatgg tgcacgcta 840
tttattcctg ccgccacggt gcacaacaac aaagaagaaa tacagctaaa ccctgacgaa 900
ctcgccctct atataaaagg cgaaacgcact taaaaaatag gcgacattgt gaaagtgaaa 960
ctcacagaag tgaagaagc aactcgcatg attgtgggcg aaatacttca ataaattgcc 1020
gttccaatat gttacggaag acggcaaaac cgttgtgaaa gtgggcaatg agtattacga 1080
agccaagcaa gacgggttcg cgcatatgga taaaaaagtc aaaaatggcg agctggtgaa 1140
aactaaagtg aaattggtat cggcaaacgg tacaaatccg gtgaaaatca gcaatggtgc 1200
ggaaggcacg gaagataccg atgcgggtcag ctttaagcag ttgaaagcct tgcaaaacaa 1260
acagggttacg ttaagcgcg gcaatgctta tgccaatggc ggtagcgatg ccgacgtcgg 1320
caaggtaact caaactttaa gcaatgggtt gaattttaaa tttaaatcca cagacggcga 1380
gttggtgaac atcaaagcag acaaggacac ggttaccatt acgcgggcaa gcggtgcgaa 1440
tggtgcggcg gcgactgatg ccgacaagat taaagtggct tcagacggca ttagcgcggg 1500
taataaagca gttaaaaacg tcgcggcagg cgaaatttcc gccacttcca ccgatgcgat 1560
taacggcagt cagttgtatg ccgtggcaaa gggggtaaca aaccttgctg gacaagtga 1620
taaagtgggc aaacgtgcag atgcaggtac agcaagtgca ttagcggtt cacagttacc 1680
acaagcctct atgcccgggt aatcaatggt ttctattgcg ggaagtagtt atcaaggtca 1740
aagtgggtta gctatcgggg tatcaagaat ttccgataat ggcaaattga ttattcgctt 1800
gtcaggcaca accaatagcc aaggtaaaaac aggcgttgca gcagggtgtg gttaccagt 1860
gtaatagaat tc
1872

```

<210> 38

<211> 616

<212> PRT

<213> Haemophilus influenzae

<400> 38

Tyr Tyr His Trp Pro Thr Pro Ala Ala Thr Pro Glu Thr Ala Gln Gln
 1 5 10 15
 Ile His Trp Leu His Gln Phe Thr Lys Ala Arg Ile Gln Trp Arg Lys
 20 25 30
 Thr His Ser Leu Phe Phe Lys Glu Lys Pro Asp Tyr Ala Phe Val Leu
 35 40 45
 Ala Glu Asn Gly Lys Val Gln Glu Ile Lys Ala Glu Tyr Arg Arg Ile
 50 55 60
 Ala Asn Gln Ile Val Glu Glu Ala Met Ile Ile Ala Asn Ile Cys Ala
 65 70 75 80
 Ala Gln Phe Leu His Glu Gln Ala Lys Thr Gly Ile Phe Asn Ala His
 85 90 95
 Ser Gly Phe Asp Lys Lys Tyr Leu Glu Asn Ala His His Phe Leu Met
 100 105 110
 Ala Asn Leu Ala Asn Glu Gln Asn Gln Thr Glu Leu Ala Glu Arg Tyr
 115 120 125
 Ser Val Glu Asn Leu Ala Thr Leu Asn Gly Tyr Cys Gln Met Arg His
 130 135 140
 Asp Ile Glu Pro Ile Glu Ser Asp Tyr Leu Glu Leu Arg Leu Arg Arg
 145 150 155 160
 Tyr Leu Thr Phe Ala Glu Phe Lys Ser Glu Leu Ala Pro His Phe Gly
 165 170 175
 Leu Gly Leu Glu Gly Tyr Ala Thr Trp Thr Ser Pro Ile Arg Lys Tyr
 180 185 190
 Ser Asp Met Val Asn His Arg Leu Ile Lys Ala Val Leu Ala Lys Gln
 195 200 205
 Pro Tyr Glu Lys Pro Gln Asn Asp Val Leu Ala Arg Leu Gln Glu Ser
 210 215 220
 Arg Arg Gln Asn Arg Leu Val Glu Arg Asp Ile Ala Asp Trp Leu Tyr
 225 230 235 240
 Cys Arg Tyr Leu Ala Asp Lys Val Ala Glu Asn Val Glu Phe Asn Ala
 245 250 255
 Glu Val Gln Asp Val Met Arg Ala Gly Leu Arg Val Gln Leu Leu Glu
 260 265 270
 Asn Gly Ala Ser Leu Phe Ile Pro Ala Ala Thr Leu His Asn Asn Lys
 275 280 285
 Glu Glu Ile Gln Leu Asn Pro Asp Glu Leu Ala Leu Tyr Ile Lys Gly
 290 295 300

Glu Arg Thr Tyr Lys Ile Gly Asp Ile Val Lys Val Lys Leu Thr Glu
 305 310 315 320
 Val Lys Glu Ala Thr Arg Ser Ile Val Gly Glu Ile Leu Gln Leu Pro
 325 330 335
 Phe Gln Tyr Val Thr Glu Asp Gly Lys Thr Val Val Lys Val Gly Asn
 340 345 350
 Glu Tyr Tyr Glu Ala Lys Gln Asp Gly Ser Ala Asp Met Asp Lys Lys
 355 360 365
 Val Lys Asn Gly Glu Leu Val Lys Thr Lys Val Lys Leu Val Ser Ala
 370 375 380
 Asn Gly Thr Asn Pro Val Lys Ile Ser Asn Val Ala Glu Gly Thr Glu
 385 390 395 400
 Asp Thr Asp Ala Val Ser Phe Lys Gln Leu Lys Ala Leu Gln Asn Lys
 405 410 415
 Gln Val Thr Leu Ser Ala Ser Asn Ala Tyr Ala Asn Gly Gly Ser Asp
 420 425 430
 Ala Asp Val Gly Lys Val Thr Gln Thr Leu Ser Asn Gly Leu Asn Phe
 435 440 445
 Lys Phe Lys Ser Thr Asp Gly Glu Leu Leu Asn Ile Lys Ala Asp Lys
 450 455 460
 Asp Thr Val Thr Ile Thr Arg Ala Ser Gly Ala Asn Gly Ala Ala Ala
 465 470 475 480
 Thr Asp Ala Asp Lys Ile Lys Val Ala Ser Asp Gly Ile Ser Ala Gly
 485 490 495
 Asn Lys Ala Val Lys Asn Val Ala Ala Gly Glu Ile Ser Ala Thr Ser
 500 505 510
 Thr Asp Ala Ile Asn Gly Ser Gln Leu Tyr Ala Val Ala Lys Gly Val
 515 520 525
 Thr Asn Leu Ala Gly Gln Val Asn Lys Val Gly Lys Arg Ala Asp Ala
 530 535 540
 Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala Ser Met
 545 550 555 560
 Pro Gly Lys Ser Met Val Ser Ile Ala Gly Ser Ser Tyr Gln Gly Gln
 565 570 575
 Ser Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly Lys Leu
 580 585 590
 Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr Gly Val
 595 600 605
 Ala Ala Gly Val Gly Tyr Gln Trp
 610 615

<210> 39
 <211> 24
 <212> DNA
 <213> Haemophilus influenzae

<400> 39
 tgcgcacccat ttcttaatgg caaa 24

<210> 40
 <211> 26
 <212> DNA
 <213> Haemophilus influenzae

<400> 40
 gacgtgttgg cacgtttgca agagtc 26

<210> 41
 <211> 23
 <212> DNA
 <213> Haemophilus influenzae

<400> 41
 gcaagacggg tcggcgggata tgg 23

<210> 42
 <211> 24
 <212> DNA
 <213> Haemophilus influenzae

<400> 42
 ccgccacttc caccgatgcg atta 24

<210> 43
 <211> 3294
 <212> DNA
 <213> Haemophilus influenzae

<400> 43
 atgaacaaaa tttttaacgt tatttggaat gttgtgactc aaacttgggt tgtcgtatct 60
 gaactcactc gcacccacac caaatgcgcc tccgccaccg tggcggttgc cgtattggca 120
 accctgttgt ccgcaacggg tgaggcgaac aacaatactc ctgttacgaa taagttgaag 180
 gcttatggcg atgcgaattt taatttcact aataattcga tagcagatgc agaaaaacaa 240
 gttcaagagg cttataaagg tttattaaat ctaaatgaaa aaaatgcgag tgataaaactg 300
 ttggtggagg acaataactgc ggcgaccgta ggcaatttgc gtaaattggg ctgggtattg 360
 tctagcaaaa acggcacaag gaacgagaaa agccaacaag tcaaactatgc ggatgaagtg 420
 ttgtttgaag gcaaaggcgg tgtgcagggt acttccacct ctgaaaacgg caaacacacc 480
 attacctttg ctttagcgaa agaccttggg gtgaaaactg cgactgtgag tgatacctta 540
 acgattggcg gtggtgctgc tgcagggtgt acaacaacac cgaaagtga tctaactagt 600
 acaactgatg gcttgaagtt cgctaaagat gctgcgggtg ctaatggcga tactacgggt 660
 cacttgaatg gtattggttc aaccttgaca gacacgcttg tgggttctcc tgctactcat 720
 attgacggag gagatcaaaag tacgcattac actcgtgcag caagtatcaa ggatgtcttg 780
 aatgcgggtt ggaatatcaa ggggtgttaa gctgggtcaa caactgggtc atcagaaaat 840
 gtcgattttg ttcatactta cgatactgtt gagttcttga gtgcgggatac agagaccacg 900

```

actgttactg tagatagcaa agaaaacggt aagagaaccg aagttaaaaat cgggtgcgaag 960
acttctgtta tcaaagaaaa agacggtaag ttattttactg gaaaagctaa caaagagaca 1020
aataaagtgt atggtgctaa cgcgactgaa gatgcagacg aaggcaaagg cttagtgtact 1080
gcgaaagatg tgattgacgc agtgaataag actggttggg gaattaaaac aaccgatgct 1140
aatggtcaaa atggcgactt cgcgaactgtt gcatcaggca caaatgtaac ctttgctagt 1200
ggtaatggta caactgcgac tgtaactaat ggcaccgatg gtattaccgt taagtatgat 1260
gcgaaagtgt gcgacggcctt aaaactagat ggcgataaaa tcgctgcaga tacgaccgca 1320
cttactgtga atgatggtaa gaacgctaata aatccgaaag gtaaagtggc tgatgttgct 1380
tcaactgacg agaagaattt ggttacagca aaagggttag taacagcctt aaacagtcta 1440
agctggacta caactgctgc tgaggcggac ggtggtacgc ttgatggaaa tgcaagttag 1500
caagaagtta aagcgggcca taaagtaacc tttaaagcag gcaagaactt aaaagtgaag 1560
caagaggggtg cgaactttac ttattcactg caagatgctt taacaggcctt aacgagcatt 1620
acttttaggtt caggaaataa tgggtgcgaaa actgaaatca acaaagacgg ctttaaccatc 1680
acaccagcaa atggtgcggg tgcaataaat gcaaacacca tcagcgtaac caaagacggc 1740
attagtgcgg gcggtcagtc ggttaaaaaac gttgtgagcg gactgaagaa atttggtgat 1800
gcgaattttcg atccgctgac tagctccgcc gacaacttaa cgaaacaaaa tgacgatgcc 1860
tataaaggct tgaccaattt ggatgaaaaa ggtacagaca agcaaaactcc agttgttgcc 1920
gacaataccg ccgcaaccgt ggcgatttg cgcggcttg gctgggtcat ttctgcggac 1980
aaaaccacag gcggtcaaac ggaatatcac gatcaagttc ggaatgcgaa cgaagtgaag 2040
ttcaaaagcg gcaacggtat caatgtttcc ggtaaaacgg tcaacggtag gcgtgaaatt 2100
acttttgaat tggctaaagg tgaagtgggt aaatcgaaat aatttaccgt caaagaaacc 2160
aatggaaagg aaacgagcct ggttaaagtt ggcgataaat attacagcaa agaggatatt 2220
gacttaacaa caggtcagcc taaattaaaa gatggcaata cagttgctgc gaaatatcaa 2280
gataaagggt gcaaagtcgt ttctgtaacg gataatactg aagctaccat aaccaacaaa 2340
ggttctggct atgtaacagg taaccaagtg gcagatgcga ttgcgaaatc aggcctttgag 2400
cttggcttgg ctgatgaagc tgatgcgaaa cgggcgtttg atgataagac aaaagcctta 2460
tctgctggta caacggaaat tgtaaatgcc caccgataag tccgttttgc taatggttta 2520
aataccaag tgagcgcggc aacgggtggaa agcaccgatg caaacggcga taaagtgacc 2580
acaacctttg tgaaaaccga tgtggaattg cctttaacgc aaatctacaa taccgatgca 2640
aacggtaaga aaatcactaa agttgtcaaa gatgggcaaa ctaaatggta tgaactgaat 2700
gctgacggta cggctgatat gaccaaagaa gttaccctcg gtaacgtgga ttcagacggc 2760
aagaaagtgt tgaaagacaa cgatggcaag tggatatcag ccaaagctga cgggtactgcg 2820
gataaaacca aaggcgaggt gagcaatgat aaagtttcta ccgatgaaaa acacgttgctc 2880
agccttgatc caaatgatca atcaaaaggt aaaggtgtcg tgattgacaa tgtggctaata 2940
ggcgatatct ctgccacttc caccgatgcg attaacggaa gtcagttgta tgctgtggca 3000
aaaggggtaa caaaccttgc tggacaagtg aataatcttg agggcaaaag gaataaagtg 3060
ggcaaacgtg cagatgcagg tacagcaagt gcattagcgg cttcacagtt accacaagcc 3120
actatgccag gtaaatcaat ggttgcattt gcgggaagta gttatcaagg tcaaaatggt 3180
ttagctatcg gggatcaag aatttccgat aatggcaaaag tgattattcg cttgtcaggc 3240
acaaccaata gtcaaggtaa aacaggcggt gcagcaggtg ttggttacca gtgg 3294

```

<210> 44

<211> 1098

<212> PRT

<213> Haemophilus influenzae

<400> 44

```

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp
  1             5             10             15

```

```

Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Cys Ala Ser Ala
          20             25             30

```

```

Thr Val Ala Val Ala Val Leu Ala Thr Leu Leu Ser Ala Thr Val Glu
    35             40             45

```

```

Ala Asn Asn Asn Thr Pro Val Thr Asn Lys Leu Lys Ala Tyr Gly Asp
    50             55             60

```

Ala Asn Phe Asn Phe Thr Asn Asn Ser Ile Ala Asp Ala Glu Lys Gln
 65 70 75 80
 Val Gln Glu Ala Tyr Lys Gly Leu Leu Asn Leu Asn Glu Lys Asn Ala
 85 90 95
 Ser Asp Lys Leu Leu Val Glu Asp Asn Thr Ala Ala Thr Val Gly Asn
 100 105 110
 Leu Arg Lys Leu Gly Trp Val Leu Ser Ser Lys Asn Gly Thr Arg Asn
 115 120 125
 Glu Lys Ser Gln Gln Val Lys His Ala Asp Glu Val Leu Phe Glu Gly
 130 135 140
 Lys Gly Gly Val Gln Val Thr Ser Thr Ser Glu Asn Gly Lys His Thr
 145 150 155 160
 Ile Thr Phe Ala Leu Ala Lys Asp Leu Gly Val Lys Thr Ala Thr Val
 165 170 175
 Ser Asp Thr Leu Thr Ile Gly Gly Gly Ala Ala Ala Gly Ala Thr Thr
 180 185 190
 Thr Pro Lys Val Asn Val Thr Ser Thr Thr Asp Gly Leu Lys Phe Ala
 195 200 205
 Lys Asp Ala Ala Gly Ala Asn Gly Asp Thr Thr Val His Leu Asn Gly
 210 215 220
 Ile Gly Ser Thr Leu Thr Asp Thr Leu Val Gly Ser Pro Ala Thr His
 225 230 235 240
 Ile Asp Gly Gly Asp Gln Ser Thr His Tyr Thr Arg Ala Ala Ser Ile
 245 250 255
 Lys Asp Val Leu Asn Ala Gly Trp Asn Ile Lys Gly Val Lys Ala Gly
 260 265 270
 Ser Thr Thr Gly Gln Ser Glu Asn Val Asp Phe Val His Thr Tyr Asp
 275 280 285
 Thr Val Glu Phe Leu Ser Ala Asp Thr Glu Thr Thr Thr Val Thr Val
 290 295 300
 Asp Ser Lys Glu Asn Gly Lys Arg Thr Glu Val Lys Ile Gly Ala Lys
 305 310 315 320
 Thr Ser Val Ile Lys Glu Lys Asp Gly Lys Leu Phe Thr Gly Lys Ala
 325 330 335
 Asn Lys Glu Thr Asn Lys Val Asp Gly Ala Asn Ala Thr Glu Asp Ala
 340 345 350
 Asp Glu Gly Lys Gly Leu Val Thr Ala Lys Asp Val Ile Asp Ala Val
 355 360 365

Asn Lys Thr Gly Trp Arg Ile Lys Thr Thr Asp Ala Asn Gly Gln Asn
 370 375 380
 Gly Asp Phe Ala Thr Val Ala Ser Gly Thr Asn Val Thr Phe Ala Ser
 385 390 395 400
 Gly Asn Gly Thr Thr Ala Thr Val Thr Asn Gly Thr Asp Gly Ile Thr
 405 410 415
 Val Lys Tyr Asp Ala Lys Val Gly Asp Gly Leu Lys Leu Asp Gly Asp
 420 425 430
 Lys Ile Ala Ala Asp Thr Thr Ala Leu Thr Val Asn Asp Gly Lys Asn
 435 440 445
 Ala Asn Asn Pro Lys Gly Lys Val Ala Asp Val Ala Ser Thr Asp Glu
 450 455 460
 Lys Lys Leu Val Thr Ala Lys Gly Leu Val Thr Ala Leu Asn Ser Leu
 465 470 475 480
 Ser Trp Thr Thr Thr Ala Ala Glu Ala Asp Gly Gly Thr Leu Asp Gly
 485 490 495
 Asn Ala Ser Glu Gln Glu Val Lys Ala Gly Asp Lys Val Thr Phe Lys
 500 505 510
 Ala Gly Lys Asn Leu Lys Val Lys Gln Glu Gly Ala Asn Phe Thr Tyr
 515 520 525
 Ser Leu Gln Asp Ala Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Thr
 530 535 540
 Gly Asn Asn Gly Ala Lys Thr Glu Ile Asn Lys Asp Gly Leu Thr Ile
 545 550 555 560
 Thr Pro Ala Asn Gly Ala Gly Ala Asn Asn Ala Asn Thr Ile Ser Val
 565 570 575
 Thr Lys Asp Gly Ile Ser Ala Gly Gly Gln Ser Val Lys Asn Val Val
 580 585 590
 Ser Gly Leu Lys Lys Phe Gly Asp Ala Asn Phe Asp Pro Leu Thr Ser
 595 600 605
 Ser Ala Asp Asn Leu Thr Lys Gln Asn Asp Asp Ala Tyr Lys Gly Leu
 610 615 620
 Thr Asn Leu Asp Glu Lys Gly Thr Asp Lys Gln Thr Pro Val Val Ala
 625 630 635 640
 Asp Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly Trp Val
 645 650 655
 Ile Ser Ala Asp Lys Thr Thr Gly Gly Ser Thr Glu Tyr His Asp Gln
 660 665 670
 Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn Gly Ile Asn
 675 680 685

Val Ser Gly Lys Thr Val Asn Gly Arg Arg Glu Ile Thr Phe Glu Leu
 690 695 700
 Ala Lys Gly Glu Val Val Lys Ser Asn Glu Phe Thr Val Lys Glu Thr
 705 710 715 720
 Asn Gly Lys Glu Thr Ser Leu Val Lys Val Gly Asp Lys Tyr Tyr Ser
 725 730 735
 Lys Glu Asp Ile Asp Leu Thr Thr Gly Gln Pro Lys Leu Lys Asp Gly
 740 745 750
 Asn Thr Val Ala Ala Lys Tyr Gln Asp Lys Gly Gly Lys Val Val Ser
 755 760 765
 Val Thr Asp Asn Thr Glu Ala Thr Ile Thr Asn Lys Gly Ser Gly Tyr
 770 775 780
 Val Thr Gly Asn Gln Val Ala Asp Ala Ile Ala Lys Ser Gly Phe Glu
 785 790 795 800
 Leu Gly Leu Ala Asp Glu Ala Asp Ala Lys Arg Ala Phe Asp Asp Lys
 805 810 815
 Thr Lys Ala Leu Ser Ala Gly Thr Thr Glu Ile Val Asn Ala His Asp
 820 825 830
 Lys Val Arg Phe Ala Asn Gly Leu Asn Thr Lys Val Ser Ala Ala Thr
 835 840 845
 Val Glu Ser Thr Asp Ala Asn Gly Asp Lys Val Thr Thr Thr Phe Val
 850 855 860
 Lys Thr Asp Val Glu Leu Pro Leu Thr Gln Ile Tyr Asn Thr Asp Ala
 865 870 875 880
 Asn Gly Lys Lys Ile Thr Lys Val Val Lys Asp Gly Gln Thr Lys Trp
 885 890 895
 Tyr Glu Leu Asn Ala Asp Gly Thr Ala Asp Met Thr Lys Glu Val Thr
 900 905 910
 Leu Gly Asn Val Asp Ser Asp Gly Lys Lys Val Val Lys Asp Asn Asp
 915 920 925
 Gly Lys Trp Tyr His Ala Lys Ala Asp Gly Thr Ala Asp Lys Thr Lys
 930 935 940
 Gly Glu Val Ser Asn Asp Lys Val Ser Thr Asp Glu Lys His Val Val
 945 950 955 960
 Ser Leu Asp Pro Asn Asp Gln Ser Lys Gly Lys Gly Val Val Ile Asp
 965 970 975
 Asn Val Ala Asn Gly Asp Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn
 980 985 990

Gly Ser Gln Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly
 995 1000 1005

Gln Val Asn Asn Leu Glu Gly Lys Val Asn Lys Val Gly Lys Arg Ala
 1010 1015 1020

Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala
 1025 1030 1035 1040

Thr Met Pro Gly Lys Ser Met Val Ala Ile Ala Gly Ser Ser Tyr Gln
 1045 1050 1055

Gly Gln Asn Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly
 1060 1065 1070

Lys Val Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr
 1075 1080 1085

Gly Val Ala Ala Gly Val Gly Tyr Gln Trp
 1090 1095

<210> 45

<211> 330

<212> PRT

<213> Haemophilus influenzae

<400> 45

Pro Thr Pro Ala Ala Thr Pro Glu Thr Ala Gln Gln Ile His Trp Leu
 1 5 10 15

His Gln Phe Thr Lys Ala Arg Ile Gln Trp Arg Lys Thr His Ser Leu
 20 25 30

Phe Phe Lys Glu Lys Pro Asp Tyr Ala Phe Val Leu Ala Glu Asn Gly
 35 40 45

Lys Val Gln Glu Ile Lys Ala Glu Tyr Arg Arg Ile Ala Asn Gln Ile
 50 55 60

Val Glu Glu Ala Met Ile Ile Ala Asn Ile Cys Ala Ala Gln Phe Leu
 65 70 75 80

His Glu Gln Ala Lys Thr Gly Ile Phe Asn Ala His Ser Gly Phe Asp
 85 90 95

Lys Lys Tyr Leu Glu Asn Ala His His Phe Leu Met Ala Asn Leu Ala
 100 105 110

Asn Glu Gln Asn Gln Thr Glu Leu Ala Glu Arg Tyr Ser Val Glu Asn
 115 120 125

Leu Ala Thr Leu Asn Gly Tyr Cys Gln Met Arg His Asp Ile Glu Pro
 130 135 140

Ile Glu Ser Asp Tyr Leu Glu Leu Arg Leu Arg Arg Tyr Leu Thr Phe
 145 150 155 160

Ala Glu Phe Lys Ser Glu Leu Ala Pro His Phe Gly Leu Gly Leu Glu
165 170 175

Gly Tyr Ala Thr Trp Thr Ser Pro Ile Arg Lys Tyr Ser Asp Met Val
180 185 190

Asn His Arg Leu Ile Lys Ala Val Leu Ala Lys Gln Pro Tyr Glu Lys
195 200 205

Pro Gln Asn Asp Val Leu Ala Arg Leu Gln Glu Ser Arg Arg Gln Asn
210 215 220

Arg Leu Val Glu Arg Asp Ile Ala Asp Trp Leu Tyr Cys Arg Tyr Leu
225 230 235 240

Ala Asp Lys Val Ala Glu Asn Val Glu Phe Asn Ala Glu Val Gln Asp
245 250 255

Val Met Arg Ala Gly Leu Arg Val Gln Leu Leu Glu Asn Gly Ala Ser
260 265 270

Leu Phe Ile Pro Ala Ala Thr Leu His Asn Asn Lys Glu Glu Ile Gln
275 280 285

Leu Asn Pro Asp Glu Leu Ala Leu Tyr Ile Lys Gly Glu Arg Thr Tyr
290 295 300

Lys Ile Gly Asp Ile Val Lys Val Lys Leu Thr Glu Val Lys Glu Ala
305 310 315 320

Thr Arg Ser Ile Val Gly Glu Ile Leu Gln
325 330

<210> 46

<211> 330

<212> PRT

<213> Haemophilus influenzae

<400> 46

Ala Trp Gln Pro Glu Met Pro Glu Thr Ala Gln Gln Ile His Trp Leu
1 5 10 15

His Gln Phe Thr Lys Ala Arg Ile Gln Trp Arg Lys Thr His Ser Leu
20 25 30

Phe Phe Lys Glu Lys Pro Asp Tyr Ala Phe Val Leu Ala Glu Asn Gly
35 40 45

Lys Val Gln Glu Ile Lys Ala Glu Tyr Arg Arg Ile Ala Asn Gln Ile
50 55 60

Val Glu Glu Ala Met Ile Ile Ala Asn Ile Cys Ala Ala Gln Phe Leu
65 70 75 80

His Glu Gln Ala Lys Thr Gly Ile Phe Asn Thr His Ser Gly Phe Asp
85 90 95

Lys Lys Phe Leu Glu Asn Ala His Asn Phe Leu Met Ala Asn Leu Ala
100 105 110

Asn Glu Gln Asn Gln Thr Glu Leu Ala Glu Arg Tyr Ser Val Glu Asn
115 120 125

Leu Ala Thr Leu Asn Gly Tyr Cys Gln Met Arg His Asp Ile Glu Pro
130 135 140

Ile Glu Ser Asp Tyr Leu Glu Leu Arg Leu Arg Arg Tyr Leu Thr Phe
145 150 155 160

Ala Glu Phe Lys Ser Glu Leu Ala Pro His Phe Gly Leu Gly Leu Glu
165 170 175

Gly Tyr Ala Thr Trp Thr Ser Pro Ile Arg Lys Tyr Ser Asp Met Val
180 185 190

Asn His Arg Leu Ile Lys Ala Val Leu Ala Lys Gln Pro Tyr Glu Lys
195 200 205

Pro Gln Asn Asp Val Leu Ala Arg Leu Gln Glu Ala Arg Arg Gln Asn
210 215 220

Arg Leu Val Glu Arg Asp Ile Ala Asp Trp Leu Tyr Cys Arg Tyr Leu
225 230 235 240

Ala Asp Lys Val Ala Ser Asn Ala Glu Phe Glu Ala Glu Val Gln Asp
245 250 255

Val Met Arg Ala Gly Leu Arg Val Gln Leu Leu Glu Asn Gly Ala Ser
260 265 270

Leu Phe Ile Pro Ala Ala Thr Leu His Asn Asn Lys Glu Glu Ile Gln
275 280 285

Leu Asn Pro Asp Glu Leu Ala Leu Tyr Ile Lys Gly Glu Arg Thr Tyr
290 295 300

Lys Ile Gly Asp Met Val Lys Val Lys Leu Thr Glu Val Lys Glu Ala
305 310 315 320

Thr Arg Ser Ile Val Gly Glu Ile Leu Gln
325 330

<210> 47

<211> 2353

<212> PRT

<213> Haemophilus influenzae

<400> 47

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
1 5 10 15

Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Arg Ala Ser Ala
20 25 30

o

Thr Val Glu Thr Ala Val Leu Ala Thr Leu Leu Phe Ala Thr Val Gln
 35 40 45
 Ala Asn Ala Thr Asp Glu Asp Glu Glu Leu Asp Pro Val Val Arg Thr
 50 55 60
 Ala Pro Val Leu Ser Phe His Ser Asp Lys Glu Gly Thr Gly Glu Lys
 65 70 75 80
 Glu Val Thr Glu Asn Ser Asn Trp Gly Ile Tyr Phe Asp Asn Lys Gly
 85 90 95
 Val Leu Lys Ala Gly Ala Ile Thr Leu Lys Ala Gly Asp Asn Leu Lys
 100 105 110
 Ile Lys Gln Asn Thr Asp Glu Ser Thr Asn Ala Ser Ser Phe Thr Tyr
 115 120 125
 Ser Leu Lys Lys Asp Leu Thr Asp Leu Thr Ser Val Ala Thr Glu Lys
 130 135 140
 Leu Ser Phe Gly Ala Asn Gly Asp Lys Val Asp Ile Thr Ser Asp Ala
 145 150 155 160
 Asn Gly Leu Lys Leu Ala Lys Thr Gly Asn Gly Asn Val His Leu Asn
 165 170 175
 Gly Leu Asp Ser Thr Leu Pro Asp Ala Val Thr Asn Thr Gly Val Leu
 180 185 190
 Ser Ser Ser Ser Phe Thr Pro Asn Asp Val Glu Lys Thr Arg Ala Ala
 195 200 205
 Thr Val Lys Asp Val Leu Asn Ala Gly Trp Asn Ile Lys Gly Ala Lys
 210 215 220
 Thr Ala Gly Gly Asn Val Glu Ser Val Asp Leu Val Ser Ala Tyr Asn
 225 230 235 240
 Asn Val Glu Phe Ile Thr Gly Asp Lys Asn Thr Leu Asp Val Val Leu
 245 250 255
 Thr Ala Lys Glu Asn Gly Lys Thr Thr Glu Val Lys Phe Thr Pro Lys
 260 265 270
 Thr Ser Val Ile Lys Glu Lys Asp Gly Lys Leu Phe Thr Gly Lys Glu
 275 280 285
 Asn Asn Asp Thr Asn Lys Val Thr Ser Asn Thr Ala Thr Asp Asn Thr
 290 295 300
 Asp Glu Gly Asn Gly Leu Val Thr Ala Lys Ala Val Ile Asp Ala Val
 305 310 315 320
 Asn Lys Ala Gly Trp Arg Val Lys Thr Thr Thr Ala Asn Gly Gln Asn
 325 330 335
 Gly Asp Phe Ala Thr Val Ala Ser Gly Thr Asn Val Thr Phe Glu Ser
 340 345 350

Gly Asp Gly Thr Thr Ala Ser Val Thr Lys Asp Thr Asn Gly Asn Gly
 355 360 365
 Ile Thr Val Lys Tyr Asp Ala Lys Val Gly Asp Gly Leu Lys Phe Asp
 370 375 380
 Ser Asp Lys Lys Ile Val Ala Asp Thr Thr Ala Leu Thr Val Thr Gly
 385 390 395 400
 Gly Lys Val Ala Glu Ile Ala Lys Glu Asp Asp Lys Lys Lys Leu Val
 405 410 415
 Asn Ala Gly Asp Leu Val Thr Ala Leu Gly Asn Leu Ser Trp Lys Ala
 420 425 430
 Lys Ala Glu Ala Asp Thr Asp Gly Ala Leu Glu Gly Ile Ser Lys Asp
 435 440 445
 Gln Glu Val Lys Ala Gly Glu Thr Val Thr Phe Lys Ala Gly Lys Asn
 450 455 460
 Leu Lys Val Lys Gln Asp Gly Ala Asn Phe Thr Tyr Ser Leu Gln Asp
 465 470 475 480
 Ala Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Gly Thr Thr Asn Gly
 485 490 495
 Gly Asn Asp Ala Lys Thr Val Ile Asn Lys Asp Gly Leu Thr Ile Thr
 500 505 510
 Pro Ala Gly Asn Gly Gly Thr Thr Gly Thr Asn Thr Ile Ser Val Thr
 515 520 525
 Lys Asp Gly Ile Lys Ala Gly Asn Lys Ala Ile Thr Asn Val Ala Ser
 530 535 540
 Gly Leu Arg Ala Tyr Asp Asp Ala Asn Phe Asp Val Leu Asn Asn Ser
 545 550 555 560
 Ala Thr Asp Leu Asn Arg His Val Glu Asp Ala Tyr Lys Gly Leu Leu
 565 570 575
 Asn Leu Asn Glu Lys Asn Ala Asn Lys Gln Pro Leu Val Thr Asp Ser
 580 585 590
 Thr Ala Ala Thr Val Gly Asp Leu Arg Lys Leu Gly Trp Val Val Ser
 595 600 605
 Thr Lys Asn Gly Thr Lys Glu Glu Ser Asn Gln Val Lys Gln Ala Asp
 610 615 620
 Glu Val Leu Phe Thr Gly Ala Gly Ala Ala Thr Val Thr Ser Lys Ser
 625 630 635 640
 Glu Asn Gly Lys His Thr Ile Thr Val Ser Val Ala Glu Thr Lys Ala
 645 650 655

Asp Cys Gly Leu Glu Lys Asp Gly Asp Thr Ile Lys Leu Lys Val Asp
 660 665 670
 Asn Gln Asn Thr Asp Asn Val Leu Thr Val Gly Asn Asn Gly Thr Ala
 675 680 685
 Val Thr Lys Gly Gly Phe Glu Thr Val Lys Thr Gly Ala Thr Asp Ala
 690 695 700
 Asp Arg Gly Lys Val Thr Val Lys Asp Ala Thr Ala Asn Asp Ala Asp
 705 710 715 720
 Lys Lys Val Ala Thr Val Lys Asp Val Ala Thr Ala Ile Asn Ser Ala
 725 730 735
 Ala Thr Phe Val Lys Thr Glu Asn Leu Thr Thr Ser Ile Asp Glu Asp
 740 745 750
 Asn Pro Thr Asp Asn Gly Lys Asp Asp Ala Leu Lys Ala Gly Asp Thr
 755 760 765
 Leu Thr Phe Lys Ala Gly Lys Asn Leu Lys Val Lys Arg Asp Gly Lys
 770 775 780
 Asn Ile Thr Phe Asp Leu Ala Lys Asn Leu Glu Val Lys Thr Ala Lys
 785 790 795 800
 Val Ser Asp Thr Leu Thr Ile Gly Gly Asn Thr Pro Thr Gly Gly Thr
 805 810 815
 Thr Ala Thr Pro Lys Val Asn Ile Thr Ser Thr Ala Asp Gly Leu Asn
 820 825 830
 Phe Ala Lys Glu Thr Ala Asp Ala Ser Gly Ser Lys Asn Val Tyr Leu
 835 840 845
 Lys Gly Ile Ala Thr Thr Leu Thr Glu Pro Ser Ala Gly Ala Lys Ser
 850 855 860
 Ser His Val Asp Leu Asn Val Asp Ala Thr Lys Lys Ser Asn Ala Ala
 865 870 875 880
 Ser Ile Glu Asp Val Leu Arg Ala Gly Trp Asn Ile Gln Gly Asn Gly
 885 890 895
 Asn Asn Val Asp Tyr Val Ala Thr Tyr Asp Thr Val Asn Phe Thr Asp
 900 905 910
 Asp Ser Thr Gly Thr Thr Thr Val Thr Val Thr Gln Lys Ala Asp Gly
 915 920 925
 Lys Gly Ala Asp Val Lys Ile Gly Ala Lys Thr Ser Val Ile Lys Asp
 930 935 940
 His Asn Gly Lys Leu Phe Thr Gly Lys Asp Leu Lys Asp Ala Asn Asn
 945 950 955 960
 Gly Ala Thr Val Ser Glu Asp Asp Gly Lys Asp Thr Gly Thr Gly Leu
 965 970 975

Val Thr Ala Lys Thr Val Ile Asp Ala Val Asn Lys Ser Gly Trp Arg
 980 985 990
 Val Thr Gly Glu Gly Ala Thr Ala Glu Thr Gly Ala Thr Ala Val Asn
 995 1000 1005
 Ala Gly Asn Ala Glu Thr Val Thr Ser Gly Thr Ser Val Asn Phe Lys
 1010 1015 1020
 Asn Gly Asn Ala Thr Thr Ala Thr Val Ser Lys Asp Asn Gly Asn Ile
 1025 1030 1035 1040
 Asn Val Lys Tyr Asp Val Asn Val Gly Asp Gly Leu Lys Ile Gly Asp
 1045 1050 1055
 Asp Lys Lys Ile Val Ala Asp Thr Thr Thr Leu Thr Val Thr Gly Gly
 1060 1065 1070
 Lys Val Ser Val Pro Ala Gly Ala Asn Ser Val Asn Asn Asn Lys Lys
 1075 1080 1085
 Leu Val Asn Ala Glu Gly Leu Ala Thr Ala Leu Asn Asn Leu Ser Trp
 1090 1095 1100
 Thr Ala Lys Ala Asp Lys Tyr Ala Asp Gly Glu Ser Glu Gly Glu Thr
 1105 1110 1115 1120
 Asp Gln Glu Val Lys Ala Gly Asp Lys Val Thr Phe Lys Ala Gly Lys
 1125 1130 1135
 Asn Leu Lys Val Lys Gln Ser Glu Lys Asp Phe Thr Tyr Ser Leu Gln
 1140 1145 1150
 Asp Thr Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Gly Thr Ala Asn
 1155 1160 1165
 Gly Arg Asn Asp Thr Gly Thr Val Ile Asn Lys Asp Gly Leu Thr Ile
 1170 1175 1180
 Thr Leu Ala Asn Gly Ala Ala Ala Gly Thr Asp Ala Ser Asn Gly Asn
 1185 1190 1195 1200
 Thr Ile Ser Val Thr Lys Asp Gly Ile Ser Ala Gly Asn Lys Glu Ile
 1205 1210 1215
 Thr Asn Val Lys Ser Ala Leu Lys Thr Tyr Lys Asp Thr Gln Asn Thr
 1220 1225 1230
 Ala Asp Glu Thr Gln Asp Lys Glu Phe His Ala Ala Val Lys Asn Ala
 1235 1240 1245
 Asn Glu Val Glu Phe Val Gly Lys Asn Gly Ala Thr Val Ser Ala Lys
 1250 1255 1260
 Thr Asp Asn Asn Gly Lys His Thr Val Thr Ile Asp Val Ala Glu Ala
 1265 1270 1275 1280

Lys Val Gly Asp Gly Leu Glu Lys Asp Thr Asp Gly Lys Ile Lys Leu
 1285 1290 1295
 Lys Val Asp Asn Thr Asp Gly Asn Asn Leu Leu Thr Val Asp Ala Thr
 1300 1305 1310
 Lys Gly Ala Ser Val Ala Lys Gly Glu Phe Asn Ala Val Thr Thr Asp
 1315 1320 1325
 Ala Thr Thr Ala Gln Gly Thr Asn Ala Asn Glu Arg Gly Lys Val Val
 1330 1335 1340
 Val Lys Gly Ser Asn Gly Ala Thr Ala Thr Glu Thr Asp Lys Lys Lys
 1345 1350 1355 1360
 Val Ala Thr Val Gly Asp Val Ala Lys Ala Ile Asn Asp Ala Ala Thr
 1365 1370 1375
 Phe Val Lys Val Glu Asn Asp Asp Ser Ala Thr Ile Asp Asp Ser Pro
 1380 1385 1390
 Thr Asp Asp Gly Ala Asn Asp Ala Leu Lys Ala Gly Asp Thr Leu Thr
 1395 1400 1405
 Leu Lys Ala Gly Lys Asn Leu Lys Val Lys Arg Asp Gly Lys Asn Ile
 1410 1415 1420
 Thr Phe Ala Leu Ala Asn Asp Leu Ser Val Lys Ser Ala Thr Val Ser
 1425 1430 1435 1440
 Asp Lys Leu Ser Leu Gly Thr Asn Gly Asn Lys Val Asn Ile Thr Ser
 1445 1450 1455
 Asp Thr Lys Gly Leu Lys Phe Ala Lys Asp Ser Lys Thr Gly Asp Asp
 1460 1465 1470
 Ala Asn Ile His Leu Asn Gly Ile Ala Ser Thr Leu Thr Asp Thr Leu
 1475 1480 1485
 Leu Asn Ser Gly Ala Thr Thr Asn Leu Gly Gly Asn Gly Ile Thr Asp
 1490 1495 1500
 Asn Glu Lys Lys Arg Ala Ala Ser Val Lys Asp Val Leu Asn Ala Gly
 1505 1510 1515 1520
 Trp Asn Val Arg Gly Val Lys Pro Ala Ser Ala Asn Asn Gln Val Glu
 1525 1530 1535
 Asn Ile Asp Phe Val Ala Thr Tyr Asp Thr Val Asp Phe Val Ser Gly
 1540 1545 1550
 Asp Lys Asp Thr Thr Ser Val Thr Val Glu Ser Lys Asp Asn Gly Lys
 1555 1560 1565
 Arg Thr Glu Val Lys Ile Gly Ala Lys Thr Ser Val Ile Lys Asp His
 1570 1575 1580
 Asn Gly Lys Leu Phe Thr Gly Lys Glu Leu Lys Asp Ala Asn Asn Asn
 1585 1590 1595 1600

Gly Val Thr Val Thr Glu Thr Asp Gly Lys Asp Glu Gly Asn Gly Leu
 1605 1610 1615
 Val Thr Ala Lys Ala Val Ile Asp Ala Val Asn Lys Ala Gly Trp Arg
 1620 1625 1630
 Val Lys Thr Thr Gly Ala Asn Gly Gln Asn Asp Asp Phe Ala Thr Val
 1635 1640 1645
 Ala Ser Gly Thr Asn Val Thr Phe Ala Asp Gly Asn Gly Thr Thr Ala
 1650 1655 1660
 Glu Val Thr Lys Ala Asn Asp Gly Ser Ile Thr Val Lys Tyr Asn Val
 1665 1670 1675 1680
 Lys Val Ala Asp Gly Leu Lys Leu Asp Gly Asp Lys Ile Val Ala Asp
 1685 1690 1695
 Thr Thr Val Leu Thr Val Ala Asp Gly Lys Val Thr Ala Pro Asn Asn
 1700 1705 1710
 Gly Asp Gly Lys Lys Phe Val Asp Ala Ser Gly Leu Ala Asp Ala Leu
 1715 1720 1725
 Asn Lys Leu Ser Trp Thr Ala Thr Ala Gly Lys Glu Gly Thr Gly Glu
 1730 1735 1740
 Val Asp Pro Ala Asn Ser Ala Gly Gln Glu Val Lys Ala Gly Asp Lys
 1745 1750 1755 1760
 Val Thr Phe Lys Ala Gly Asp Asn Leu Lys Ile Lys Gln Ser Gly Lys
 1765 1770 1775
 Asp Phe Thr Tyr Ser Leu Lys Lys Glu Leu Lys Asp Leu Thr Ser Val
 1780 1785 1790
 Glu Phe Lys Asp Ala Asn Gly Gly Thr Gly Ser Glu Ser Thr Lys Ile
 1795 1800 1805
 Thr Lys Asp Gly Leu Thr Ile Thr Pro Ala Asn Gly Ala Gly Ala Ala
 1810 1815 1820
 Gly Ala Asn Thr Ala Asn Thr Ile Ser Val Thr Lys Asp Gly Ile Ser
 1825 1830 1835 1840
 Ala Gly Asn Lys Ala Val Thr Asn Val Val Ser Gly Leu Lys Lys Phe
 1845 1850 1855
 Gly Asp Gly His Thr Leu Ala Asn Gly Thr Val Ala Asp Phe Glu Lys
 1860 1865 1870
 His Tyr Asp Asn Ala Tyr Lys Asp Leu Thr Asn Leu Asp Glu Lys Gly
 1875 1880 1885
 Ala Asp Asn Asn Pro Thr Val Ala Asp Asn Thr Ala Ala Thr Val Gly
 1890 1895 1900

Asp Leu Arg Gly Leu Gly Trp Val Ile Ser Ala Asp Lys Thr Thr Gly
 1905 1910 1915 1920
 Glu Pro Asn Gln Glu Tyr Asn Ala Gln Val Arg Asn Ala Asn Glu Val
 1925 1930 1935
 Lys Phe Lys Ser Gly Asn Gly Ile Asn Val Ser Gly Lys Thr Leu Asn
 1940 1945 1950
 Gly Thr Arg Val Ile Thr Phe Glu Leu Ala Lys Gly Glu Val Val Lys
 1955 1960 1965
 Ser Asn Glu Phe Thr Val Lys Asn Ala Asp Gly Ser Glu Thr Asn Leu
 1970 1975 1980
 Val Lys Val Gly Asp Met Tyr Tyr Ser Lys Glu Asp Ile Asp Pro Ala
 1985 1990 1995 2000
 Thr Ser Lys Pro Met Thr Gly Lys Thr Glu Lys Tyr Lys Val Glu Asn
 2005 2010 2015
 Gly Lys Val Val Ser Ala Asn Gly Ser Lys Thr Glu Val Thr Leu Thr
 2020 2025 2030
 Asn Lys Gly Ser Gly Tyr Val Thr Gly Asn Gln Val Ala Asp Ala Ile
 2035 2040 2045
 Ala Lys Ser Gly Phe Glu Leu Gly Leu Ala Asp Ala Glu Ala Glu
 2050 2055 2060
 Lys Ala Phe Ala Glu Ser Ala Lys Asp Lys Gln Leu Ser Lys Asp Lys
 2065 2070 2075 2080
 Ala Glu Thr Val Asn Ala His Asp Lys Val Arg Phe Ala Asn Gly Leu
 2085 2090 2095
 Asn Thr Lys Val Ser Ala Ala Thr Val Glu Ser Thr Asp Ala Asn Gly
 2100 2105 2110
 Asp Lys Val Thr Thr Thr Phe Val Lys Thr Asp Val Glu Leu Pro Leu
 2115 2120 2125
 Thr Gln Ile Tyr Asn Thr Asp Ala Asn Gly Asn Lys Ile Val Lys Lys
 2130 2135 2140
 Ala Asp Gly Lys Trp Tyr Glu Leu Asn Ala Asp Gly Thr Ala Ser Asn
 2145 2150 2155 2160
 Lys Glu Val Thr Leu Gly Asn Val Asp Ala Asn Gly Lys Lys Val Val
 2165 2170 2175
 Lys Val Thr Glu Asn Gly Ala Asp Lys Trp Tyr Tyr Thr Asn Ala Asp
 2180 2185 2190
 Gly Ala Ala Asp Lys Thr Lys Gly Glu Val Ser Asn Asp Lys Val Ser
 2195 2200 2205
 Thr Asp Glu Lys His Val Val Arg Leu Asp Pro Asn Asn Gln Ser Asn
 2210 2215 2220

Gly Lys Gly Val Val Ile Asp Asn Val Ala Asn Gly Glu Ile Ser Ala
2225 2230 2235 2240

Thr Ser Thr Asp Ala Ile Asn Gly Ser Gln Leu Tyr Ala Val Ala Lys
2245 2250 2255

Gly Val Thr Asn Leu Ala Gly Gln Val Asn Asn Leu Glu Gly Lys Val
2260 2265 2270

Asn Lys Val Gly Lys Arg Ala Asp Ala Gly Thr Ala Ser Ala Leu Ala
2275 2280 2285

Ala Ser Gln Leu Pro Gln Ala Thr Met Pro Gly Lys Ser Met Val Ala
2290 2295 2300

Ile Ala Gly Ser Ser Tyr Gln Gly Gln Asn Gly Leu Ala Ile Gly Val
2305 2310 2315 2320

Ser Arg Ile Ser Asp Asn Gly Lys Val Ile Ile Arg Leu Ser Gly Thr
2325 2330 2335

Thr Asn Ser Gln Gly Lys Thr Gly Val Ala Ala Gly Val Gly Tyr Gln
2340 2345 2350

Trp

<210> 48

<211> 2047

<212> PRT

<213> Haemophilus influenzae

<400> 48

Met Asn His Ile Tyr Lys Val Ile Phe Asn Lys Ala Thr Gly Thr Phe
1 5 10 15

Met Ala Val Ala Glu Tyr Ala Lys Ser His Ser Thr Gly Gly Gly Ser
20 25 30

Cys Ala Thr Gly Gln Val Gly Ser Val Cys Thr Leu Ser Phe Ala Arg
35 40 45

Ile Ala Ala Leu Ala Val Leu Val Ile Gly Ala Thr Leu Ser Gly Ser
50 55 60

Ala Tyr Ala Gln Lys Lys Asp Thr Lys His Ile Ala Ile Gly Glu Gln
65 70 75 80

Asn Gln Pro Arg Arg Ser Gly Thr Ala Lys Ala Asp Gly Asp Arg Ala
85 90 95

Ile Ala Ile Gly Glu Asn Ala Asn Ala Gln Gly Gly Gln Ala Ile Ala
100 105 110

Ile Gly Ser Ser Asn Lys Thr Val Asn Gly Ser Ser Leu Asp Lys Ile
115 120 125

Gly Thr Asp Ala Thr Gly Gln Glu Ser Ile Ala Ile Gly Gly Asp Val
 130 135 140
 Lys Ala Ser Gly Asp Ala Ser Ile Ala Ile Gly Ser Asp Asp Leu His
 145 150 155 160
 Leu Leu Asp Gln His Gly Asn Pro Lys His Pro Lys Gly Thr Leu Ile
 165 170 175
 Asn Asp Leu Ile Asn Gly His Ala Val Leu Lys Glu Ile Arg Ser Ser
 180 185 190
 Lys Asp Asn Asp Val Lys Tyr Arg Arg Thr Thr Ala Ser Gly His Ala
 195 200 205
 Ser Thr Ala Val Gly Ala Met Ser Tyr Ala Gln Gly His Phe Ser Asn
 210 215 220
 Ala Phe Gly Thr Arg Ala Thr Ala Lys Ser Ala Tyr Ser Leu Ala Val
 225 230 235 240
 Gly Leu Ala Ala Thr Ala Glu Gly Gln Ser Thr Ile Ala Ile Gly Ser
 245 250 255
 Asp Ala Thr Ser Ser Ser Leu Gly Ala Ile Ala Leu Gly Ala Gly Thr
 260 265 270
 Arg Ala Gln Leu Gln Gly Ser Ile Ala Leu Gly Gln Gly Ser Val Val
 275 280 285
 Thr Gln Ser Asp Asn Asn Ser Arg Pro Ala Tyr Thr Pro Asn Thr Gln
 290 295 300
 Ala Leu Asp Pro Lys Phe Gln Ala Thr Asn Asn Thr Lys Ala Gly Pro
 305 310 315 320
 Leu Ser Ile Gly Ser Asn Ser Ile Lys Arg Lys Ile Ile Asn Val Gly
 325 330 335
 Ala Gly Val Asn Lys Thr Asp Ala Val Asn Val Ala Gln Leu Glu Ala
 340 345 350
 Val Val Lys Trp Ala Lys Glu Arg Arg Ile Thr Phe Gln Gly Asp Asp
 355 360 365
 Asn Ser Thr Asp Val Lys Ile Gly Leu Asp Asn Thr Leu Thr Ile Lys
 370 375 380
 Gly Gly Ala Glu Thr Asn Ala Leu Thr Asp Asn Asn Ile Gly Val Val
 385 390 395 400
 Lys Glu Ala Asp Asn Ser Gly Leu Lys Val Lys Leu Ala Lys Thr Leu
 405 410 415
 Asn Asn Leu Thr Glu Val Asn Thr Thr Thr Leu Asn Ala Thr Thr Thr
 420 425 430
 Val Lys Val Gly Ser Ser Ser Ser Thr Thr Ala Glu Leu Leu Ser Asp
 435 440 445

Ser Leu Thr Phe Thr Gln Pro Asn Thr Gly Ser Gln Ser Thr Ser Lys
 450 455 460
 Thr Val Tyr Gly Val Asn Gly Val Lys Phe Thr Asn Asn Ala Glu Thr
 465 470 475 480
 Thr Ala Ala Ile Gly Thr Thr Arg Ile Thr Arg Asp Lys Ile Gly Phe
 485 490 495
 Ala Arg Asp Gly Asp Val Asp Glu Lys Gln Ala Pro Tyr Leu Asp Lys
 500 505 510
 Lys Gln Leu Lys Val Gly Ser Val Ala Ile Thr Ile Asp Asn Gly Ile
 515 520 525
 Asp Ala Gly Asn Lys Lys Ile Ser Asn Leu Ala Lys Gly Ser Ser Ala
 530 535 540
 Asn Asp Ala Val Thr Ile Glu Gln Leu Lys Ala Ala Lys Pro Thr Leu
 545 550 555 560
 Asn Ala Gly Ala Gly Ile Ser Val Thr Pro Thr Glu Ile Ser Val Asp
 565 570 575
 Ala Lys Ser Gly Asn Val Thr Ala Pro Thr Tyr Asn Ile Gly Val Lys
 580 585 590
 Thr Thr Glu Leu Asn Ser Asp Gly Thr Ser Asp Lys Phe Ser Val Lys
 595 600 605
 Gly Ser Gly Thr Asn Asn Ser Leu Val Thr Ala Glu His Leu Ala Ser
 610 615 620
 Tyr Leu Asn Glu Val Asn Arg Thr Ala Asp Ser Ala Leu Gln Ser Phe
 625 630 635 640
 Thr Val Lys Glu Glu Asp Asp Asp Asp Ala Asn Ala Ile Thr Val Ala
 645 650 655
 Lys Asp Thr Thr Lys Asn Ala Gly Ala Val Ser Ile Leu Lys Leu Lys
 660 665 670
 Gly Lys Asn Gly Leu Thr Val Ala Thr Lys Lys Asp Gly Thr Val Thr
 675 680 685
 Phe Gly Leu Ser Gln Asp Ser Gly Leu Thr Ile Gly Lys Ser Thr Leu
 690 695 700
 Asn Asn Asp Gly Leu Thr Val Lys Asp Thr Asn Glu Gln Ile Gln Val
 705 710 715 720
 Gly Ala Asn Gly Ile Lys Phe Thr Asn Val Asn Gly Ser Asn Pro Gly
 725 730 735
 Thr Gly Ile Ala Asn Thr Ala Arg Ile Thr Arg Asp Lys Ile Gly Phe
 740 745 750

Ala Gly Ser Asp Gly Ala Val Asp Thr Asn Lys Pro Tyr Leu Asp Gln
 755 760 765
 Asp Lys Leu Gln Val Gly Asn Val Lys Ile Thr Asn Thr Gly Ile Asn
 770 775 780
 Ala Gly Gly Lys Ala Ile Thr Gly Leu Ser Pro Thr Leu Pro Ser Ile
 785 790 795 800
 Ala Asp Gln Ser Ser Arg Asn Ile Glu Leu Gly Asn Thr Ile Gln Asp
 805 810 815
 Lys Asp Lys Ser Asn Ala Ala Ser Ile Asn Asp Ile Leu Asn Thr Gly
 820 825 830
 Phe Asn Leu Lys Asn Asn Asn Asn Pro Ile Asp Phe Val Ser Thr Tyr
 835 840 845
 Asp Ile Val Asp Phe Ala Asn Gly Asn Ala Thr Thr Ala Thr Val Thr
 850 855 860
 His Asp Thr Ala Asn Lys Thr Ser Lys Val Val Tyr Asp Val Asn Val
 865 870 875 880
 Asp Asp Thr Thr Ile His Leu Thr Gly Thr Asp Asp Asn Lys Lys Leu
 885 890 895
 Gly Val Lys Thr Thr Lys Leu Asn Lys Thr Ser Ala Asn Gly Asn Thr
 900 905 910
 Ala Thr Asn Phe Asn Val Asn Ser Ser Asp Glu Asp Ala Leu Val Asn
 915 920 925
 Ala Lys Asp Ile Ala Glu Asn Leu Asn Thr Leu Ala Lys Glu Ile His
 930 935 940
 Thr Thr Lys Gly Thr Ala Asp Thr Ala Leu Gln Thr Phe Thr Val Lys
 945 950 955 960
 Lys Val Asp Glu Asn Asn Asn Ala Asp Asp Ala Asn Ala Ile Thr Val
 965 970 975
 Gly Gln Lys Asn Ala Asn Asn Gln Val Asn Thr Leu Thr Leu Lys Gly
 980 985 990
 Glu Asn Gly Leu Asn Ile Lys Thr Asp Lys Asn Gly Thr Val Thr Phe
 995 1000 1005
 Gly Ile Asn Thr Thr Ser Gly Leu Lys Ala Gly Lys Ser Thr Leu Asn
 1010 1015 1020
 Asp Gly Gly Leu Ser Ile Lys Asn Pro Thr Gly Ser Glu Gln Ile Gln
 1025 1030 1035 1040
 Val Gly Ala Asp Gly Val Lys Phe Ala Lys Val Asn Asn Asn Gly Val
 1045 1050 1055
 Val Gly Ala Gly Ile Asp Gly Thr Thr Arg Ile Thr Arg Asp Glu Ile
 1060 1065 1070

Gly Phe Thr Gly Thr Asn Gly Ser Leu Asp Lys Ser Lys Pro His Leu
 1075 1080 1085
 Ser Lys Asp Gly Ile Asn Ala Gly Gly Lys Lys Ile Thr Asn Ile Gln
 1090 1095 1100
 Ser Gly Glu Ile Ala Gln Asn Ser His Asp Ala Val Thr Gly Gly Lys
 1105 1110 1115 1120
 Ile Tyr Asp Leu Lys Thr Glu Leu Glu Asn Lys Ile Ser Ser Thr Ala
 1125 1130 1135
 Lys Thr Ala Gln Asn Ser Leu His Glu Phe Ser Val Ala Asp Glu Gln
 1140 1145 1150
 Gly Asn Asn Phe Thr Val Ser Asn Pro Tyr Ser Ser Tyr Asp Thr Ser
 1155 1160 1165
 Lys Thr Ser Asp Val Ile Thr Phe Ala Gly Glu Asn Gly Ile Thr Thr
 1170 1175 1180
 Lys Val Asn Lys Gly Val Val Arg Val Gly Ile Asp Gln Thr Lys Gly
 1185 1190 1195 1200
 Leu Thr Thr Pro Lys Leu Thr Val Gly Asn Asn Asn Gly Lys Gly Ile
 1205 1210 1215
 Val Ile Asp Ser Gln Asn Gly Gln Asn Thr Ile Thr Gly Leu Ser Asn
 1220 1225 1230
 Thr Leu Ala Asn Val Thr Asn Asp Lys Gly Ser Val Arg Thr Thr Glu
 1235 1240 1245
 Gln Gly Asn Ile Ile Lys Asp Glu Asp Lys Thr Arg Ala Ala Ser Ile
 1250 1255 1260
 Val Asp Val Leu Ser Ala Gly Phe Asn Leu Gln Gly Asn Gly Glu Ala
 1265 1270 1275 1280
 Val Asp Phe Val Ser Thr Tyr Asp Thr Val Asn Phe Ala Asp Gly Asn
 1285 1290 1295
 Ala Thr Thr Ala Lys Val Thr Tyr Asp Asp Thr Ser Lys Thr Ser Lys
 1300 1305 1310
 Val Val Tyr Asp Val Asn Val Asp Asp Thr Thr Ile Glu Val Lys Asp
 1315 1320 1325
 Lys Lys Leu Gly Val Lys Thr Thr Thr Leu Thr Ser Thr Gly Thr Gly
 1330 1335 1340
 Ala Asn Lys Phe Ala Leu Ser Asn Gln Ala Thr Gly Asp Ala Leu Val
 1345 1350 1355 1360
 Lys Ala Ser Asp Ile Val Ala His Leu Asn Thr Leu Ser Gly Asp Ile
 1365 1370 1375

Gln Thr Ala Lys Gly Ala Ser Gln Ala Asn Asn Ser Ala Gly Tyr Val
 1380 1385 1390
 Asp Ala Asp Gly Asn Lys Val Ile Tyr Asp Ser Thr Asp Asn Lys Tyr
 1395 1400 1405
 Tyr Gln Ala Lys Asn Asp Gly Thr Val Asp Lys Thr Lys Glu Val Ala
 1410 1415 1420
 Lys Asp Lys Leu Val Ala Gln Ala Gln Thr Pro Asp Gly Thr Leu Ala
 1425 1430 1435 1440
 Gln Met Asn Val Lys Ser Val Ile Asn Lys Glu Gln Val Asn Asp Ala
 1445 1450 1455
 Asn Lys Lys Gln Gly Ile Asn Glu Asp Asn Ala Phe Val Lys Gly Leu
 1460 1465 1470
 Glu Lys Ala Ala Ser Asp Asn Lys Thr Lys Asn Ala Ala Val Thr Val
 1475 1480 1485
 Gly Asp Leu Asn Ala Val Ala Gln Thr Pro Leu Thr Phe Ala Gly Asp
 1490 1495 1500
 Thr Gly Thr Thr Ala Lys Lys Leu Gly Glu Thr Leu Thr Ile Lys Gly
 1505 1510 1515 1520
 Gly Gln Thr Asp Thr Asn Lys Leu Thr Asp Asn Asn Ile Gly Val Val
 1525 1530 1535
 Ala Gly Thr Asp Gly Phe Thr Val Lys Leu Ala Lys Asp Leu Thr Asn
 1540 1545 1550
 Leu Asn Ser Val Asn Ala Gly Gly Thr Lys Ile Asp Asp Lys Gly Val
 1555 1560 1565
 Ser Phe Val Asp Ser Ser Gly Gln Ala Lys Ala Asn Thr Pro Val Leu
 1570 1575 1580
 Ser Ala Asn Gly Leu Asp Leu Gly Gly Lys Val Ile Ser Asn Val Gly
 1585 1590 1595 1600
 Lys Gly Thr Lys Asp Thr Asp Ala Ala Asn Val Gln Gln Leu Asn Glu
 1605 1610 1615
 Val Arg Asn Leu Leu Gly Leu Gly Asn Ala Gly Asn Asp Asn Ala Asp
 1620 1625 1630
 Gly Asn Gln Val Asn Ile Ala Asp Ile Lys Lys Asp Pro Asn Ser Gly
 1635 1640 1645
 Ser Ser Ser Asn Arg Thr Val Ile Lys Ala Gly Thr Val Leu Gly Gly
 1650 1655 1660
 Lys Gly Asn Asn Asp Thr Glu Lys Leu Ala Thr Gly Gly Ile Gln Val
 1665 1670 1675 1680
 Gly Val Asp Lys Asp Gly Asn Ala Asn Gly Asp Leu Ser Asn Val Trp
 1685 1690 1695

Val Lys Thr Gln Lys Asp Gly Ser Lys Lys Ala Leu Leu Ala Thr Tyr
 1700 1705 1710
 Asn Ala Ala Gly Gln Thr Asn Tyr Leu Thr Asn Asn Pro Ala Glu Ala
 1715 1720 1725
 Ile Asp Arg Ile Asn Glu Gln Gly Ile Arg Phe Phe His Val Asn Asp
 1730 1735 1740
 Gly Asn Gln Glu Pro Val Val Gln Gly Arg Asn Gly Ile Asp Ser Ser
 1745 1750 1755 1760
 Ala Ser Gly Lys His Ser Val Ala Ile Gly Phe Gln Ala Lys Ala Asp
 1765 1770 1775
 Gly Glu Ala Ala Val Ala Ile Gly Arg Gln Thr Gln Ala Gly Asn Gln
 1780 1785 1790
 Ser Ile Ala Ile Gly Asp Asn Ala Gln Ala Thr Gly Asp Gln Ser Ile
 1795 1800 1805
 Ala Ile Gly Thr Gly Asn Val Val Ala Gly Lys His Ser Gly Ala Ile
 1810 1815 1820
 Gly Asp Pro Ser Thr Val Lys Ala Asp Asn Ser Tyr Ser Val Gly Asn
 1825 1830 1835 1840
 Asn Asn Gln Phe Thr Asp Ala Thr Gln Thr Asp Val Phe Gly Val Gly
 1845 1850 1855
 Asn Asn Ile Thr Val Thr Glu Ser Asn Ser Val Ala Leu Gly Ser Asn
 1860 1865 1870
 Ser Ala Ile Ser Ala Gly Thr His Ala Gly Thr Gln Ala Lys Lys Ser
 1875 1880 1885
 Asp Gly Thr Ala Gly Thr Thr Thr Thr Ala Gly Ala Thr Gly Thr Val
 1890 1895 1900
 Lys Gly Phe Ala Gly Gln Thr Ala Val Gly Ala Val Ser Val Gly Ala
 1905 1910 1915 1920
 Ser Gly Ala Glu Arg Arg Ile Gln Asn Val Ala Ala Gly Glu Val Ser
 1925 1930 1935
 Ala Thr Ser Thr Asp Ala Val Asn Gly Ser Gln Leu Tyr Lys Ala Thr
 1940 1945 1950
 Gln Ser Ile Ala Asn Ala Thr Asn Glu Leu Asp His Arg Ile His Gln
 1955 1960 1965
 Asn Glu Asn Lys Ala Asn Ala Gly Ile Ser Ser Ala Met Ala Met Ala
 1970 1975 1980
 Ser Met Pro Gln Ala Tyr Ile Pro Gly Arg Ser Met Val Thr Gly Gly
 1985 1990 1995 2000

Ile Ala Thr His Asn Gly Gln Gly Ala Val Ala Val Gly Leu Ser Lys
 2005 2010 2015

Leu Ser Asp Asn Gly Gln Trp Val Phe Lys Ile Asn Gly Ser Ala Asp
 2020 2025 2030

Thr Gln Gly His Val Gly Ala Ala Val Gly Ala Gly Phe His Phe
 2035 2040 2045

<210> 49

<211> 2314

<212> PRT

<213> Haemophilus influenzae

<400> 49

Met Asn His Ile Tyr Lys Val Ile Phe Asn Lys Ala Thr Gly Thr Phe
 1 5 10 15

Met Ala Val Ala Glu Cys Ala Lys Ser His Ser Gly Gly Ser Ser Ser
 20 25 30

Ser Thr Ala Gly Gln Val Gly Ser Ser Pro Val Ile Arg Leu Thr Arg
 35 40 45

Val Ala Thr Leu Ala Ile Leu Val Ile Gly Ala Thr Leu Asn Gly Ser
 50 55 60

Ala Tyr Ala Gln Asn Asn Ser Lys Ile Ala Phe Gly Thr Thr Gly Asn
 65 70 75 80

Asn Asp Asn Ala Ser Ala Ser Asn Glu Ala Ser Ile Ala Ile Gly Ser
 85 90 95

Leu Ala Lys Ala His Ala Asn Gln Ala Ile Ala Ile Gly Gly Ser Lys
 100 105 110

Pro Asp Pro Arg Asn Gln Ala Ala Asn Gln Lys Ala Gly Ser His Ala
 115 120 125

Lys Gly Lys Glu Ser Ile Ala Ile Gly Gly Asp Val Leu Ala Glu Gly
 130 135 140

Asp Ala Ser Ile Ala Ile Gly Ser Asp Asp Leu Tyr Leu Asp Arg Asn
 145 150 155 160

Ser Thr Asn Ser Lys Tyr Pro Asn Gly Leu Leu Ser Thr Leu Ile Gln
 165 170 175

Asn His Thr Val Leu Arg Gln Ile Arg Asp Ser Asn Gly Ser Gln Lys
 180 185 190

Tyr Arg Arg Thr Ala Ala Glu Gly His Ala Ser Thr Ala Val Gly Ala
 195 200 205

Met Ala Tyr Ala Lys Gly His Phe Ala Asn Ala Phe Gly Thr Arg Ser
 210 215 220

Thr Ala Glu Gly Asn Tyr Ser Leu Ala Val Gly Leu Thr Ala Lys Ala
 225 230 235 240
 Glu Lys Gly Tyr Thr Ile Ala Ile Gly Ser Asn Ala Gln Ala Ile Asn
 245 250 255
 Tyr Gly Ala Leu Ala Leu Gly Ala Asp Thr Arg Val Asp Leu Asp Tyr
 260 265 270
 Gly Ile Ala Leu Gly Tyr Gly Ser Gln Ile Leu Asn Asn Asn Asn Asn
 275 280 285
 Asn Asn Asn Lys Ala Tyr Val Pro Glu Gly Asn Gly Ser Asn Ile Lys
 290 295 300
 Ser Ser Lys Ala Thr Gly Asn Gly Leu Phe Ser Ile Gly Ser Ser Thr
 305 310 315 320
 Ile Lys Arg Lys Ile Ile Asn Val Gly Ala Gly Tyr Glu Asp Thr Asp
 325 330 335
 Ala Val Asn Val Ala Gln Leu Lys Ala Val Glu Asn Leu Ala Lys Arg
 340 345 350
 Gln Ile Thr Phe Lys Gly Asp Asp Asn Gly Thr Gly Val Lys Lys Lys
 355 360 365
 Leu Gly Glu Thr Leu Thr Ile Lys Gly Gly Glu Thr Gln Ala Asp Lys
 370 375 380
 Leu Thr Asp Asn Asn Asn Ile Gly Val Val Thr Asp Asn Asn Thr Gly
 385 390 395 400
 Leu Lys Val Lys Leu Ala Lys Asn Leu Ser Gly Leu Glu Thr Val Ser
 405 410 415
 Thr Lys Asn Leu Thr Ala Ser Glu Lys Val Thr Val Gly Ser Gly Asn
 420 425 430
 Asn Thr Ala Glu Leu Gln Ser Gly Gly Leu Thr Phe Thr Pro Thr Thr
 435 440 445
 Asn Ala Ser Thr Asp Lys Thr Val Tyr Gly Thr Asp Gly Leu Lys Phe
 450 455 460
 Thr Asp Asn Ser Asn Thr Ala Leu Glu Asp Thr Thr Arg Ile Thr Lys
 465 470 475 480
 Asp Lys Ile Gly Phe Ser Asn Lys Ala Gly Thr Val Asp Glu Asn Lys
 485 490 495
 Pro Tyr Leu Asp Lys Asp Lys Leu Lys Val Gly Asn Ser Thr Leu Asn
 500 505 510
 Asn Gly Gly Leu Thr Val Asn Asn Thr Ile Gly Gly Ser Asn Lys Gln
 515 520 525
 Ile Gln Val Gly Ala Asp Gly Ile Lys Phe Ala Asp Val Asn Val Asn
 530 535 540

```

Val Ser Asn Ala Ala Lys Phe Gly Thr Thr Arg Ile Thr Glu Glu Glu
545                      550                      555                      560

Ile Gly Phe Ala Asp Ala Asp Gly Lys Val Asp Lys Lys Ser Pro Tyr
565                      570                      575

Leu Asp Lys Lys Gln Leu Gln Val Gly Gly Val Lys Ile Thr Lys Asp
580                      585                      590

Ser Gly Ile Asn Ala Gly Asp Gln Lys Ile Ser Asn Val Lys Asp Ala
595                      600                      605

Thr Asp Asp Thr Asp Ala Val Thr Tyr Lys Gln Leu Lys Gln Val Gln
610                      615                      620

Gln Asp Ala Asp Gly Ala Leu Gln Ser Phe Ser Ile Arg Asp Glu Lys
625                      630                      635                      640

Gly Gln Glu Phe Thr Ile Ser Asn Leu Tyr Ser Asn Gly Asn Thr Pro
645                      650                      655

Asn Thr Phe Glu Thr Ile Thr Phe Ala Gly Glu Asn Gly Ile Ser Ile
660                      665                      670

Ser Asn Asp Ile Ala Lys Gly Lys Val Lys Val Gly Ile Asp Pro Ile
675                      680                      685

Asn Gly Leu Thr Thr Pro Lys Leu Thr Val Gly Ser Asp Lys Asp Gly
690                      695                      700

Lys Thr Gln Leu Val Ile Glu Gln Val Ala Ser Gly Asn Asp Thr Lys
705                      710                      715                      720

Asn Ile Thr Arg Gly Leu Ser Pro Thr Leu Pro Ser Ile Thr Asn Ala
725                      730                      735

Gly Gly Val Arg Thr Thr Glu Gln Gly Asn Thr Ile Thr Ser Asp Glu
740                      745                      750

Asp Lys Ser Lys Ala Ala Ser Ile Gly Asp Ile Leu Asn Thr Gly Phe
755                      760                      765

Asn Leu Lys Asn Asn Ser Asn Ser Val Gly Phe Val Ser Thr Tyr Asn
770                      775                      780

Thr Val Asp Phe Ile Asp Gly Asn Ala Thr Thr Ala Lys Val Thr Tyr
785                      790                      795                      800

Asp Glu Thr Asn Gln Thr Ser Lys Val Thr Tyr Asp Val Asn Val Asp
805                      810                      815

Glu Lys Thr Ile Glu Leu Thr Gly Asp Asn Gly Lys Thr Asn Lys Ile
820                      825                      830

Gly Val Lys Thr Thr Thr Leu Thr Thr Thr Asn Ala Asn Gly Lys Ala
835                      840                      845

```

Thr Asn Phe Ser Thr Thr Asp Asn Asp Ala Leu Val Asn Ala Lys Asp
 850 855 860
 Ile Ala Glu Asn Leu Asn Thr Leu Ala Lys Glu Ile His Thr Thr Lys
 865 870 875 880
 Gly Thr Ala Asp Thr Ala Leu Gln Thr Phe Lys Val Lys Lys Asp Gly
 885 890 895
 Ala Thr Asp Asp Glu Thr Ile Thr Val Gly Lys Asp Gly Thr Gln Asn
 900 905 910
 Gly Lys Thr Val Asn Thr Leu Lys Leu Lys Gly Glu Asn Gly Leu Thr
 915 920 925
 Val Ala Thr Asn Lys Asp Gly Thr Val Thr Phe Gly Ile Asn Thr Gln
 930 935 940
 Ser Gly Leu Lys Ala Gly Asp Ser Thr Thr Leu Asn Lys Asp Gly Leu
 945 950 955 960
 Ser Ile Lys Asn Pro Ala Ser Asn Glu Gln Ile Gln Val Gly Ala Asp
 965 970 975
 Gly Val Lys Phe Ala Lys Val Asp Lys Gly Asn Ser Ser Thr Gly Ile
 980 985 990
 Asp Gly Thr Ser Arg Ile Thr Lys Asp Gln Ile Gly Phe Thr Gly Ala
 995 1000 1005
 Asn Gly Ser Leu Asp Thr Thr Lys Pro His Leu Thr Lys Asp Lys Leu
 1010 1015 1020
 Lys Val Gly Glu Val Glu Ile Thr Asn Thr Gly Ile Asn Ala Gly Gly
 1025 1030 1035 1040
 Lys Lys Ile Thr Asn Ile Gln Ser Gly Asp Ile Thr Gln Asn Ser Asn
 1045 1050 1055
 Asp Ala Val Thr Gly Gly Arg Val Tyr Asp Leu Lys Thr Glu Leu Glu
 1060 1065 1070
 Ser Lys Ile Asn Ser Ala Ala Lys Thr Ala Gln Asn Ser Leu His Glu
 1075 1080 1085
 Phe Ser Val Ala Asp Glu Gln Gly Asn His Phe Thr Val Ser Asn Pro
 1090 1095 1100
 Tyr Ser Ser Tyr Asp Thr Ser Lys Thr Ser Asp Val Ile Thr Phe Ala
 1105 1110 1115 1120
 Gly Glu Asn Gly Ile Thr Thr Lys Val Asn Lys Gly Val Val Arg Val
 1125 1130 1135
 Gly Ile Asp Gln Thr Lys Gly Leu Thr Thr Pro Lys Leu Thr Val Gly
 1140 1145 1150
 Asn Asn Asn Gly Lys Gly Ile Val Ile Asp Ser Lys Asp Gly Gln Asn
 1155 1160 1165

Thr Ile Thr Gly Leu Ser Asn Thr Leu Ala Asn Val Thr Asn Asp Gly
 1170 1175 1180
 Ala Gly His Ala Leu Ser Gln Gly Leu Ala Asn Asp Thr Asp Lys Thr
 1185 1190 1195 1200
 Arg Ala Ala Ser Ile Gly Asp Val Leu Asn Ala Gly Phe Asn Leu Gln
 1205 1210 1215
 Gly Asn Gly Glu Ala Val Asp Phe Val Ser Thr Tyr Asp Thr Val Asp
 1220 1225 1230
 Phe Ile Asp Gly Asn Ala Thr Thr Ala Lys Val Thr Tyr Asp Asp Thr
 1235 1240 1245
 Ser Lys Thr Ser Lys Val Val Tyr Asp Val Asn Val Asp Asn Lys Thr
 1250 1255 1260
 Ile Glu Val Thr Ser Asp Lys Lys Leu Gly Val Lys Thr Thr Thr Leu
 1265 1270 1275 1280
 Thr Lys Thr Ser Ala Asn Gly Asn Ala Thr Lys Phe Ser Ala Ala Asp
 1285 1290 1295
 Gly Asp Ala Leu Val Lys Ala Ser Asp Ile Ala Thr His Leu Asn Thr
 1300 1305 1310
 Leu Ser Gly Asp Ile Gln Thr Ala Lys Gly Ala Ser Gln Ala Ser Ser
 1315 1320 1325
 Ser Ala Ser Tyr Val Asp Ala Asp Gly Asn Lys Val Ile Tyr Asp Ser
 1330 1335 1340
 Thr Asp Lys Lys Tyr Tyr Gln Val Asn Asp Lys Gly Gln Val Asp Lys
 1345 1350 1355 1360
 Asn Lys Glu Val Ala Lys Asp Lys Leu Val Ala Gln Ala Gln Thr Pro
 1365 1370 1375
 Asp Gly Thr Leu Ala Gln Met Asn Val Lys Ser Val Ile Asn Lys Glu
 1380 1385 1390
 Gln Val Asn Asp Ala Asn Lys Lys Gln Gly Ile Asn Glu Asp Asn Ala
 1395 1400 1405
 Phe Ile Lys Gly Leu Glu Asn Ala Ala Lys Asp Thr Lys Thr Lys Asn
 1410 1415 1420
 Ala Ala Val Thr Val Gly Asp Leu Asn Ala Val Ala Gln Thr Pro Leu
 1425 1430 1435 1440
 Thr Phe Ala Gly Asp Thr Gly Thr Thr Ala Lys Lys Leu Gly Glu Thr
 1445 1450 1455
 Leu Thr Ile Lys Gly Gly Gln Thr Asp Thr Asn Lys Leu Thr Asp Asn
 1460 1465 1470

Asn Ile Gly Val Val Ala Gly Thr Asp Gly Phe Thr Val Lys Leu Ala
 1475 1480 1485
 Lys Asp Leu Thr Asn Leu Asn Ser Val Asn Ala Gly Gly Thr Arg Ile
 1490 1495 1500
 Asp Glu Lys Gly Ile Ser Phe Val Asp Ala Asn Gly Gln Ala Lys Ala
 1505 1510 1515 1520
 Asn Thr Pro Val Leu Ser Ala Asn Gly Leu Asp Leu Gly Gly Lys Arg
 1525 1530 1535
 Ile Ser Asn Ile Gly Ala Ala Val Asp Asp Asn Asp Ala Val Asn Phe
 1540 1545 1550
 Lys Gln Phe Asn Glu Val Ala Lys Thr Val Asn Asn Leu Asn Asn Gln
 1555 1560 1565
 Ser Asn Ser Gly Ala Ser Leu Pro Phe Val Val Thr Asp Ala Asn Gly
 1570 1575 1580
 Lys Pro Ile Asn Gly Thr Asp Gly Lys Pro Gln Lys Ala Ile Lys Gly
 1585 1590 1595 1600
 Ala Asp Gly Lys Tyr Tyr His Ala Asn Ala Asn Gly Val Pro Val Asp
 1605 1610 1615
 Lys Asp Gly Lys Pro Ile Thr Asp Ala Asp Lys Leu Ala Asn Leu Ala
 1620 1625 1630
 Ala His Gly Lys Pro Leu Asp Ala Gly His Gln Val Val Ala Ser Leu
 1635 1640 1645
 Gly Gly Asn Ser Asp Ala Ile Thr Leu Thr Asn Ile Lys Ser Thr Leu
 1650 1655 1660
 Pro Gln Ile Asp Thr Pro Asn Thr Gly Asn Ala Asn Ala Gly Gln Ala
 1665 1670 1675 1680
 Gln Ser Leu Pro Ser Leu Ser Ala Ala Gln Gln Ser Asn Ala Ala Ser
 1685 1690 1695
 Val Lys Asp Val Leu Asn Val Gly Phe Asn Leu Gln Thr Asn His Asn
 1700 1705 1710
 Gln Val Asp Phe Val Lys Ala Tyr Asp Thr Val Asn Phe Val Asn Gly
 1715 1720 1725
 Thr Gly Ala Asp Ile Thr Ser Val Arg Ser Ala Asp Gly Thr Met Ser
 1730 1735 1740
 Asn Ile Thr Val Asn Thr Ala Leu Ala Ala Thr Asp Asp Asp Gly Asn
 1745 1750 1755 1760
 Val Leu Ile Lys Ala Lys Asp Gly Lys Phe Tyr Lys Ala Asp Asp Leu
 1765 1770 1775
 Met Pro Asn Gly Ser Leu Lys Ala Gly Lys Ser Ala Ser Asp Ala Lys
 1780 1785 1790

Thr Pro Thr Gly Leu Ser Leu Val Asn Pro Asn Ala Gly Lys Gly Ser
1795 1800 1805

Thr Gly Asp Ala Val Ala Leu Asn Asn Leu Ser Lys Ala Val Phe Lys
1810 1815 1820

Ser Lys Asp Gly Thr Thr Thr Thr Thr Val Ser Ser Asp Gly Ile Ser
1825 1830 1835 1840

Ile Gln Gly Lys Asp Asn Ser Ser Ile Thr Leu Ser Lys Asp Gly Leu
1845 1850 1855

Asn Val Gly Gly Lys Val Ile Ser Asn Val Gly Lys Gly Thr Lys Asp
1860 1865 1870

Thr Asp Ala Ala Asn Val Gln Gln Leu Asn Glu Val Arg Asn Leu Leu
1875 1880 1885

Gly Leu Gly Asn Ala Gly Asn Asp Asn Ala Asp Gly Asn Gln Val Asn
1890 1895 1900

Ile Ala Asp Ile Lys Lys Asp Pro Asn Ser Gly Ser Ser Ser Asn Arg
1905 1910 1915 1920

Thr Val Ile Lys Ala Gly Thr Val Leu Gly Gly Lys Gly Asn Asn Asp
1925 1930 1935

Thr Glu Lys Leu Ala Thr Gly Gly Val Gln Val Gly Val Asp Lys Asp
1940 1945 1950

Gly Asn Ala Asn Gly Asp Leu Ser Asn Val Trp Val Lys Thr Gln Lys
1955 1960 1965

Asp Gly Ser Lys Lys Ala Leu Leu Ala Thr Tyr Asn Ala Ala Gly Gln
1970 1975 1980

Thr Asn Tyr Leu Thr Asn Asn Pro Ala Glu Ala Ile Asp Arg Ile Asn
1985 1990 1995 2000

Glu Gln Gly Ile Arg Phe Phe His Val Asn Asp Gly Asn Gln Glu Pro
2005 2010 2015

Val Val Gln Gly Arg Asn Gly Ile Asp Ser Ser Ala Ser Gly Lys His
2020 2025 2030

Ser Val Ala Ile Gly Phe Gln Ala Lys Ala Asp Gly Glu Ala Ala Val
2035 2040 2045

Ala Ile Gly Arg Gln Thr Gln Ala Gly Asn Gln Ser Ile Ala Ile Gly
2050 2055 2060

Asp Asn Ala Gln Ala Thr Gly Asp Gln Ser Ile Ala Ile Gly Thr Gly
2065 2070 2075 2080

Asn Val Val Thr Gly Lys His Ser Gly Ala Ile Gly Asp Pro Ser Thr
2085 2090 2095

Val Lys Ala Asp Asn Ser Tyr Ser Val Gly Asn Asn Asn Gln Phe Ile
 2100 2105 2110

Asp Ala Thr Gln Thr Asp Val Phe Gly Val Gly Asn Asn Ile Thr Val
 2115 2120 2125

Thr Glu Ser Asn Ser Val Ala Leu Gly Ser Asn Ser Ala Ile Ser Ala
 2130 2135 2140

Gly Thr His Ala Gly Thr Gln Ala Lys Lys Ser Asp Gly Thr Ala Gly
 2145 2150 2155 2160

Thr Thr Thr Thr Ala Gly Ala Thr Gly Thr Val Lys Gly Phe Ala Gly
 2165 2170 2175

Gln Thr Ala Val Gly Ala Val Ser Val Gly Ala Ser Gly Ala Glu Arg
 2180 2185 2190

Arg Ile Gln Asn Val Ala Ala Gly Glu Val Ser Ala Thr Ser Thr Asp
 2195 2200 2205

Ala Val Asn Gly Ser Gln Leu Tyr Lys Ala Thr Gln Gly Ile Ala Asn
 2210 2215 2220

Ala Thr Asn Glu Leu Asp His Arg Ile His Gln Asn Glu Asn Lys Ala
 2225 2230 2235 2240

Asn Ala Gly Ile Ser Ser Ala Met Ala Met Ala Ser Met Pro Gln Ala
 2245 2250 2255

Tyr Ile Pro Gly Arg Ser Met Val Thr Gly Gly Ile Ala Thr His Asn
 2260 2265 2270

Gly Gln Gly Ala Val Ala Val Gly Leu Ser Lys Leu Ser Asp Asn Gly
 2275 2280 2285

Gln Trp Val Phe Lys Ile Asn Gly Ser Ala Asp Thr Gln Gly His Val
 2290 2295 2300

Gly Ala Ala Val Gly Ala Gly Phe His Phe
 2305 2310

<210> 50

<211> 163

<212> DNA

<213> Haemophilus influenzae

<400> 50

aattgcagcc cgcctaata ggcgggtttt ttgaattccc gggtaacctg agatctctag 60
 actgcagata tagttcctcc tttagcaaaa aaacctctca agaccggtt agaggcccca 120
 aggggttatg ctagttattg ctcagcgggtg gcagcagcgt gca 163

<210> 51

<211> 155

<212> DNA

<213> Haemophilus influenzae

<400> 51

cgtcgggcgg attactcgcc cgaaaaaact taagggccca tggagctcta gagatctgac 60
 gtctatatca aggaggaaag tcgttttttg gggagttctg ggcaaatctc cgggggtccc 120
 caatacgatc aataacgagt cgccaccgtc gtcgc 155

<210> 52

<211> 110

<212> DNA

<213> Haemophilus influenzae

<400> 52

tatgaacaaa atttttaacg ttatttgga tggatgact caaacttggg ctgtcgtatc 60
 tgaactcact cgcgcccaca ccaaacgtgc ctcgcaacc gtggcagccg 110

<210> 53

<211> 105

<212> DNA

<213> Haemophilus influenzae

<400> 53

acttggttta aaaattgcaa taaaccttac aatactgagt ttgaaccga cagcatagac 60
 ttgagtgaagc gcgggtgtgg tttgcacgga ggcgttggca ccgtc 105

<210> 54

<211> 36

<212> PRT

<213> Haemophilus influenzae

<400> 54

Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Met Thr Gln Thr Trp
 1 5 10 15
 Ala Val Val Ser Glu Leu Thr Arg Ala His Thr Lys Arg Ala Ser Ala
 20 25 30
 Thr Val Ala Ala
 35

<210> 55

<211> 43

<212> DNA

<213> Haemophilus influenzae

<400> 55

gggaattcat atgtccgcaa cggttgaggc gaacaacaat act 43

<210> 56

<211> 11

<212> PRT

<213> Haemophilus influenzae

<400> 56

Met Ser Ala Thr Val Glu Ala Asn Asn Asn Thr
 1 5 10

<210> 57
 <211> 45
 <212> DNA
 <213> Haemophilus influenzae

<400> 57
 cgggatccca ccaaggtctt tcgctaaagc aaaggtaatg gtgtg 45

<210> 58
 <211> 36
 <212> DNA
 <213> Haemophilus influenzae

<400> 58
 cacaccatta cctttgcttt agcgaaagac cttggt 36

<210> 59
 <211> 12
 <212> PRT
 <213> Haemophilus influenzae

<400> 59
 His Thr Ile Thr Phe Ala Leu Ala Lys Asp Leu Gly
 1 5 10

<210> 60
 <211> 40
 <212> DNA
 <213> Haemophilus influenzae

<400> 60
 ggggaattcat atggtatttg cgaccgtatt gtctgcaacg 40

<210> 61
 <211> 10
 <212> PRT
 <213> Haemophilus influenzae

<400> 61
 Met Val Leu Ala Thr Val Leu Ser Ala Thr
 1 5 10

Patentansprüche

1. Ein isoliertes und gereinigtes Nukleinsäuremolekül, welches ein N-trunkiertes Haemophilus influenzae Adhäsin(Hia)-Protein eines Stammes von Haemophilus influenzae codiert, das als Einschlusskörper exprimiert wird, wobei das N-trunkierte Protein die Fähigkeit aufweist, gegen eine Besiedlung zu schützen, und dieses das N-trunkierte Haemophilus influenzae Adhäsin(Hia)-Protein beginnend bei der Aminosäure V38 von SEQ ID NO: 24 (Haemophilus influenzae Stamm 33) ist.

2. Ein Vektor zum Transformieren eines Wirtes, umfassend das Nukleinsäuremolekül von Anspruch 1.

3. Ein Plasmidvektor gemäß Anspruch 2.

4. Ein Vektor zum Transformieren eines Wirtes, welcher ein Nukleinsäuremolekül, welches ein N-trunkiertes Haemophilus influenzae Adhäsin(Hia)-Protein wie in Anspruch 1 beansprucht codiert, und einen Promoter, der operativ mit dem Nukleinsäuremolekül verbunden ist, zur Expression des trunkierten Hia-Proteins umfasst.

5. Der Vektor, der in Anspruch 4 beansprucht wird, welcher weiterhin das cer-Gen von E. coli umfasst.

6. Der Vektor, der in Anspruch 4 oder 5 beansprucht wird, wobei der Vektor die Identifizierungscharakteristika eines Plasmidvektors aufweist, der ausgewählt wird aus der Gruppe, bestehend aus:

DS-2340-2-3, enthaltend eine T7-Hia-Genkassette, welche das trunkierte Hia-Gen von Anspruch 1 umfasst, wie in [Fig. 8A](#) gezeigt, und
DS-2448-17, enthaltend Tandem-T7-Hia-Genkassetten, welche das trunkierte Hia-Gen von Anspruch 1 umfassen, wie in [Fig. 9B](#) gezeigt.

7. Eine Wirtszelle, die mit einem Vektor wie in irgendeinem der Ansprüche 2 bis 6 beansprucht transformiert ist, und welche ein schützendes N-trunkiertes Haemophilus influenzae Adhäsion(Hia)-Protein beginnend bei der Aminosäure V38 der SEQ ID NO: 24 (nicht typisierbarer Stamm 33 von Haemophilus influenzae) exprimiert.

8. Ein rekombinantes schützendes N-trunkiertes Hia-Protein beginnend bei der Aminosäure V38 der SEQ ID NO: 24 Haemophilus influenzae Stamm 33).

9. Eine immunogene Zusammensetzung, umfassend wenigstens eine immunologisch aktive Komponente, die ausgewählt ist aus der Gruppe, bestehend aus:

(A) einem isolierten und gereinigten Nukleinsäuremolekül gemäß Anspruch 1;

(B) einem rekombinanten Protein gemäß Anspruch 8; und

einen pharmazeutisch verträglichen Träger dafür.

10. Ein in vitro-Verfahren zur Erzeugung eines schützenden N-trunkierten Haemophilus influenzae Adhäsion(Hia)-Proteins beginnend bei der Aminosäure V38 der SEQ ID NO: 24 (nicht typisierbarer Stamm 33 von Haemophilus influenzae), welches umfasst:

das Transformieren einer Wirtszelle mit einem Vektor wie in Anspruch 2 beansprucht,
das Wachsen lassen der Wirtszelle, um das codierte trunkierte Hia zu exprimieren, und
das Isolieren und Reinigen des exprimierten Hia-Proteins.

11. Das Verfahren, das in Anspruch 10 beansprucht wird, wobei die Wirtszelle E. coli ist.

12. Das Verfahren, das in Anspruch 10 oder 11 beansprucht wird, wobei das codierte trunkierte Hia in Einschlusskörpern exprimiert wird.

13. Das Verfahren, das in Anspruch 12 beansprucht wird, wobei die Isolation und Reinigung des exprimierten Hia bewirkt wird durch:

das Aufbrechen der gewachsenen transformierten Zellen, um einen Überstand und die Einschlusskörper zu erzeugen,

das Solubilisieren der Einschlusskörper, um eine Lösung des rekombinanten Hia zu erzeugen,

das chromatographische Reinigen der Lösung des rekombinanten Hia, so dass es frei von Zelltrümmern ist, und

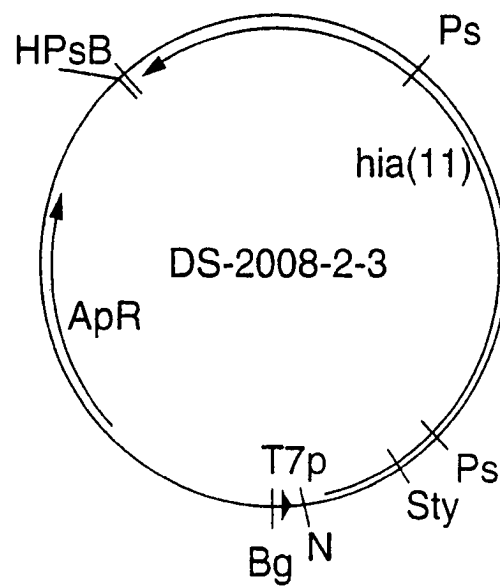
das Isolieren des gereinigten rekombinanten Hia-Proteins.

14. Das Verfahren, das in irgendeinem der Ansprüche 10 bis 13 beansprucht wird, wobei der Vektor den T7-Promoter beinhaltet und der Wirt E. coli ist, welcher in der Gegenwart einer induzierenden Menge an Lactose kultiviert wird.

15. Eine Vakzine zur in vivo-Verabreichung, um gegen eine Krankheit zu schützen, die durch Haemophilus verursacht wird, umfassend eine immunogene Zusammensetzung gemäß Anspruch 9.

Es folgen 204 Blatt Zeichnungen

Restriktionskarte von DS-2008-2-3, pT7 hia (11).



pT7 hia (11)

FIG.1A

FIG.1B

Oligonukleotide, die verwendet wurden, um das *hla*-Gen des Stamms 11 in voller Länge für Expressionsstudien mit PCR zu amplifizieren.

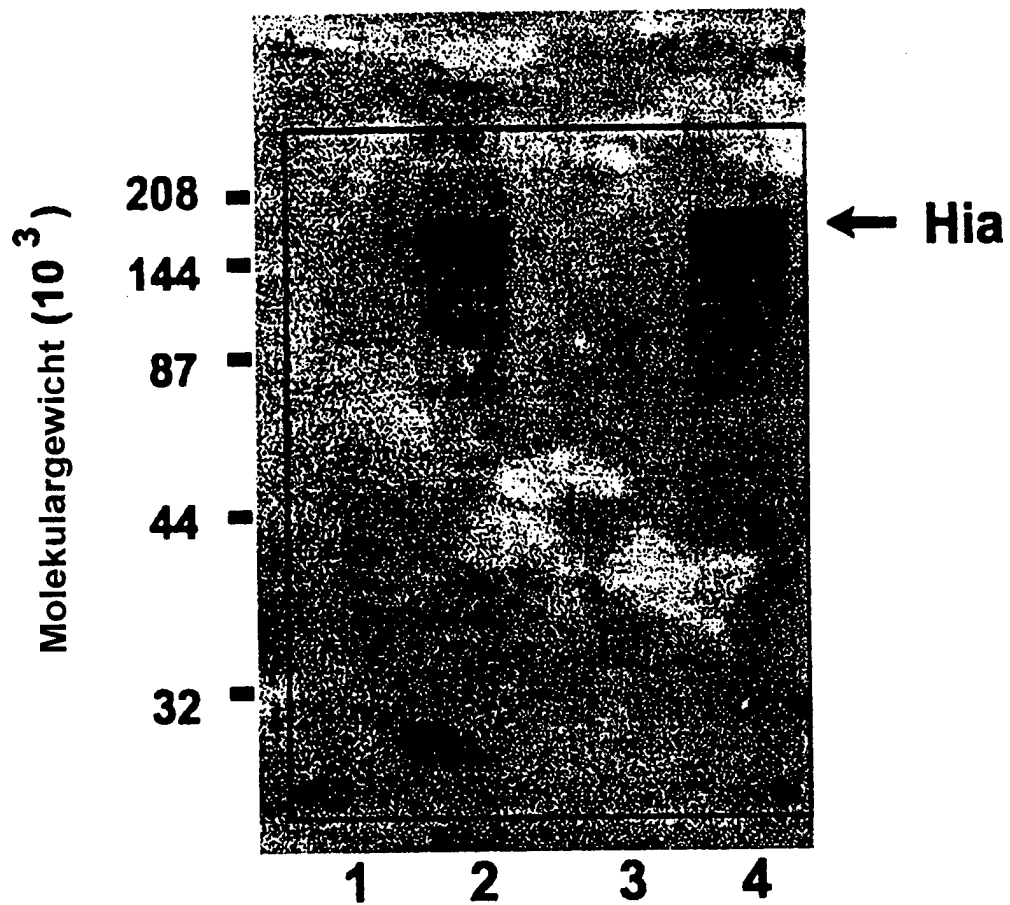
Sinn

EcoR I Nde I
↓ ↓
M N K I F N V I W N P
5' GCGAATTCATATGAACAAAATTTTAAACGTTATTGGAAT 3' SEQ ID NO:2
5038.SL SEQ ID NO:1

Gegensinn

K T G V A A G V G Y Q W * *
5' AAACAGCGGTGACAGGAGGTCGTTGTTACCAAGTGGTAATAG 3' SEQ ID NO:5
3' TTTTGTCCGAACGTCGTCACCAACCAATGGTCACCATTTATCTTAAAGGCCTAGGCG 5' SEQ ID NO:4
5039.SL SEQ ID NO:3
↑ ↑
EcoR I BamH I

FIG.2



Konstruktion von DS-2092-1 und DS-2092-40,
Plasmiden, welche Tandem-T7-hia (11)-Gene enthalten.

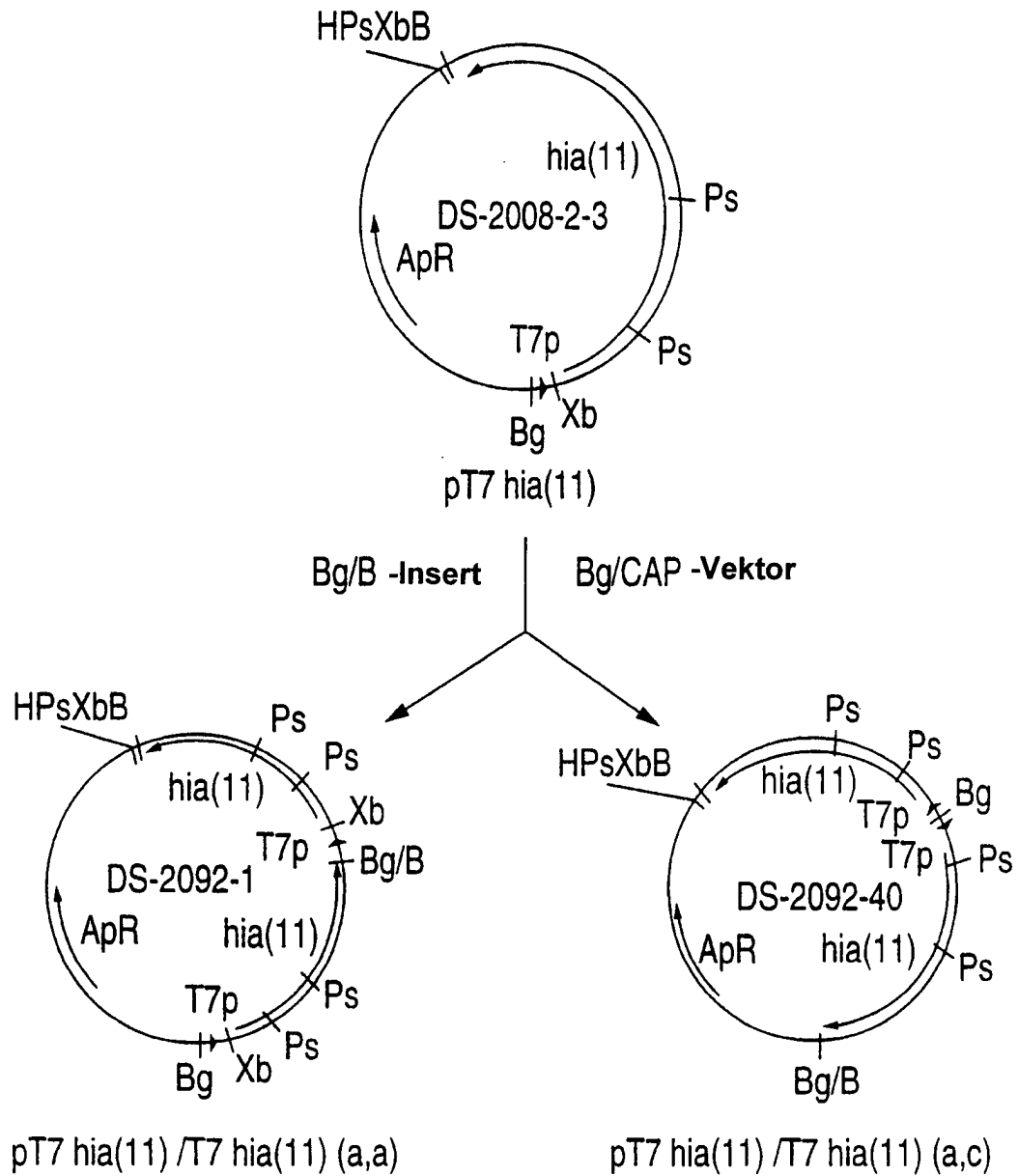
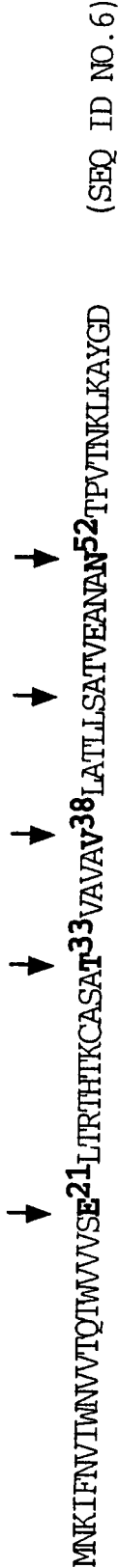


FIG.3

FIG.4

Stellen für N-terminale Trunkierungen von rHia-Proteinen.



Konstruktion von Plasmiden, welche trunkierte hia (11)-Gene exprimieren.

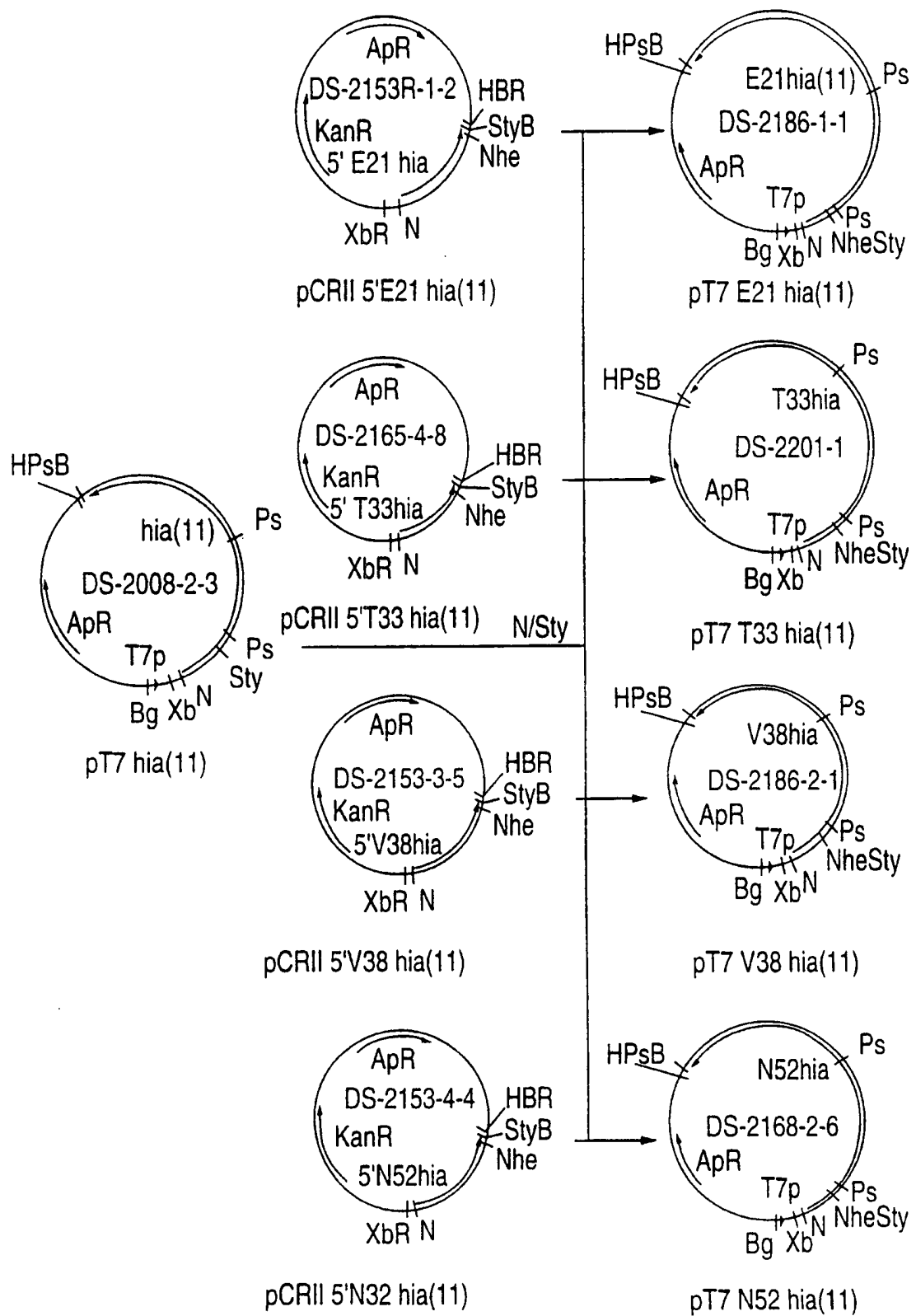


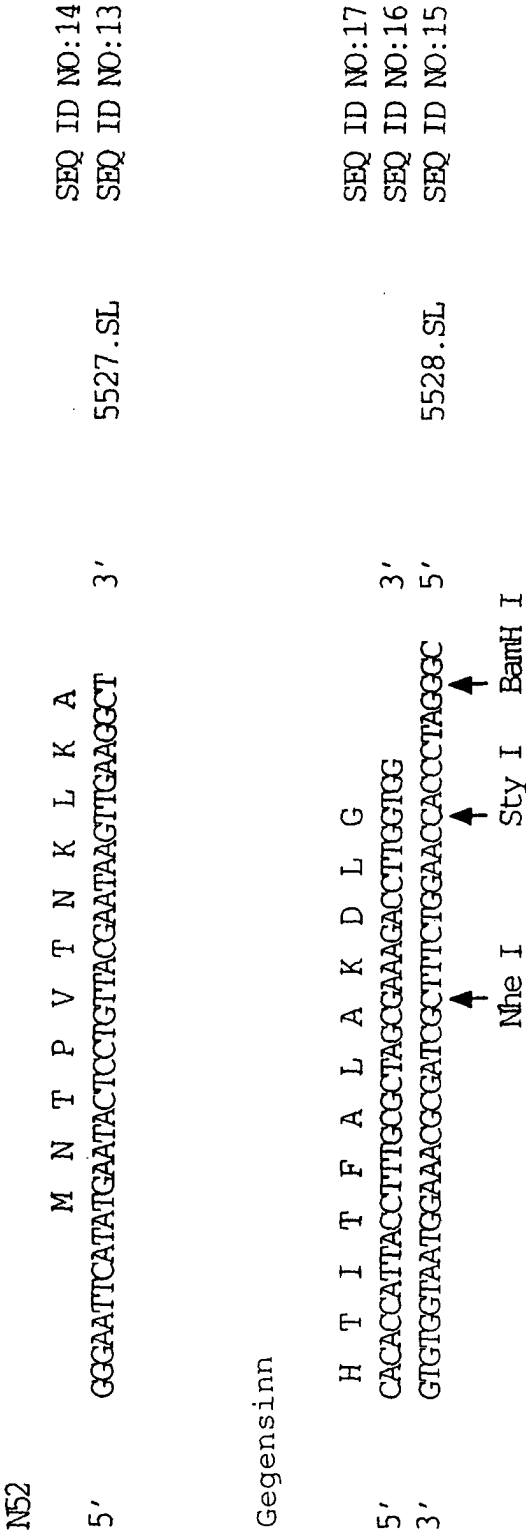
FIG.5A

FIG.5B

Oligonukleotidprimer, um trunkierte *hia*-Gene des Stamms 11 mit PCR zu amplifizieren.

E21	EcoR I Nde I ↓ ↓ M E L T R T H T K C A GGGAATTCATATGGAACTCACCTCGCACCCACACCAATGGGCC	3'	5524.SL	SEQ ID NO: 8 SEQ ID NO: 7
T33	M T V A V A V L A T L GGGAATTCATATGACCGTGGCGTTGCCGTATTGGCAACCCCTG	3'	5525.SL	SEQ ID NO:10 SEQ ID NO: 9
V38	M V L A T L L S A T GGGAATTCATATGGTATTGGCAACCCCTGTCGCAACG	3'	5526.SL	SEQ ID NO:12 SEQ ID NO:11

FIG.5B'



Konstruktion von BK-96-2-11,
einem Plasmid, das T7-V38-hia (11) und cer enthält.

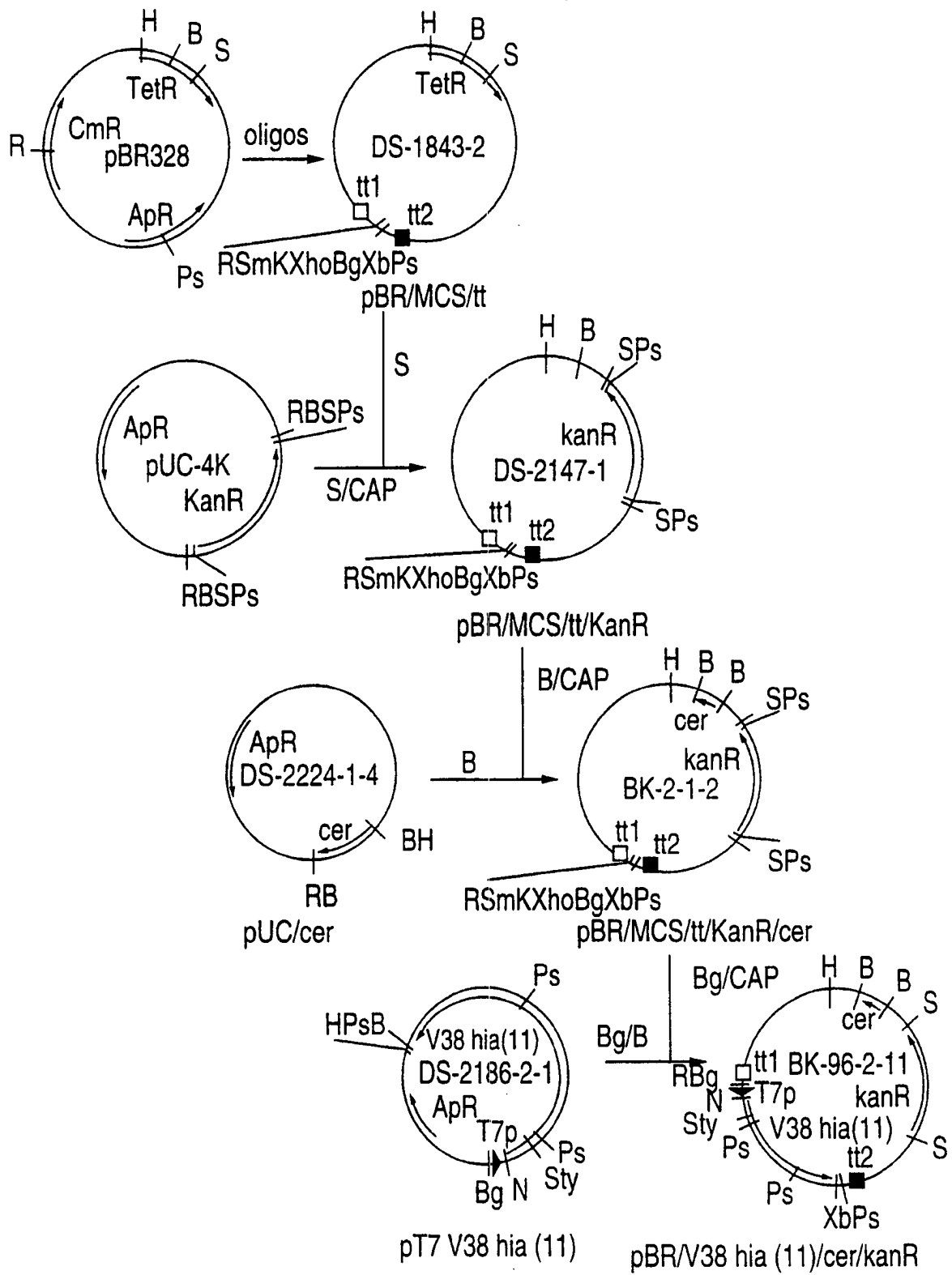


FIG.6A

Konstruktion von DS-2242-1 und DS-2242-2,
Plasmiden, welche T7-hia (33) und cer enthalten.

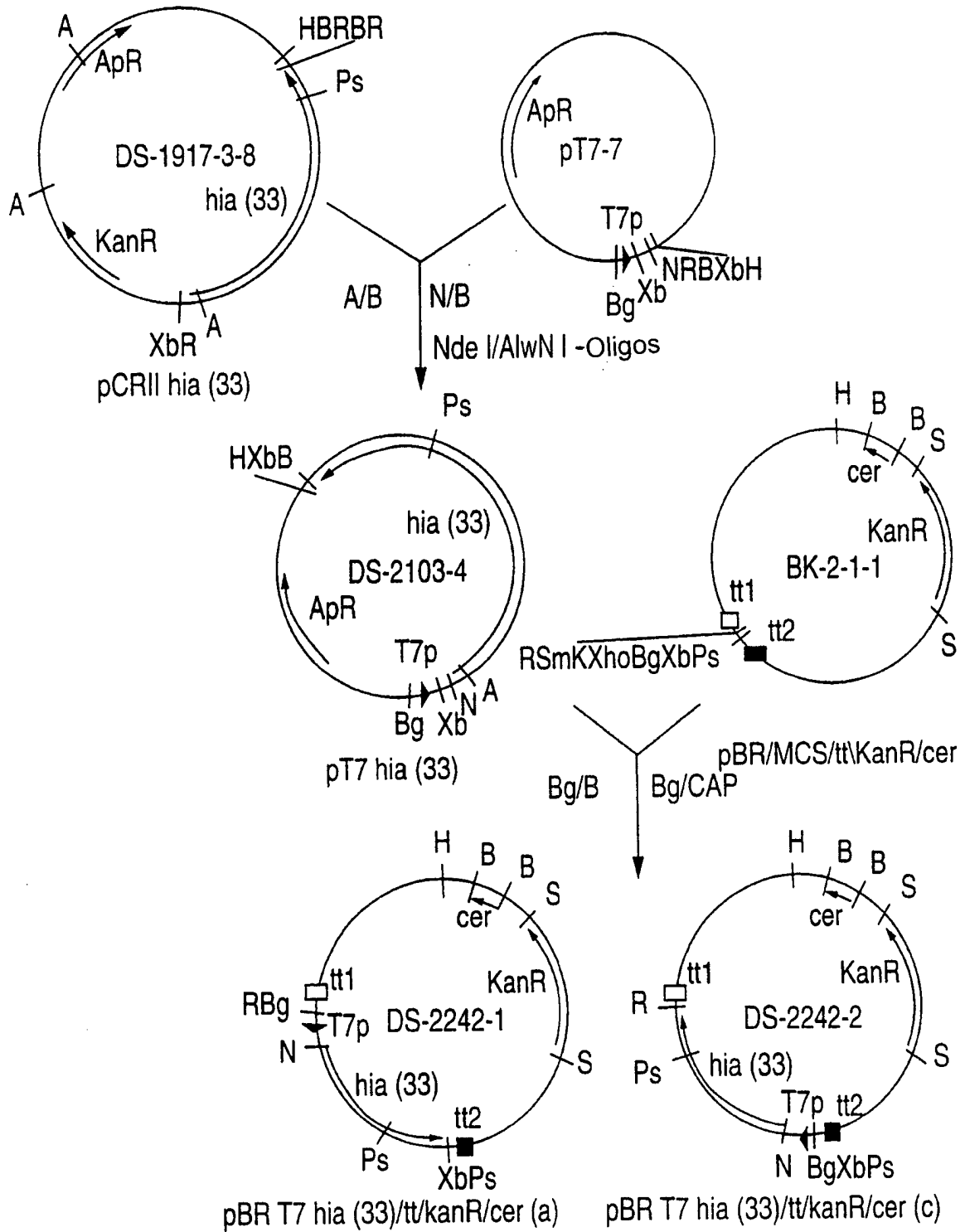


FIG.7A

FIG.7B

Oligonukleotide, die verwendet wurden, um das 5'-Ende des *hla*-Gens des Stamms 33 für Expressionsstudien zu erzeugen.

```
Nde I
↓
M N K I F N V I W N V M T Q T W A V V S E L T R A H T K...
TATGAACAAAATTTTAAAGTTATTGGGAATGTTATGACTCAAACCTTGGGCTGTGG
TATCTGAACCTCACTGGCGGCCACACCA...
ACTTGTTTTAAAAATTGCAATPAAACCTTACAATACTGAGTTT
GAACCCGACAGCATAGACTTGAGTGAGCGGGGTGGT...
...
... R A S A T V A A SEQ ID NO:54
... AACGTGCTCCGCAACCGTGGCAGCCG SEQ ID NO:52
... TTGCACGGAGGCGTTGGCACCCGTC SEQ ID NO:53
AlwNI
↑
...
```


Konstruktion von DS-2340-2-3,
einem Plasmid, das T7-V38-hia (33) und cer enthält.

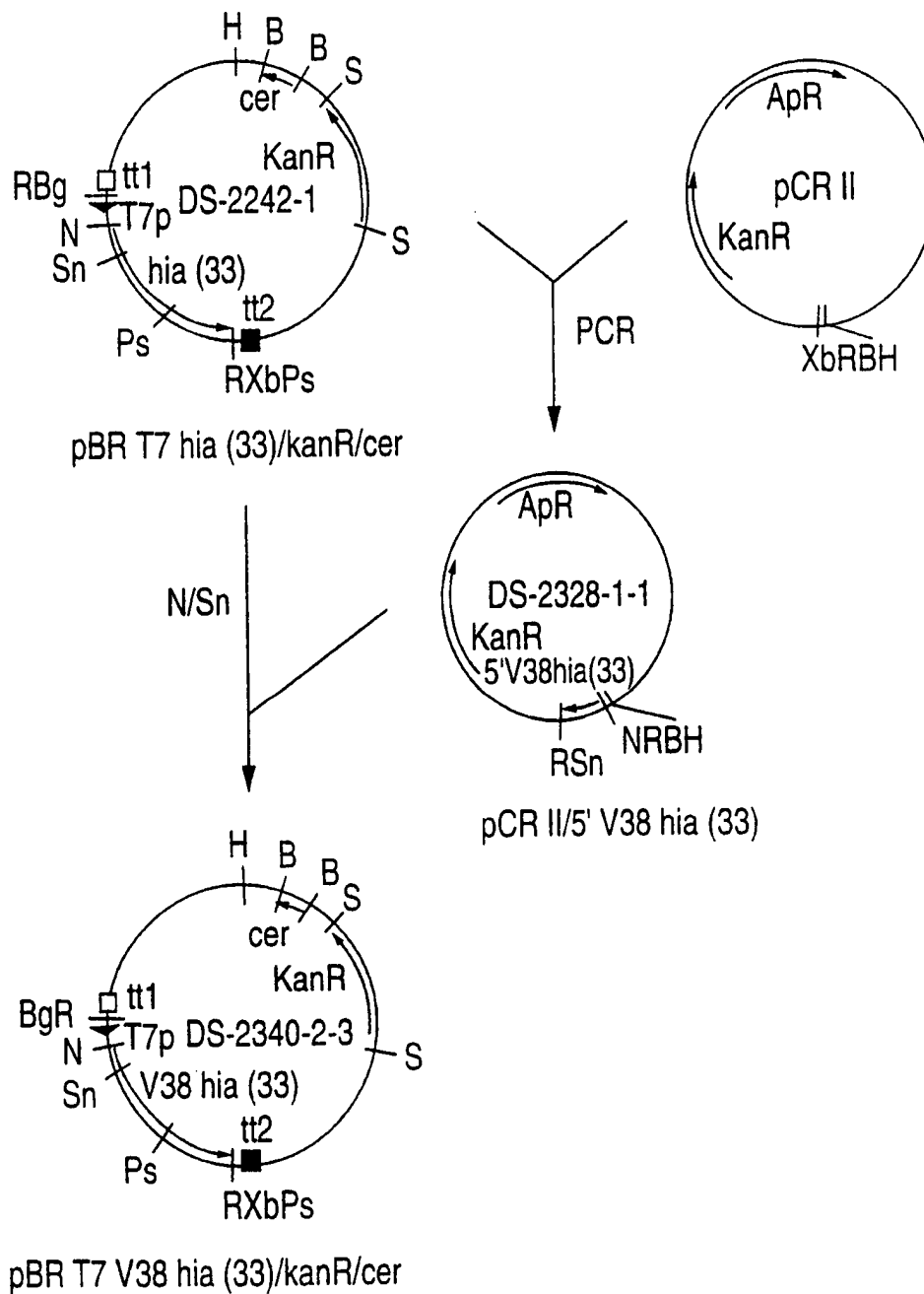


FIG.8A

FIG. 8B

Oligonukleotide, die verwendet wurden, um das *hla*-Gen des Stamms 33 ausgehend von dem V38-Codon bis zu der *Sna*B I-Stelle mit PCR zu amplifizieren.

Sinn

Nde I
 ↓ M V L A T V L S A T
 5' GCGAATTCATGTTATGGCCACCGTATTGTCGCAACG 3' 6286.SL

Gegensinn

		SnaB I															
		↓															
		D	E	T	T	A	T	V	G	N	L	R	K	L			
5'		GACGAACACACCGCAACGTAAGCAATTACGTAATTTGAAGCTTGG													3'	SEQ ID NO:19	
3'		CTGCTTTGGTGGGTTGGCATCGTTAAATGCATTAACTTGAAGC													5'	6287.SL	
																SEQ ID NO:18	
																SEQ ID NO:20	

Konstruktion von DS-2447-2,
einem Plasmid, welches Tandem-T7-V38-hia(11)-Kassetten und cer enthält.

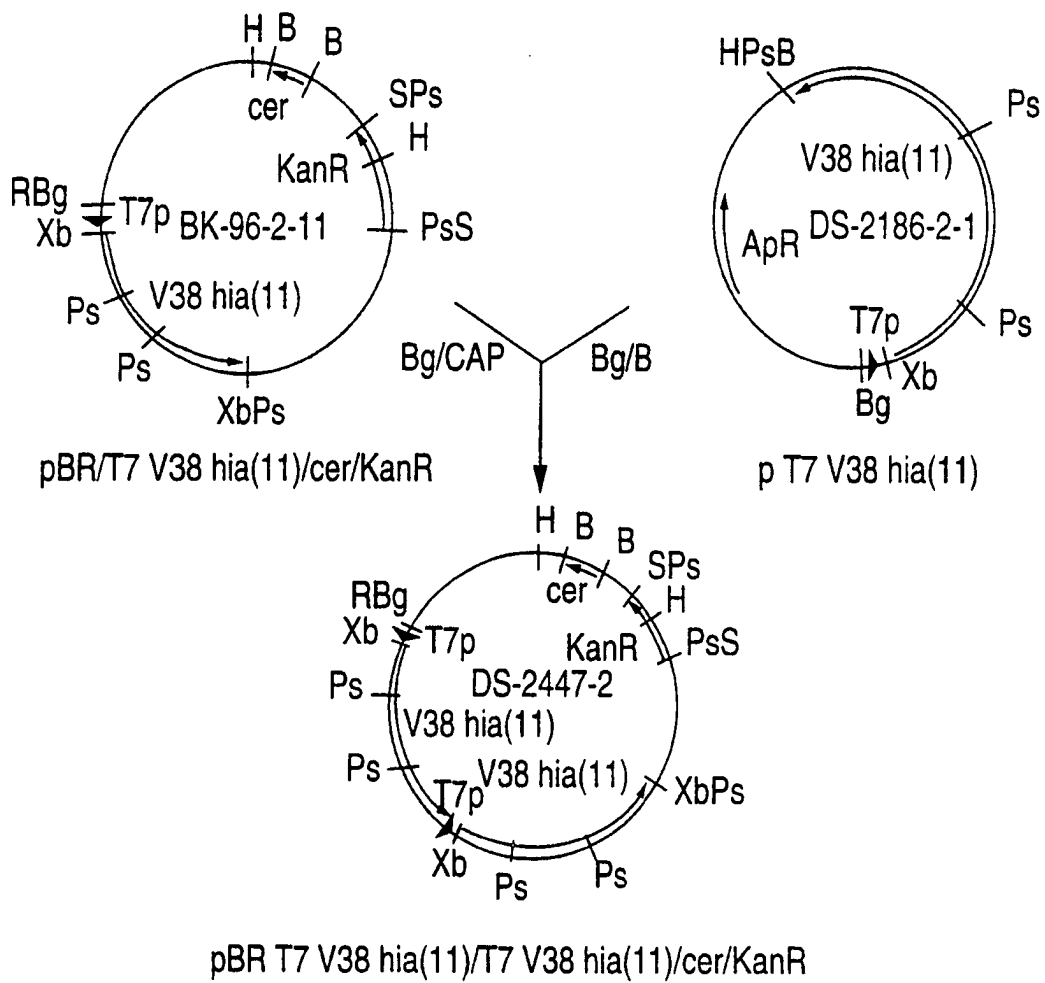


FIG.9A

Konstruktion von DS-2448-17,
einem Plasmid, welches Tandem-T7-V38-hia (33)-Kassetten und cer enthält.

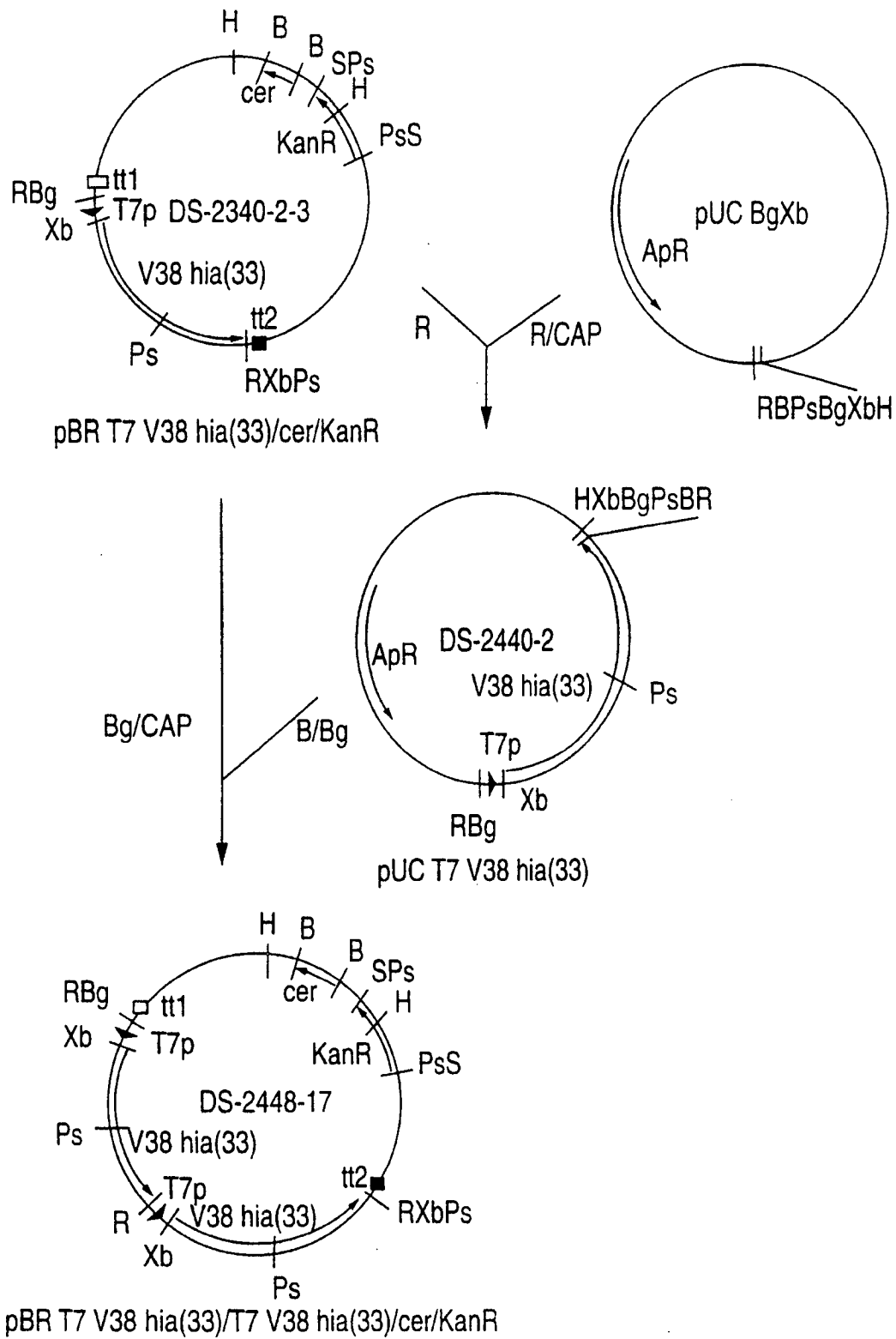
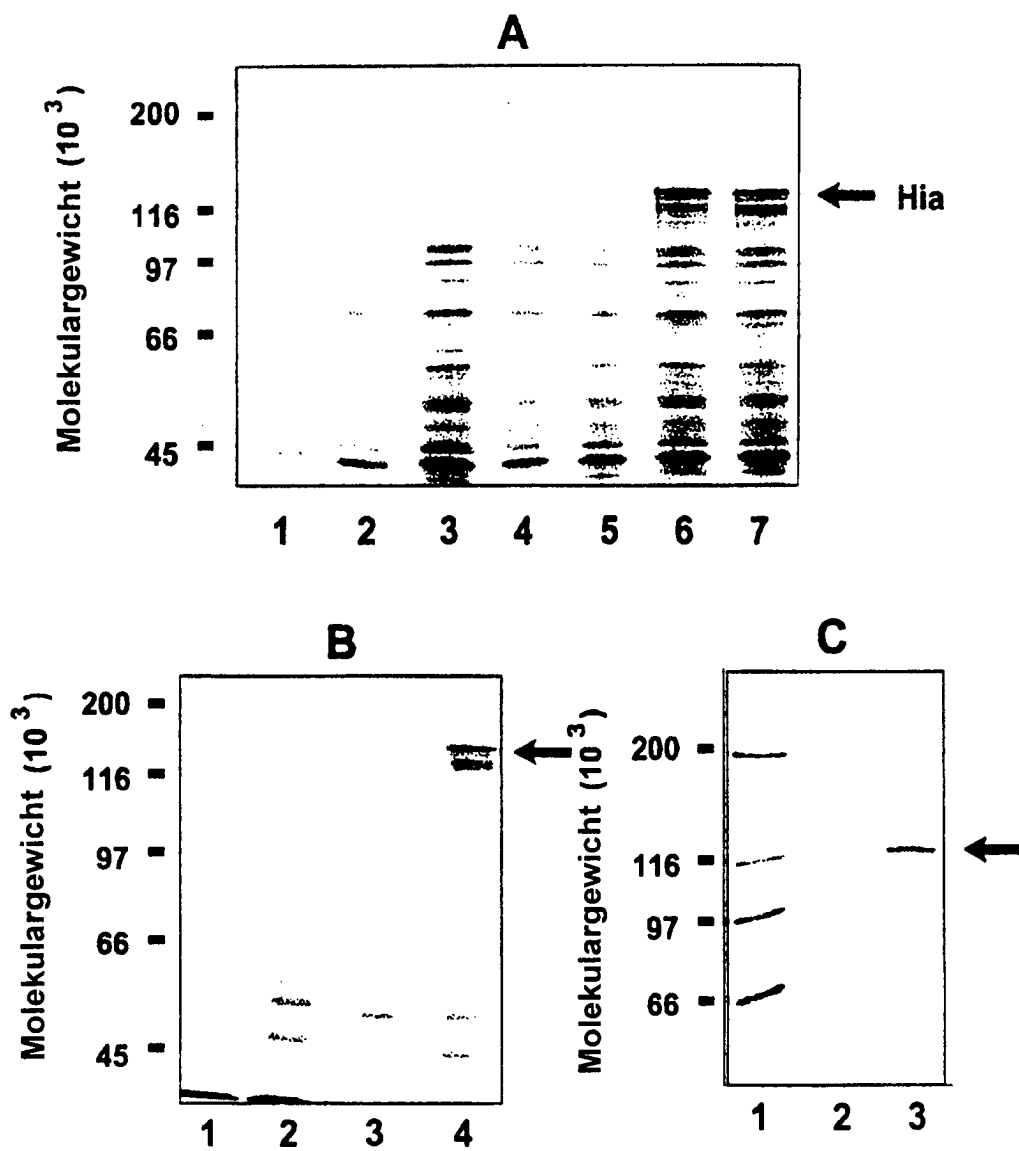


FIG.9B

FIG.10



Reinigung von rHia-Proteinen aus E. coli

E. coli Ganze Zellen

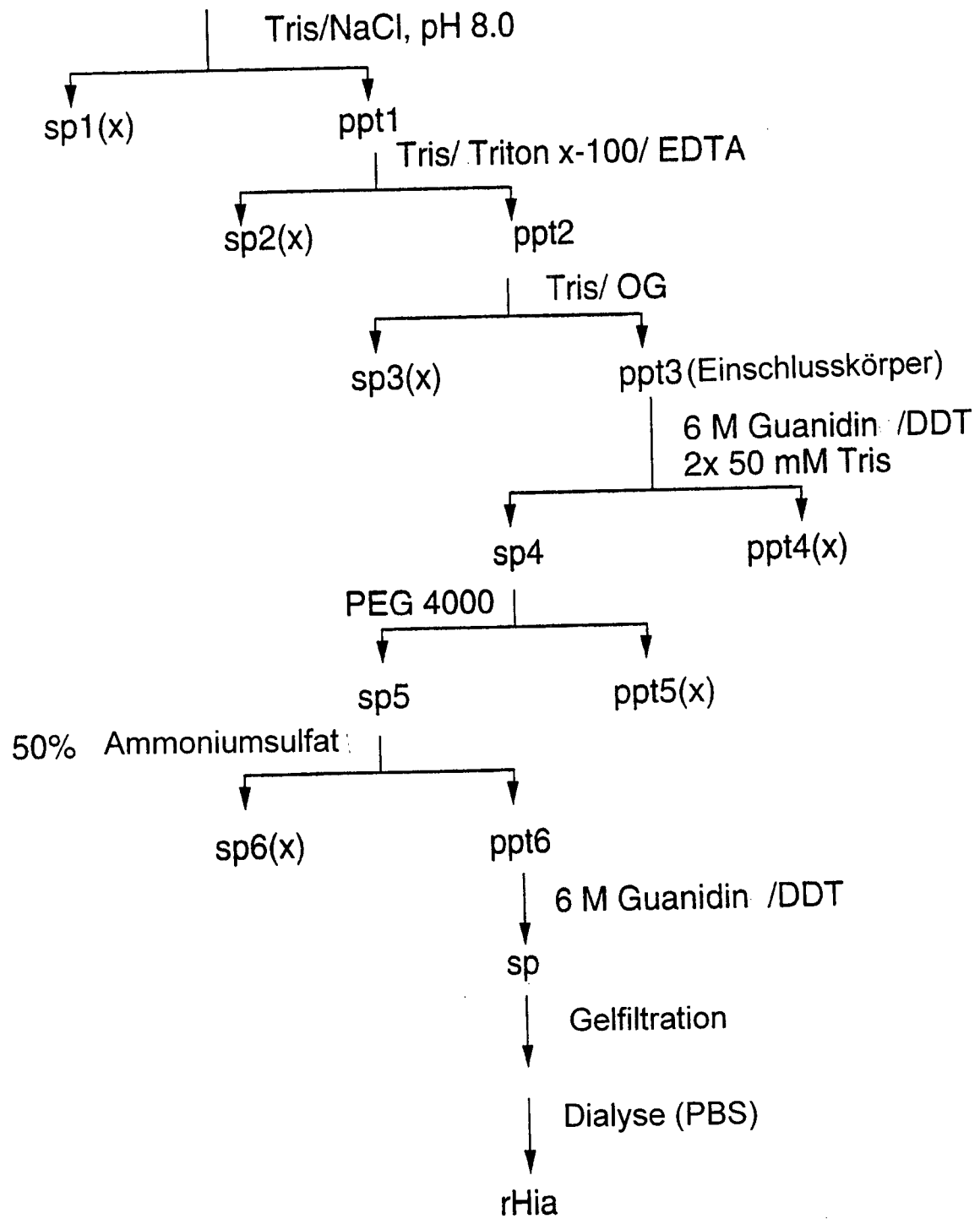
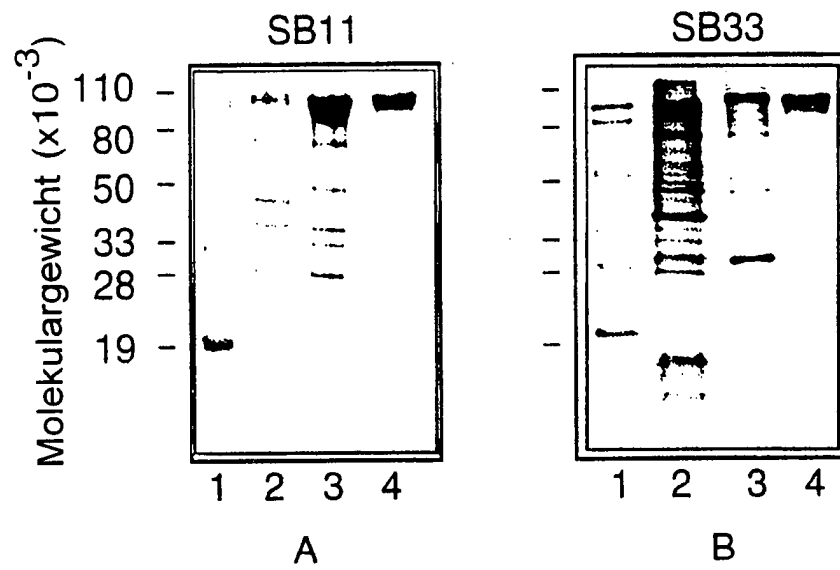


FIG.11

Reinigung von rHia (V38) aus E. coli



1. Vorgefärbte Molekulargewichtsmarker
2. E. coli-Ganzzelllysat
3. Roher Extrakt
4. Gereinigtes rHia-Protein

FIG.12

Die Stabilität von rHia (V38/SB11)

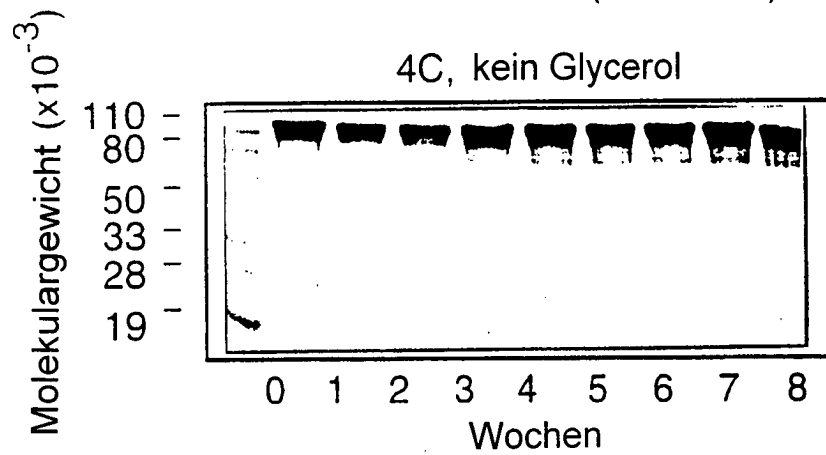


FIG.13A

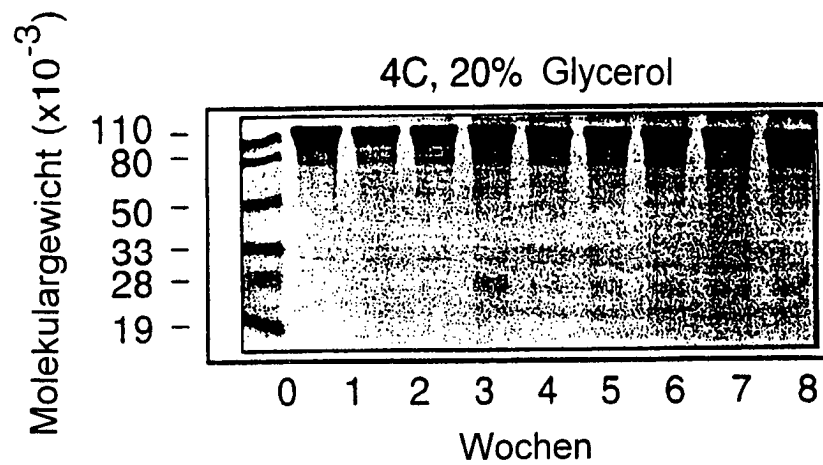


FIG.13B

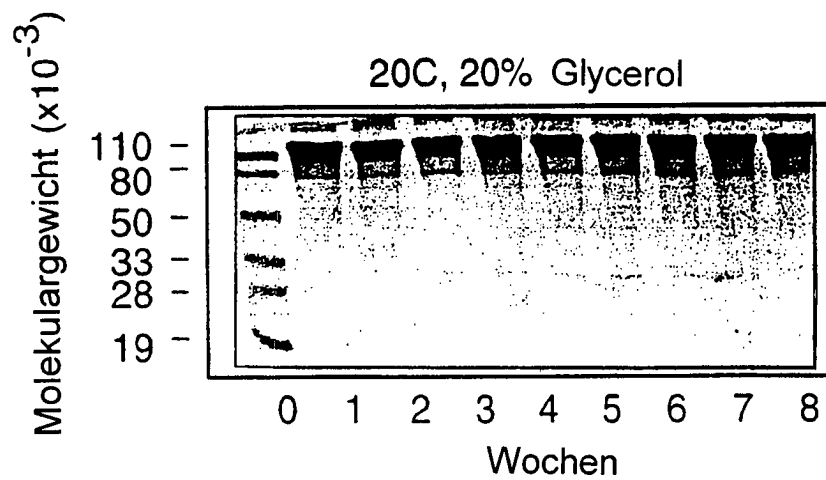


FIG.13C

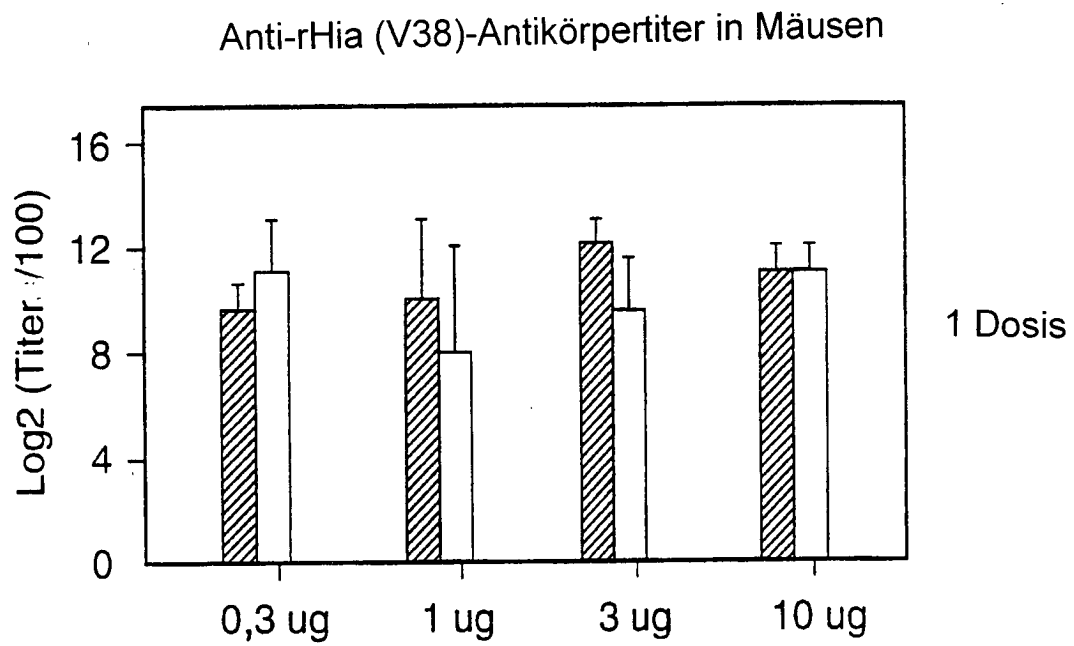
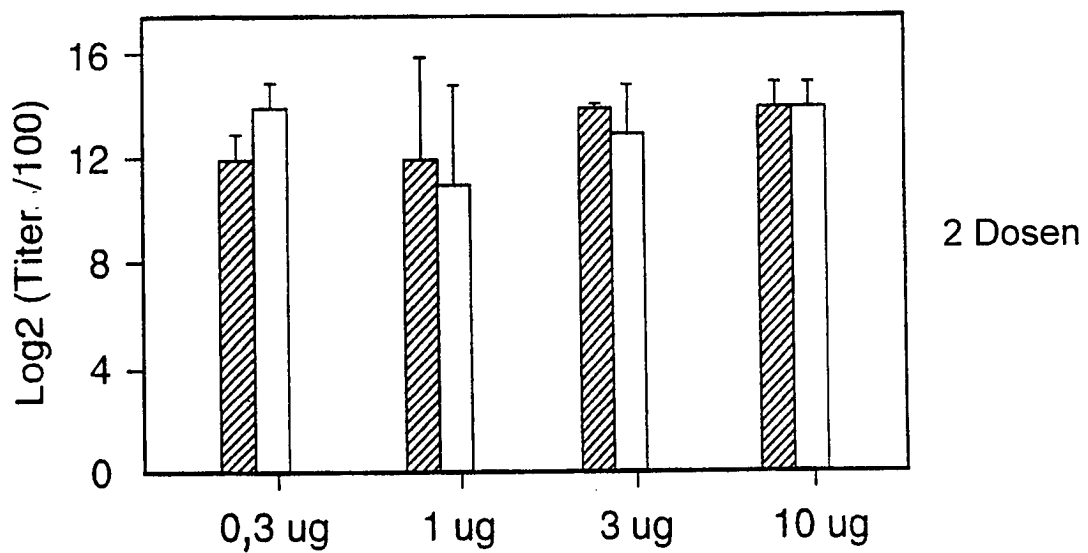


FIG.14A



SB11
SB33

FIG.14B

Anti-V38 rHia (SB11)-Antikörpertiter in BALB/c-Mäusen

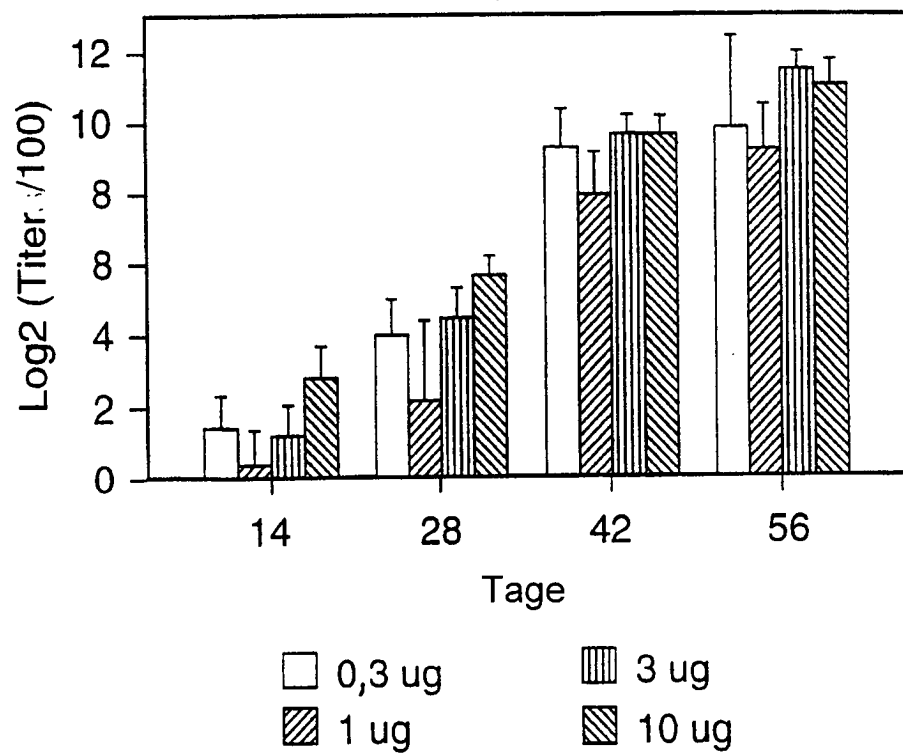


FIG.15A

Anti-V38 rHia (SB11)-Antikörpertiter in Meerschweinchen

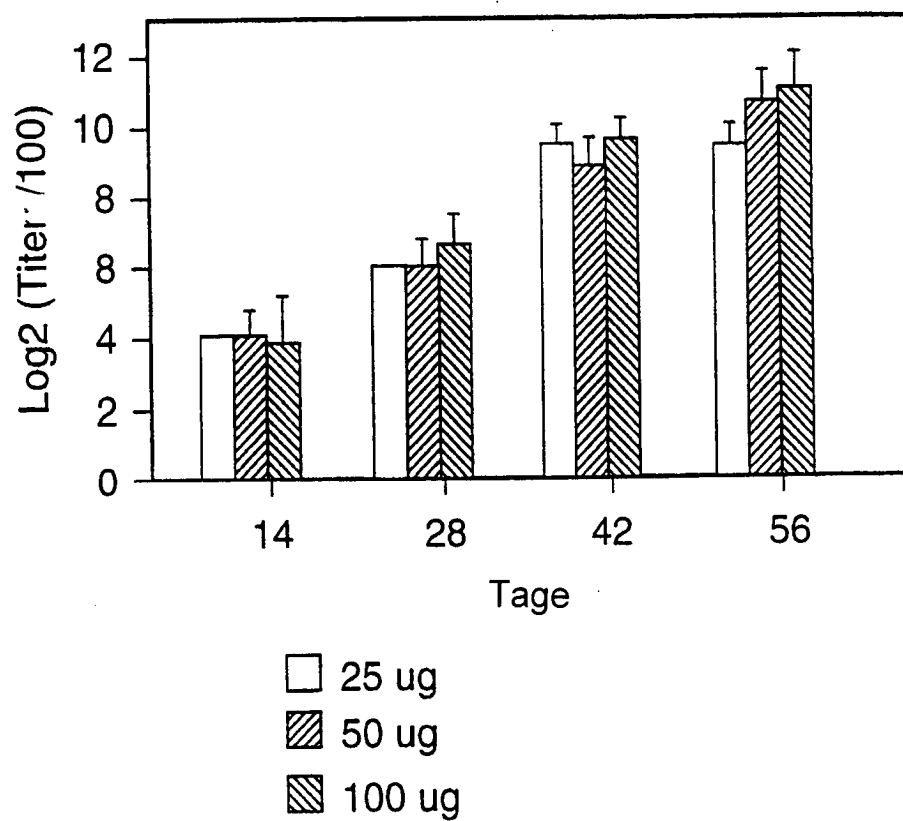


FIG.15B

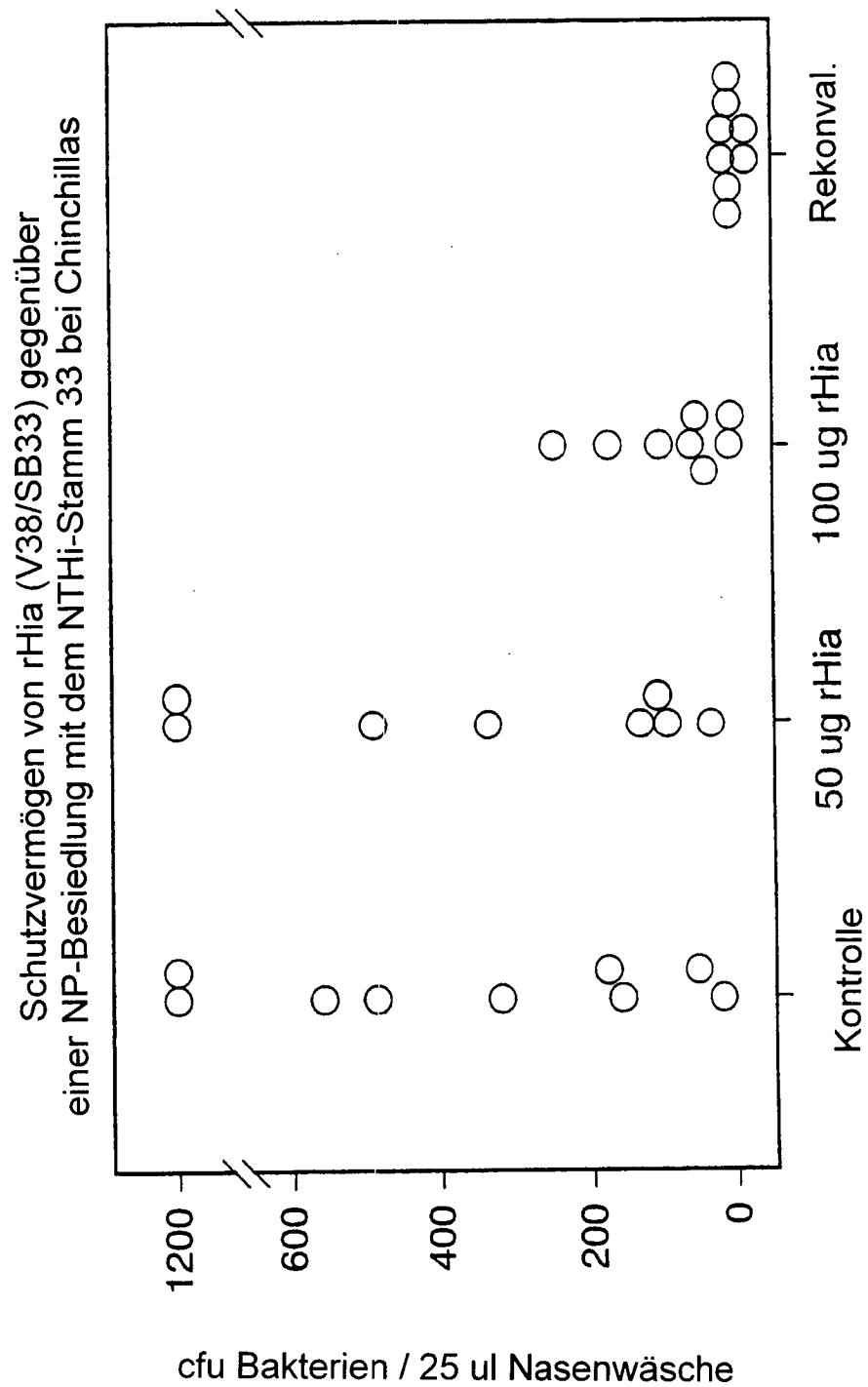


FIG. 16

FIG.17

Oligonukleotide, die verwendet wurden, um zusätzliche *hia*-Gene mit PCR zu amplifizieren.

Sinn

	M N K I F N V		
5'	TTAAATATAAGGTAAATPAAAAATGAACAAAATTTTAAAGTT	3'	5040.SL
			SEQ ID NO:21

Gegensinn

	K T G V A A G V G Y Q W * *		
5'	AAAACAGCGTTGCAGCAGGTCGTTGGTTACCACTGCTAATAG	3'	
3'	TTTGTGCGCAAGTCGTCACCAACCAATGGTCACCATTTATCTTAAGGCCTAGCG	5'	5039.SL
			SEQ ID NO:5
			SEQ ID NO:4
			SEQ ID NO:3

↑ ↑
EcoR I BamH I

FIG.18A

NTHi-Stamm 33 Hia

```

      MET ASN LYS...
GAATTCGGCTTAATAATAAATGAACA...
10      ... ILE PHE ASN VAL ILE TRP ASN VAL MET THR GLN
      ...AATTTTAAACGTTATTTGGAATGTTATGACTCA
      ... 30      40      50      60

      THR TRP ALA VAL VAL SER GLU LEU THR...
      AACTTGGGCTGTCGTATCTGAACTCAC...
70      ... ARG ALA HIS THR LYS ARG ALA SER ALA THR VAL
      ...TCGGCCCCACCAACGTCCTCCGCAACCGT
      ... 90      100      110      120

      ALA ALA ALA VAL LEU ALA THR VAL LEU...
      GGCAAGCCGTGTATTGGCGACCGTATT...
130      ... SER ALA THR VAL GLN ALA SER ALA GLY SER THR
      ...GTC TGCAACGGTTTCAGGCGAGTGCAAGGCAGTAC
      ... 150      160      170      180

      THR GLY THR ASN SER LEU ASN VAL TYR...
      GACAGGTACAAATAGTTTGAAATGTTTA...
190      ... 200

```

FIG.18B

```

...   GLY   LYS   ASN   ASN   SER   ASN   PHE   ASN   SER   ALA   ASN
...TGGAAAGAAATAATTCTGAAATTCTCAATTCTAGCCAA
...   210   220   230   240

ASN   SER   ILE   ALA   ASP   LEU   ASN   LYS   GLN...
TAATTCATAAGCAGATTTAATAACA...
250   260   ...
...   ASN   ASP   SER   VAL   TYR   ASP   GLY   LEU   ASN   LEU
...AATGATAGTTTACGATGGTTTATAAATCT
...   270   280   290   300

ASN   GLU   LYS   GLY   THR   ASP   LYS   SER   LYS...
GAATGAATAAGGTACGGATAAGTCAAA...
310   320   ...
...   PHE   LEU   VAL   ALA   ASP   GLU   THR   THR   ALA   THR   VAL
...ATCCTGGTTGCTGACGAAACCAACCGCAACCGT
...   330   340   350   360

GLY   ASN   LEU   ARG   LYS   LEU   GLY   TRP   VAL...
AGGCAATTACGTAAATTGGGTTGGGT...
370   380   ...
...   VAL   SER   THR   LYS   ASN   SER   THR   LYS   GLU   GLU   SER
...AGTATCAACCAAAACAGTACGAAAGAGAGAG
...   390   400   410   420

```

FIG.18C

```

ASN  GLN  VAL  LYS  GLN  ALA  ASP  GLU  VAL...
CAATCAAGTCAAAACAGGCGGATGAGT...
430
...
...  LEU  PHE  GLU  GLY  LYS  ASP  GLY  VAL  THR  VAL  THR
...GTTGTTTGAAAGGCAAAAGACGGGTGTAACGGTTAC
460
... 450
... 470
... 480

SER  LYS  SER  GLU  ASN  GLY  LYS  HIS  THR...
TTCCAAATCTGAATAACGGCAACACAC...
490
...  VAL  THR  PHE  ALA  LEU  ALA  ASN  ASP  LEU  ASN  VAL
...CGTTACTTTTGCCCTTGCGCAATGACCTTAATGT
500
... 510
... 520
... 530
... 540

LYS  ASN  ALA  THR  VAL  SER  ASP  LYS  LEU...
AAACACGCAACCGTTAGCGATAAATT...
550
...  SER  LEU  GLY  ALA  ASN  GLY  LYS  VAL  ASP  ILE
...ATCGCTTGGTGCAACCGGCAAGAGATCGATAT
560
... 570
... 580
... 590
... 600

THR  SER  ASP  ALA  ASN  GLY  LEU  LYS  PHE...
TACCAAGTGCAACCGGCTTGAAATT...
610
... 620
...

```


FIG.18D

```

...   ALA  LYS  GLN  GLY  THR  ASN  GLY  GLN  ASN  GLY  ASN
...TGC GA A A C A G G G T A C G A A T G G T C A A A A C G G T A A
...   630                                     650       660

VAL  HIS  LEU  ASN  GLY  ILE  ALA  SER  THR...
TGT TCA CTTTAAACGGTATTGCTTCGAC...
670                                     680
...   LEU  ASP  ASP  PRO  ARG  VAL  GLY  GLY  LYS  THR  ALA
...TTAGATGATCCCTCGTG TGGGTGGA A A A C A G C
...   690                                     700       710       720

HIS  LEU  THR  LYS  GLU  ILE  SER  ASP  THR...
ACACCTTACAAAGAAATCAGCGATAC...
730                                     740
...   GLU  ARG  ASN  ARG  ALA  ALA  SER  VAL  GLY  ASP  VAL
...AGACGTAA CCGTGCTGCCGAGCCGTGGGCGGATGT
...   750                                     760       770       780

LEU  ASN  ALA  GLY  TRP  ASN  ILE  ARG  GLY...
ATTGAATGCCGGGTGGGAATAATTCGTGG...
790                                     800
...   ALA  LYS  THR  ILE  GLY  GLY  THR  VAL  ASP  ASN  VAL
...CGCAAAACGAT TGGCGGTACAGTGGATATGT
...   810                                     820       830       840

```

FIG.18E

```

ASP  PHE  VAL  SER  THR  TYR  ASP  THR  VAL...
TGA TTT GTT TCA ACTT ATGACACTGT ...
850
...  GLU  PHE  ALA  SER  GLY  ALA  ASN  ALA  ASN  VAL  SER
...TGAATTTGCCAGCGGCGCAACGCAATGTGAG
... 870 880 890 900

VAL  THR  THR  ASP  ASP  ASN  LYS  LYS  THR...
CGTTACGACTGATGATTAACAATAAAC ...
910
...  THR  VAL  ARG  VAL  ASP  VAL  THR  GLY  LEU  PRO  VAL
...AACCGTCCGTGTGGATGTAAACAGGCTTGCCGGT
... 930 940 950 960

GLN  TYR  VAL  THR  GLU  ASP  SER  LYS  THR...
CCATAATGTTACGGAGAGACAGCAAAAC ...
970
...  VAL  VAL  LYS  VAL  GLY  ASN  GLU  TYR  TYR  GLU  ALA
...CGTTGTGAAGAAGTGGGCAATGAGTATTACGAGAGC
... 990 1000 1010 1020

LYS  GLN  ASP  GLY  SER  ALA  ASP  MET  ASP...
CAGCAAGACGGTTCGGCGGATATGGA ...
1030 1040

```

FIG.18F

... LYS LYS VAL GLU ASN GLY LYS LEU ALA LYS THR
 ...T A A A A A G T C G A A A T G G C A A G C T G G C G A A A C
 ... 1050 1060 1070 1080

LYS VAL LYS LEU VAL SER ALA ASN GLY...
 T A A G T G A A A T T G G T A T C G G C A A A C G G ...
 1090 ...
 ... THR ASN PRO VAL LYS ILE SER ASN VAL ALA ASP
 ...T A C A A A T C C G G T G A A A T C A G C A A T G T T G C G G A
 ... 1110 1120 1130 1140

GLY THR GLU ASP THR ASP ALA VAL SER...
 C G G C A C G G A A G A T A C C G A T G C G G T C A G ...
 1150 ...
 ... PHE LYS GLN LEU LYS ALA LEU GLN ASP LYS GLN
 ...C T T T A A G C A G T T G A A A G C C T T G C A A G A T A A C A
 ... 1170 1180 1190 1200

VAL THR LEU SER ALA SER ASN ALA TYR...
 G G T T A C G T T A G T G C G A G C A A T G C T T A ...
 1210 ...
 ... ALA ASN GLY GLY SER ASP ALA ASP GLY GLY LYS
 ...T G C C A A T G G C G G T A G C G A T G C C G A C G G C G G C A
 ... 1230 1240 1250 1260

FIG.18G

```

ALA  THR  GLN  THR  LEU  GLY  ASN  ASP  LEU...
GGCAACTCAAACTTTAGGCAATGATTT...
1270
...  ASN  PHE  LYS  PHE  LYS  SER  THR  ASP  SER  GLU  LEU
...GAAATTTTAAATTTAAATCCACACAGCAGTTT
... 1290 1300 1310 1320
LEU  ASN  ILE  LYS  ALA  ALA  GLY  ASP  THR...
GTTGAACATCAAGCAGCAGGTGACAC...
1330
...  VAL  THR  PHE  THR  PRO  LYS  LYS  GLY  SER  VAL  GLN
...GGTTACCTTTACGCCGAAATAAGGTTCCGGTGCA
... 1350 1360 1370 1380
VAL  GLY  ASP  ASP  GLY  LYS  ALA  THR  ILE...
GGTTGGCGATGATGGTAAGGCTACGAT...
1390
...  GLN  ASP  GLY  ALA  LYS  THR  THR  THR  GLY  LEU  VAL
...TCAAGACGGCGCGAATAACAATACTACCGGTTTGGT
... 1410 1420 1430 1440
GLU  ALA  SER  GLU  LEU  VAL  ASP  SER  LEU...
TGAGGCTTCTGAATTGGTTTGACAGCCT...
1450

```

FIG.18H

```

... ASN LYS LEU GLY TRP LYS VAL GLY LYS
...G A A C A A T T G G C C T G G A A A G T G G C C G T T G G T A A 1500
... 1470 1480 1490

ASP GLY THR GLY ALA THR ASP GLY THR...
A G A C G G C A C A G G A G C G A C C G A T G G C A C ...
1510
... HIS THR ASP THR LEU VAL LYS SER GLY ASP LYS
...G C A T A C C G A C A C T T A G T G A A G T C G G C G A T A A 1560
... 1530 1540 1550

VAL THR LEU LYS ALA GLY ASP ASN LEU...
A G T A A C T T G A A A G C C G G C G A T A T C T ...
1570
... LYS VAL LYS GLN GLU GLY THR ASN PHE THR TYR
...G A A G G T C A A A C A A G A G G G T A C A A C T T C A C T T A 1620
... 1590 1600 1610

VAL LEU ARG ASP GLU LEU THR GLY VAL...
C G T G C T C A G A G A T G A A T T G A C G G C G T ...
1630
... LYS SER VAL GLU PHE LYS ASP THR GLU ASN GLY
...A A A G A G C C G T G G A G T T T A A A G A C A C G A G A A T G G 1680
... 1650 1660 1670

```

FIG.18I

```

ALA ASN GLY ALA SER THR LYS ILE THR...
TGC A A A C G G T G C A A G C A C G A G A T T A C ...
1690
... LYS ASP GLY LEU THR ILE THR PRO ALA ASN ASP
...C A A A G A C G G C T T G A C C A T T A C G C C G G C A A A C G A
... 1710 1720 1730 1740

ALA ASN GLY ALA ALA THR ASP ALA...
TGC G A A T G G T G C G C G C G A C T G A T G C ...
1750
... ASP LYS ILE LYS VAL ALA SER ASP GLY ILE SER
...T G A C A A G A T T A A A G T G G C T T C A G A C G G C A T T A G
... 1770 1780 1790 1800

ALA GLY ASN LYS ALA VAL LYS ASN VAL...
TGC G G T A A T A A G C A G T T A A A A C G T ...
1810
... VAL SER GLY LEU LYS LYS PHE GLY ASP ALA ASN
...T G T G A G C G G A C T G A A G A A A T T T G G T G A T G C G A A
... 1830 1840 1850 1860

PHE ASN PRO LEU THR SER SER ALA ASP...
T T T C A A T C C G C T G A C T A G C T C A G C C G A ...
1870 1880

```

FIG.18J

```

... ASN LEU THR LYS GLN TYR ASP ASN ALA TYR LYS
...C A A C T T A C G A A A C A T A T G A C A A T G C C T A T A A
... 1890 1900 1910 1920

GLY LEU THR ASN LEU ASP GLU LYS SER...
A G G C T T G A C C A A T C T G G A T G A A A A G ...
1930 1940 ...
... LYS GLY LYS GLN THR PRO THR VAL ALA ASP ASN
...T A A A G G C A A G C A A A C T C C G A C C G T T G C T G A C A A
... 1950 1960 1970 1980

THR ALA ALA THR VAL GLY ASP LEU ARG...
T A C C G C T G C A A C C G T G G G C G A T T T G C G ...
1990 2000 ...
... GLY LEU GLY TRP VAL ILE SER ALA ASP LYS THR
...C G G T T T G G G C T G G G T C A T T T C T G C A G A C A A A C
... 2010 2020 2030 2040

THR GLY GLU SER LYS GLU TYR SER ALA...
C A C A G G C G A G T C A A A G G A A T A T A G C G C ...
2050 2060 ...
... GLN VAL ARG ASN ALA ASN GLU VAL LYS PHE LYS
...G C A A G T G C G T A A C G C C A A T G A A G T G A A T T C A A
... 2070 2080 2090 2100

```

FIG.18K

```

SER  GLY  ASN  GLY  ILE  ASN  VAL  SER  GLY...
G A G C G G C A A C G G T A T C A A T G T T T C C G G ...
2110
...
...  LYS  THR  LEU  ASP  ASN  GLY  THR  ARG  GLU  ILE  THR
...T A A A C A T T G G A T A C G G T A C G C G C G A A T T A C
... 2130 2140 2150 2160

PHE  GLU  LEU  ALA  LYS  ASP  GLU  ASN  ALA...
T T T T G A A T T G G C T A A G A C G A A A T G C ...
2170 2180
...
...  ILE  ALA  PHE  GLY  SER  GLY  SER  LYS  ALA  LEU  ARG
...C A T T G C T T T C G G T T C T G G C T C A A A G C C T T G C G
... 2190 2200 2210 2220

ASP  ASN  THR  VAL  ALA  ILE  GLY  THR  GLY...
C G A T A C A C G G T G G C G A T T G G T A C G G G ...
2230 2240
...
...  ASN  VAL  VAL  ASN  ALA  GLU  LYS  SER  GLY  ALA  PHE
...C A A C G T T G T G A A T G C G G A A A A T C T G G T G C A T T
... 2250 2260 2270 2280

GLY  ASP  PRO  ASN  TYR  ILE  GLU  ASP  LYS...
C G G C G A T C C G A A C T A C A T C G A A G A T A A ...
2290 2300
...
```


FIG.18L

```

...   ALA  GLY  GLY  SER  TYR  ALA  PHE  GLY  ASN  ASP  ASN
...AGCCGGTGGCAGCTACGCTTTCGGTAACGATAA  2340
... 2310 2320 2330

ARG  ILE  THR  SER  LYS  ASN  THR  PHE  VAL...
CCGTAATTACTCTAATAACACTTTTGT ...
2350
...   LEU  GLY  ASN  GLY  VAL  ASN  ALA  LYS  TYR  LYS  ALA
...GTTGGGTAAATGGAGTTTAATGCCGAATAATAAGC  2400
... 2370 2380 2390

ASN  GLY  ASP  VAL  ASP  THR  GLU  THR  VAL...
CAATGGAGATGTTGTATACGGAAACCGT ...
2410
...   THR  VAL  LYS  ASP  LYS  ASP  GLY  LYS  GLU  THR  THR
...ACCTGTTAAGGACACAAGACGGTAAGAAGACTAC  2460
... 2430 2440 2450

VAL  THR  VAL  PRO  LYS  ALA  LEU  GLY  ALA...
CGTTACTGTTCTTAAGCGTTAGGGGC ...
2470
...   THR  VAL  GLU  ASN  SER  VAL  TYR  LEU  GLY  ASN  LYS
...TACGGTTGAATAACCTCCGTTTATTGGGTATAATA  2520
... 2490 2500 2510

```

FIG.18M

```

SER  THR  ALA  THR  LYS  ASP  LYS  GLY  LYS...
A T C G A C T G C G A C A A A G A T A G G G T A A ...
2530
... ASN  LEU  LYS  SER  ASP  GLY  THR  ALA  GLY  ASN  THR
...A A A T C T G A A A T C T G A T G G T A C G G C G G T A A C A C
... 2550 2560 2570 2580

THR  THR  ALA  GLY  THR  THR  GLY  THR  VAL...
T A C A A C T G C T G G T A C A A C G G G T A C G G T ...
2590
... ASN  GLY  PHE  ALA  GLY  ALA  THR  ALA  HIS  GLY  ALA
...A A A C G G C T T T G C C G G T G C A A C G G C G C A C G G T G C
... 2610 2620 2630 2640

VAL  SER  VAL  GLY  ALA  SER  GLY  GLU...
G G T T T C T G T C G G C G C A A G C G G C G A A G A ...
2650
... ARG  ARG  ILE  GLN  ASN  VAL  ALA  ALA  GLY  GLU  ILE
...A A G A C G T A T C C A A A A C G T T G C G G C A G G C G A A A T
... 2670 2680 2690 2700

SER  ALA  THR  SER  THR  ASP  ALA  ILE  ASN...
T T C C G C T A C T T C C A C C G A T G C G A T T A A ...
2710
...

```

FIG.18N

```

...   GLY   SER   GLN   LEU   TYR   ALA   VAL   ALA   LYS   GLY   VAL
...C G G C A G C C A G T T G T A T G C C G T G G C A A A G G G G T
... 2730                                2750                2760

THR   ASN   LEU   ALA   GLY   GLN   VAL   ASN   LYS...
A C A A A C C T T G C T G G A C A A G T G A A T A A ...
2770                                2780

...   VAL   GLY   LYS   ARG   ALA   ASP   ALA   GLY   THR   ALA   SER
...A G T G G G C A A A C G T G C A G A T G C A G G T A C A G C A A G
... 2790                                2800                2810                2820

ALA   LEU   ALA   ALA   SER   GLN   LEU   PRO   GLN...
T G C A T T A G C G G C C T T C A C A G T T A C C A C A ...
2830                                2840

...   ALA   SER   MET   SER   GLY   LYS   SER   MET   VAL   SER   ILE
...A G C C T C T A T G T C A G G T A A A T C A A T G G T T T C T A T
... 2850                                2860                2870                2880

ALA   GLY   SER   SER   TYR   GLN   GLY   GLN   SER...
T G C G G G A A G T A G T T A T C A A G G T C A A A G ...
2890                                2900

...   GLY   LEU   ALA   ILE   GLY   VAL   SER   ARG   ILE   SER   ASP
...T G G T T T A G C T A T C G G G G T A T C A A G A A T T T C C G A
... 2910                                2920                2930                2940

```

FIG. 180

```

ASN  GLY  LYS  VAL  ILE  ARG  LEU  SER...
T A A T G G C A A A G T G A T T A T T C G C T T G T C ...
2950
... GLY  THR  THR  ASN  SER  GLN  GLY  LYS  THR  GLY  VAL
...A G G C A C A C C A A T A G C C A A G G T A A A C A G G C G T
... 2970 2980 2990 3000

ALA  ALA  GLY  VAL  GLY  TYR  GLN  TRP  ***
T G C A G C A G G T G T T G G T T A C C A G T G G T A ...
3010 3020 ...

...A T A G A A T T C
... 3030

```

FIG.19A

NTHi-Stamm 32 *hia*

```

G A A T C G G C T T T A A T A T A A G G T A A A T A A A ...
10                               20 30 ...
                               MET ASN LYS ILE PHE ASN VAL ILE TRP ASN
                               ...A A T G A A C A A A T T T A A C G T T A T T G G A A
                               ...                               40 50 60

VAL VAL THR GLN THR TRP VAL VAL VAL SER...
T G T T G T G A C T C A A A C T T G G G T T G T C G T A T C ...
70                               80 90 ...
                               ... GLU LEU THR ARG THR HIS THR LYS CYS ALA
                               ...T G A A C T C A C T C G C A C C C A C A C C A A A T G C G C
                               ...                               100 110 120

SER ALA THR VAL ALA VAL ALA VAL LEU ALA...
C T C C G C C A C C G T G G C A G T T G C C G T A T T G G C ...
130                               140 150 ...
                               ... THR LEU LEU SER ALA THR VAL GLN ALA ASN
                               ...A A C C C T G T T G T C C G C A C C G T T C A G G C G A A
                               ...                               160 170 180

ALA THR ASP GLU ASN GLU ASP GLU GLU...
T G C T A C C G A T G A A A C G A A G A T G A T G A G A ...
190                               200 210 ...

```

FIG.19B

```

...   GLU  LEU  GLU  PRO  VAL  GLN  ARG  SER  VAL  LEU
...A G A G T T A G A A C C C G T A C A A C G C T C T G T T T
...                                     220      230      240

ARG  TRP  SER  PHE  LYS  SER  ALA  LYS  GLU  GLY...
A A G T G G A G C T T C A A A T C C G C T A A G G A A G G ...
250                                     260      270 ...

...   THR  GLY  GLU  GLN  GLU  GLY  THR  THR  GLU  VAL
...C A C T G G A G A A C A A G A G G G A A C A C A G A G G T
...                                     280      290      300

ILE  ASN  LEU  ASN  THR  ASP  SER  SER  GLY  ASN...
A A T A A T T G A A C A C A G A T T C A T C A G G A A A ...
310                                     320      330 ...

...   ALA  VAL  GLY  SER  SER  THR  ILE  THR  PHE  LYS
...T G C A G T A G G A A G C A G C A C A A T C A C C T T C A A
...                                     340      350      360

ALA  GLY  ASP  ASN  LEU  LYS  ILE  LYS  GLN  SER...
A G C C G G C G A C A A C C T G A A A A T C A A A C A A A G ...
370                                     380      390 ...

...   GLY  ASN  ASP  PHE  THR  TYR  SER  LEU  LYS  LYS
...C G G C A A T G A C T T C A C C T A C T C G C T G A A A A
...                                     400      410      420

```

FIG.19C

```

GLU  LEU  LYS  ASN  LEU  THR  SER  VAL  GLU  THR...
A G A G C T G A A A A C C T G A C C A G T G T T G A A A C ...
430
...
... GLU  LYS  LEU  SER  PHE  GLY  ALA  ASN  GLY  ASN
...T G A A A A T T A T C G T T T G G C G C A A A C G G C A A
440
...
450 ...
460
470
480

```

```

LYS  VAL  ASP  ILE  THR  SER  ASP  ALA  ASN  GLY...
T A A G T T G A T A T T A C C A G T G A T G C A A A T G G ...
490
...
... LEU  LYS  LEU  ALA  LYS  THR  GLY  ASN  GLY  ASN
...C T T G A A A T T G G C G A A A C A G G T A A C G G A A A
500
...
510 ...
520
530
540

```

```

GLY  GLN  ASN  SER  ASN  VAL  HIS  LEU  ASN  GLY...
T G G T C A A A C A G T A A T G T T C A C T T A A C G G ...
550
...
... ILE  ALA  SER  THR  LEU  THR  ASP  THR  LEU  ALA
...T A T T G C T T C G A C T T T G A C C G A T A C G C T T G C
560
...
570 ...
580
590
600

```

```

GLY  GLY  THR  THR  GLY  HIS  VAL  ASP  THR  ASN...
C G G T G G C A C A C A G G A C A C G T T G A C A C C A A ...
610
620
630 ...

```

FIG.19D

```

...   ILE  ASP  ALA  VAL  ASN  TYR  HIS  ARG  ALA  ALA
...CATGTGATGCGGTTAATAATATCATCGCGCTGC
...
640
650
660

SER  VAL  GLN  ASP  VAL  LEU  ASN  SER  GLY  TRP...
AAGCGTACAAGATGTGTTAAACAGCGGTG...
670
680
690
...   ASN  ILE  GLN  GLY  ASN  GLY  ASN  ASN  VAL  ASP
...GAATAATCCCAAGGCCAATGGGAACAATGTCGA
700
710
720

PHE  VAL  ARG  THR  TYR  ASP  THR  VAL  ASP  PHE...
TTTGTCCGTACTTACGACACCGTGGACTT...
730
740
750
...   VAL  ASN  GLY  ALA  ASN  ALA  ASN  VAL  SER  VAL
...TGTCAAATGGCGCGCAATGCCAATGTGAGCGT
760
770
780

THR  ALA  ASP  THR  ALA  HIS  LYS  LYS  THR  THR...
TACGGCTGATACGGCTCACAAAGACAC...
790
800
810
...   VAL  ARG  VAL  ASP  VAL  THR  GLY  LEU  PRO  VAL
...TGTCGGTGTGGAATGTACAGGCTTGCCGGT
820
830
840

```


FIG.19E

```

GLN  THR  VAL  THR  GLU  ASP  GLY  LYS  THR  VAL...
T C A A T A T G T T A C G G A A G A C G G C A A A A C C G T ...
850
... VAL  LYS  VAL  GLY  ASN  GLU  TYR  TYR  LYS  ALA
...T G T G A A A G T G G G C A A T G A G T A T T A C A A A G C
880
...
890
900

LYS  ASP  ASP  GLY  SER  ALA  ASP  MET  ASN  GLN...
C A A G A T G A C G G T T C G G C G G A T A T G A A T C A ...
910
... LYS  VAL  GLU  ASN  GLY  GLU  LEU  ALA  LYS  THR
...A A A G T C G A A A C G G C G A G C T G G C G A A A C
940
...
950
960

LYS  VAL  LYS  LEU  VAL  SER  ALA  SER  GLY  THR...
C A A G T G A A A T T G G T A T C G G C A A G C G G T A C ...
970
... ASN  PRO  VAL  LYS  ILE  SER  ASN  VAL  ALA  ASP
...A A T C C G G T G A A A T T A G C A A T G T T G C A G A
1000
...
1010
1020

GLY  THR  GLU  ASP  THR  ASP  ALA  VAL  SER  PHE...
C G G C A C G G A A G A C A C C G A T G C G G T C A G C T T ...
1030
1040
1050 ...

```

FIG.19F

```

...   LYS   GLN   LEU   LYS   ALA   LEU   GLN   ASP   LYS   GLN
...T A G C A A T T A A A G C C T T G C A A G A C A A C A
...
1060
1070
1080

VAL  THR  LEU  SER  THR  SER  ASN  ALA  TYR  ALA...
G G T T A C G T T G A G C A C G A G C T T A T G C ...
1090
1100
1110 ...
...   ASN   GLY   GLY   THR   ASP   ASN   ASP   GLY   GLY   LYS
...C A A T G G C G G T A C A G A T A A C G A C G G C G G C A A
...
1120
1130
1140

ALA  THR  GLN  THR  LEU  SER  ASN  GLY  LEU  ASN...
G G C A A C T C A A C T T A G C C A A T G G T T T G A A ...
1150
1160
1170 ...
...   PHE   LYS   PHE   LYS   SER   SER   ASP   GLY   GLU   LEU
...T T T A A A T T T A A A T C T A G C G A T G G C G A G T T
...
1180
1190
1200

LEU  LYS  ILE  SER  ALA  THR  GLY  ASP  THR  VAL...
G T T G A A A A T T A G C G C G A C C G G C G A T A C G G T ...
1210
1220
1230 ...
...   THR   PHE   THR   PRO   LYS   LYS   GLY   SER   VAL   GLN
...T A C T T T T A C G C C G A A A A A G G T T C G G T A C A
...
1240
1250
1260

```

FIG.19G

```

VAL  GLY  ASP  ASP  GLY  LYS  ALA  SER  ILE  SER...
GGT TGG CGA TGA TGG CAA GGC TTCA ATTTC ...
1270
...  LYS  GLY  ALA  ASN  THR  GLU  GLY  LEU  VAL
...A A A GGTG CAAATA CACA CTGAAG GTTGGT
1280
...  1300
1310
1320

```

```

GLU  ALA  SER  GLU  LEU  VAL  GLU  SER  LEU  ASN...
TGA GGC TTCTGA ATTG GTTGAAAG CCTGA A ...
1330
...  LYS  LEU  GLY  TRP  LYS  VAL  GLY  VAL  GLU  LYS
...CAA ACTGG GTTGGA AAGTAG GGGTTGAGAA
1340
...  1360
1370
1380

```

```

VAL  GLY  SER  GLY  GLU  LEU  ASP  GLY  THR  SER...
AGTC GGC AGCG CGAG CTTG ATGT GTACATC ...
1390
...  LYS  GLU  THR  LEU  VAL  LYS  SER  GLY  ASP  LYS
...CAAG GAACCTTTAG TGAAGTCGGCGGATAA
1400
...  1420
1430
1440

```

```

VAL  THR  LEU  LYS  ALA  GLY  ASP  ASN  LEU  LYS...
AGTA ACTTTGA AAGCCGGCGGACAACTGA A ...
1450
1460
1470 ...

```

FIG.19H

```

... VAL LYS GLN GLU GLY THR ASN PHE THR TYR
...GGTCAACAAGAGGGCCACAACTTCACTTA 1500
... 1480 1490

ALA LEU LYS ASP GLU LEU THR GLY VAL LYS...
CGCGCTCAAGAATGAAATTGACGGCGTGAA ... 1510
1520 ... 1530 ...
... SER VAL GLU PHE LYS ASP THR ALA ASN GLY
...GAGCGTGGAAGTTTAAAGACACGGCGAATGG 1540 1550 1560
... 1540

ALA ASN GLY ALA SER THR LYS ILE THR LYS...
TGCAACGGGTGCAAGCACGAATTACCA ... 1570
1580 ... 1590 ...
... ASP GLY LEU THR ILE THR LEU ALA ASN GLY
...AGACGGCTTGACCAATTACGGCTGGCAACGGG 1600 1610 1620
... 1600

ALA ASN GLY ALA THR VAL THR ASP ALA ASP...
TGCGAATGGTGCGACGGTGACTGATGCCGA ... 1630
1640 ... 1650 ...
... LYS ILE LYS VAL ALA SER ASP GLY ILE SER
...CAAGATTAAAGTTGCTTCGGACGGCAATTAG 1660 1670 1680
... 1660

```

FIG.19I

```

ALA  GLY  ASN  LYS  ALA  VAL  LYS  ASN  VAL  ALA...
CGCGGGTAAATAAGCAGTTAAACGTCGC ...
1690
...      ALA  GLY  GLU  ILE  SER  ALA  THR  SER  THR  ASP
...GGCAGCGGAATTTCTGCCACCTTCCACCGA
1700
...      1720
...      1730
...      1740

ALA  ILE  ASN  GLY  SER  GLN  LEU  TYR  ALA  VAL...
TGGATTAAACGGAGCAGTTGTATGCCGT ...
1750
...      1760
...      1770
...      ALA  LYS  GLY  VAL  THR  ASN  LEU  ALA  GLY  GLN
...GGCAAAAGGGGTAAACAACCTTGCTGGACA
1780
...      1790
...      1800

VAL  ASN  ASN  LEU  GLY  LYS  VAL  ASN  LYS...
AGTGAATAATCTTGAGGGCAAGTGATAA ...
1810
...      1820
...      1830
...      VAL  GLY  LYS  ARG  ALA  ASP  ALA  GLY  THR  ALA
...AGTGGCAAAACGTGCAGATGCAGGTACTGC
1840
...      1850
...      1860

SER  ALA  LEU  ALA  ALA  SER  GLN  LEU  PRO  GLN...
AGTGCAATTAGCGGCTTCACAGTTACCA ...
1870
...      1880
...      1890

```

FIG.19J

```

... ALA THR MET PRO GLY LYS SER MET VAL SER
...AGCCACTATGCCAGGTAAATCAATGGTTTC
...
1900
1910
1920

ILE ALA GLY SER SER TYR GLN GLY ASN...
TATTCGGGGAAGTAGTTATCAAGGTCAAA...
1930
1940
1950 ...
... GLY LEU ALA ILE GLY VAL SER ARG ILE SER
...TGGTTTAGCTATCGGGGTATCAAGATTTC
...
1960
1970
1980

ASP ASN GLY LYS VAL ILE ILE ARG LEU SER...
CGATAATGGCAAGTAGATTATTCGCTTGTCT...
1990
2000
2010 ...
... GLY THR THR ASN SER GLN GLY LYS THR GLY
...AGGCACACCAATAGTCAAGGTAAACAGG
...
2020
2030
2040

VAL ALA ALA GLY VAL GLY TYR GLN TRP ***
CGTTGCAGCAGGTGTTGGTTACCAGTGGTA...

...ATAGAAATTC

```

FIG.20A

NTHi-Stamm 29 Hia

```

      MET ASN LYS ...
T T A A T A T A A G G T A A A T A A A T G A A C A A ...
10      20      30...
      ... ILE PHE ASN VAL ILE TRP ASN VAL VAL THR
      ... A T T T T A A C G T T A T T T G G A A T G T T G T G A C T
      ...      40      50      60

      GLN THR TRP VAL VAL VAL SER GLU LEU THR ...
C A A C T T G G G T T G T C G T A T C T G A A C T C A C T ...
70      80      90...
      ... ARG ALA HIS THR LYS CYS ALA SER ALA THR
      ... C G C G C C C A C A C C A A A T G C G C C T C C G C C A C C
      ...      100      110      120

      VAL ALA VAL ALA VAL LEU ALA THR ALA LEU ...
G T G C C G G T T G C C G T A T T G G C A A C T G C G T T G ...
130      140      150...
      ... SER ALA THR ALA GLU ALA ASN ASN ASN THR
      ... T C T G C A A C G G C T G A G C G A A C A C A A T A C T
      ...      160      170      180

      SER VAL THR ASN GLY LEU ASN ALA TYR GLY ...
T C T G T T A C G A A T G G G T T G A A T G C T T A T G G C ...
190      200      210...

```

FIG.20B

```

... ASP THR ASN PHE ASN THR THR ASN ASN SER
... GATACATAATTATAACCAATAATTCG
... 220 230 240

ILE ALA ASP LEU GLU LYS HIS VAL GIN ASP ...
ATAGCAGATTGGAATAACAGTTCAAGAT...
250 260 270...
... ALA TYR LYS LYS GLY LEU LEU ASN LEU ASN GLU
... GCTTATAAAGGCTTTATAATACTGAATGAA
... 280 290 300

LYS ASP THR ASN LYS SER SER PHE LEU VAL ...
AAGATACAAATAAGTCAAGTTCTTGTT...
310 320 330...
... ALA ASP ASN THR ALA ALA THR VAL GLY ASN
... GCCGACAAATACCCGCCCAACCGTAGGCAAT
... 340 350 360

LEU ARG LYS LYS LEU GLY TRP VAL LEU SER SER ...
TTGCGTAAATGGGCTGGGTATTGTCTAGC...
370 380 390...
... LYS ASN GLY THR ARG ASN GLU LYS SER TYR
... AAAACGGCCACAGGAACGAGAAAGCTAT
... 400 410 420

```


FIG.20C

```

GLN VAL LYS GLN ALA ASP GLU VAL LEU PHE ...
C A G T A A A C A A G C T G A T G A A G T T C T C T T T ...
430                                     440 450...
... THR GLY SER GLY ALA THR VAL SER SER
... A C T G G A T C T G G T G C C A C G G T T A G T T C C
...                                     460 470 480

SER SER LYS ASP GLY LYS HIS THR ILE THR ...
A G C T C T A A A G A C G G T A A A C A T A C C A T T A C C ...
490                                     500 510...
... ILE SER VAL THR LYS GLY SER PHE ALA GLU
... A T T C T G T T A C C A A A G G T A G T T T G C T G A G
...                                     520 530 540

VAL LYS THR ASP ALA THR THR GLY GLY GLN ...
G T A A A A C T G A T G C A A C T A C T G G A G G T C A A ...
550                                     560 570...
... VAL ASN ALA ASP ARG GLY LYS VAL LYS ALA
... G T A A A C G C C G A C C G T G G T A A A G T G A A A G C T
...                                     580 590 600

GLU ASP GLU ASN GLY ALA ASP VAL ASP LYS ...
G A G G A C G A G A A T G G A G C T G A T G T T G A T A A G ...
610                                     620 630...

```

FIG.20D

... LYS VAL ALA THR VAL LYS ASP VAL ALA LYS
 ... A A A G T T G C C A A C T G T A A A G A T G T T G C T A A G
 ... 640 650 660

ALA ILE ASN ASP ALA ALA THR PHE VAL LYS ...
 G C G A T T A A C G A T G C C G C A A C T T T C G T G A A A ...
 ... 670 680 690...
 ... VAL GLU SER THR ASP ASP ASP ILE GLU ASN
 ... G T G G A A A G C A C A G A T G A T G A C A T T G A A A A T
 ... 700 710 720

GLY ALA ALA GLY LYS ASN GLU THR THR ASP ...
 G G T G C T G C A G G C A A A A T G A A C T A C A G A C ...
 ... 730 740 750...
 ... GLN ALA LEU LYS ALA GLY ASP THR LEU THR
 ... C A A G C T C T C A A A G C A G G C G A C A C C T T A A C C
 ... 760 770 780

LEU LYS ALA GLY LYS ASN LEU LYS ALA LYS ...
 T T A A A G C G G G T A A A A C T T A A A A G C T A A G ...
 ... 790 800 810...
 ... LEU ASP GLN ASN GLY LYS SER VAL THR PHE
 ... T T A G A C C A A A A T G G T A A A T C A G T A A C C T T T
 ... 820 830 840

FIG.20E

ALA LEU ALA LYS ASP LEU ASP VAL THR SER ...
 G C T T A G C G A A A G A C C T T G A T G T G A C C T C T ...
 850 860 870...
 ... ALA LYS VAL SER ASP LYS LEU SER ILE GLY
 ... G C G A A A G T G A G T G A T A A G T T G T C T A T T G G T
 ... 880 890 900

LYS ASP THR ASN LYS VAL ASP ILE THR SER ...
 A A G A T A C G A A T A A A G T T G A T A T T A C C A G T ...
 910 920 930...
 ... ASP ALA ASN GLY LEU LYS LEU ALA LYS THR
 ... G A T G C A A A T G G C T T G A A A T T G G C G A A A C A
 ... 940 950 960

GLY ASN GLY ASN GLY GLN ASN GLY ASN VAL ...
 G G T A A C G G A A A T G G T C A A A C G G T A A T G T C ...
 970 980 990...
 ... HIS LEU ASN GLY ILE ALA SER THR LEU THR
 ... C A C T T A A A T G G T A T T G C T T C G A C T T T G A C C
 ... 1000 1010 1020

ASP THR ILE THR GLY MET THR THR GLN ALA ...
 G A T A C C A T T A C A G G T A T G A C A A C A C A G C A ...
 1030 1040 1050...

FIG.20F

```

... SER ASN GLY VAL ALA VAL GIN ASN HIS ASN
... A G C A A T G G C G T G G C T G T G C A G A A T C A T A A T
...
1060
1070
1080

ARG ALA ALA SER VAL ALA ASP VAL LEU ASN ...
C G T G C T G C G A G T G T G G C T G A T G T A T A A T ...
1090
1100
1110...
... ALA GLY TRP ASN ILE GIN GLY ASN GLY ALA
... G C A G G C T G G A A T A T T C A A G G C A A C G G A G C G
...
1120
1130
1140

SER VAL ASP PHE VAL ASN ALA TYR ASP THR ...
A G C G T T G A T T T T G T C A A T G C T T A C G A C A C A ...
1150
1160
1170...
... VAL ASP PHE VAL ASN GLY THR ASN THR ASN
... G T A G A T T T T G T C A A T G G T A C A A C A C C A A T
...
1180
1190
1200

VAL ASN VAL THR THR ASP THR ALA HIS LYS ...
G T G A A C G T T A C G A C T G A T A C G G C T C A C A A A ...
1210
1220
1230...
... LYS THR THR VAL ARG VAL ASP VAL THR GLY
... A A G A C A A C C G T C C G T G G T G G A T G T A C A G G C
...
1240
1250
1260

```

FIG.20G

```

LEU  PRO  VAL  GLN  TYR  VAL  THR  GLU  ASP  GLY  ...
TTGCCGGTTTCAATAATGTTACGGAGAAGCGGC...
1270
...  LYS  THR  VAL  VAL  LYS  VAL  ASP  ASN  LYS  TYR
1280
...  AAAACCGTTGTGAAAGTGGACAATAAGTAT
1300
...
1310
1320

TYR  GLU  ALA  LYS  GLN  ASP  GLY  SER  ALA  ASP  ...
TAGGAAGCTAAGCAAGACGGTTTCGGCGGAT...
1330
...  MET  ASP  LYS  LYS  VAL  GLU  ASN  GLY  GLU  LEU
1340
...  ATGGATAAATAAGTCCGAAATAATGGCGAGCTG
1360
...
1370
1380

ALA  LYS  THR  LYS  VAL  LYS  LEU  VAL  SER  ALA  ...
GCCAAACCAAGTGAAATTTGGTGTCGGCA...
1390
...  SER  GLY  GLN  ASN  PRO  VAL  LYS  ILE  SER  ASN
1400
...  AGCGGTCAATAATCCGGTGAAATAATCAGCAAT
1420
...
1430
1440

VAL  ALA  GLU  GLY  THR  GLU  GLN  ASP  ALA  ...
GTTGCCGAAGGCCACGGACAAGAAACGATCGG...
1450
1460
1470...

```

FIG.20H

```

... VAL SER PHE LYS GLN LEU ALA LEU GLN
... GTCAGCTTTTAAGCAATTGAAAGCCCTTGCAA
... 1480 1490 1500

GLU LYS GLN VAL THR LEU THR ALA SER ASN ...
GAGAAACAGGTTACTTTAACTGCGAGCAAT...
1510 1520 1530...
... ALA TYR ALA ASN GLY GLY ASN ASP ALA ASP
... GCTTATGCCCAATGGTGGTTAACGATGCCGAC
... 1540 1550 1560

GLY GLY LYS ALA THR GLN THR LEU ASN ASN ...
GGCGGCAAGGCAACTCAACCTTTAAACAAT...
1570 1580 1590...
... GLY LEU ASN PHE LYS PHE LYS SER THR ASP
... GGTTTGAATTTTTAAATTTAATCCACAGAC
... 1600 1610 1620

GLY GLU LEU LEU ASN ILE LYS VAL GLU ASN ...
GGCGAGTTGTTGAAACAATCAAGTAGAAAT...
1630 1640 1650...
... ASP THR VAL THR PHE THR PRO LYS LYS GLY
... GACACAGTTACCTTTACGCCGAAATAAGGT
... 1660 1670 1680

```

FIG.20I

```

SER VAL GLN VAL GLY GLU ASP GLY LYS ALA ...
TCGGTACAGGT TGGCGAAGACGGTAAGGCT...
1690
... THR ILE GLN ASN GLY THR LYS THR ASP
... ACGATTCAAAATGGTACGAAACACCAACGAC
1700
...
1720
1730
1740

GLY LEU VAL GLU ALA SER GLU LEU VAL GLU ...
GGTTGGTTGAAGCTTCCGAATTGGTTGA...
1750
... SER LEU ASN LYS LEU GLY TRP LYS VAL GLY
... AGCCTGAACAACA CTGGGCTGGAAAGTGGGC
1760
...
1780
1790
1800

VAL ASP LYS ASP GLY SER GLY LEU ASP ...
GTGATAAAGACGGCAGCGGCGAGCTTGAT...
1810
... GLY ALA SER ASN GLU THR LEU LYS SER
... GGTCATCCCAATGAACCTTAGTGAGTCG
1820
...
1840
1850
1860

GLY ASP LYS VAL THR LEU LYS ALA GLY GLU ...
GGCGATAAAGTACTTGTGAAGCCGGCGAG...
1870
1880
1890

```

FIG.20J

```

... ASN LEU LYS VAL LYS GLN ASP GLY THR ASN
... A A T C T G A A G G T C A A C A A G A C G G C A C A A C
...
1900
1910
1920

PHE THR TYR ALA LEU LYS ASP GLU LEU THR ...
T T C A C T T A C G C G C T C A A A G A T G A A T T G A C G ...
1930
1940
1950...
... GLY VAL LYS SER VAL GLU PHE LYS ASP THR
... G G C G T G A A G A G C C G T G G A G T T T A A A G A C A C G
...
1960
1970
1980

ALA ASN GLY SER ASN GLY ALA SER THR LYS ...
G C G A A T G G T T C A A A C G G T G C A A G C A C G A A G ...
1990
2000
2010...
... ILE THR LYS ASP GLY LEU THR ILE THR SER
... A T T A C C A A A G A C G G C T T G A C C A T T A C G T C G
...
2020
2030
2040

ALA ASN GLY ALA ASN GLY ALA ALA THR ...
G C A A A C G G T G C G A A T G G T G C G G C G G C A C T ...
2050
2060
2070...
... ASP ALA ASP LYS ILE LYS VAL ALA SER ASP
... G A T G C G G A C A A G A T T A A A G T G G C T T C A G A C
...
2080
2090
2100

```


FIG.20K

```

GLY  ILE  SER  ALA  GLY  ASN  LYS  ALA  VAL  LYS  ...
GGCATCAGTGCGGGTAATAAGCGGTTAA...
2110
...  ASN  VAL  VAL  SER  GLY  LEU  LYS  LYS  PHE  GLY
2120
...  AACGTTGTGAGCGGACTGAAGAAATTGGT
2130...
...
2140
2150
2160

ASP  ALA  ASN  PHE  ASN  PRO  LEU  THR  SER  SER  ...
GATGGCAATTTCATAATCCACTGACCAAGTTC...
2170
...  ALA  ASP  ASN  LEU  THR  LYS  GLN  TYR  ASP  ASP
2180
...  GCCGACAACTTACGAAACCAATATGACGAT
2190...
...
2200
2210
2220

ALA  TYR  LYS  GLY  LEU  THR  ASN  LEU  ASP  GLU  ...
GCCATAAGGCTTGACCAATTGGAAGA...
2230
...  LYS  GLY  ALA  ASP  LYS  GLN  THR  LEU  THR  VAL
2240
...  AAGGTGCGGACACAGCAACCTGACTGTT
2250...
...
2260
2270
2280

ALA  ASP  ASN  THR  ALA  ALA  THR  VAL  GLY  ASP  ...
GCCACAATAC TGCCGCAACCGTGCGGCAAT...
2290
2300
2310...

```

FIG.20L

... LEU ARG GLY LEU GLY TRP VAL ILE SER ALA
 ... TTGGCGGCTTGGGCTCATTTCTGCG
 ... 2320 2330 2340

ASP LYS THR THR GLY GLU LEU ASN LYS GLU ...
 GACAAACACAGGCGAATCAATAAGGA...
 2350 2360 2370...
 ... TYR ASN ALA GLN VAL ARG ASN ALA ASN GLU
 ... TACAACGGCGCAAGTGCGTAACGCCAATGAA
 ... 2380 2390 2400

VAL LYS PHE LYS SER GLY ASN GLY ILE HIS ...
 GTGAATTCAGAGCGGCAACGGTATCCAT...
 2410 2420 2430...
 ... VAL SER GLY LYS THR VAL ASN GLY ARG ARG
 ... GTTTCGGGTAAACGGTCAACGGTAGGCGC
 ... 2440 2450 2460

GLU ILE THR PHE GLU LEU ALA LYS ASP GLU ...
 GAATTAATTGAAATTGGCTAAAGACGA...
 2470 2480 2490...
 ... ASN ALA ILE ALA PHE GLY TYR GLY SER LYS
 ... AATGCCCATTGCTTTCGGTTATGGCTCAAA
 ... 2500 2510 2520

FIG.20M

ALA LEU ARG ASP ASN THR VAL ALA ILE GLY ...
 GCC TTGCCGCGATACACGGTGGCAATTGGT...
 2530 2540 2550...
 ... THR GLY ASN VAL VAL ASN ALA GLU LYS SER
 ... ACGGCAACGTTGTGAATGCGGAATAATCT
 ... 2560 2570 2580

GLY ALA PHE GLY ASP PRO ASN TYR ILE GLU ...
 GGTGCATTTCGGCGATCCGAACCTACATCGAA...
 2590 2600 2610...
 ... ASP LYS ALA GLY GLY SER TYR ALA PHE GLY
 ... GATAAGCCGGTGGCAGCTACGCTTTCGGT
 ... 2620 2630 2640

ASN ASP ASN ARG ILE THR SER LYS ASN THR ...
 AACGATAACCGTATTACTCTTAATAACACT...
 2650 2660 2670...
 ... PHE VAL LEU GLY ASN GLY VAL ASN ALA LYS
 ... TTGTGTTGGGTAAATGGAGTTAATGCGAA
 ... 2680 2690 2700

TYR LYS ALA ASN GLY ASP VAL ASP THR GLU ...
 TATAAGCCCAATGGAGATGTTGTATACGGA...
 2710 2720 2730...

FIG.20N

```

...   THR   VAL   THR   VAL   LYS   ASP   LYS   ASP   GLY   LYS
...   ACCGTAAACCGTTTAAAGGACAAAGACGGTAA  2750
...                                     2740  2760

GLU   THR   THR   VAL   THR   VAL   PRO   LYS   ALA   LEU   ...
GAGACTACCGTTACTGTTCCTAAAGCGTTA...  2780
2770                                     2790...
...   GLY   ALA   THR   VAL   GLU   ASN   SER   VAL   TYR   LEU
...   GGGGCTACGGTTTGAAACCTCCGTTTATTG  2810
...                                     2800  2820

GLY   ASN   LYS   SER   THR   ALA   THR   LYS   ASP   LYS   ...
GGTAATAATCGACTGCGACAAAGATAAG...  2840
2830                                     2850...
...   GLY   LYS   ASN   LEU   LYS   SER   ASP   GLY   THR   ALA
...   GGTAATAAACCTGAAATCTGTATGGTACGGCG  2870
...                                     2860  2880

GLY   ASN   THR   THR   THR   ALA   GLY   THR   THR   GLY   ...
GGTAACTACACTGCTGGCACAAACGGGT...  2900
2890                                     2910...
...   THR   VAL   ASN   GLY   PHE   ALA   GLY   ALA   THR   ALA
...   ACGGTAAACGGCTTTGCCGGTGCAACGGCG  2930
...                                     2920  2940

```

FIG.200

```

HIS  GLY  ALA  VAL  SER  VAL  GLY  ALA  SER  GLY  ...
C A C G G T G C G G T T T C T G T C G G C G C A A G C G G C ...
2950
...  GLU  GLU  ARG  ARG  ILE  GLN  ASN  VAL  ALA  ALA
...  G A G A A A G A C G T A T C C A A A C G T C G C G G C A
2960
...
2970...
2980
...
3000

GLY  GLU  ILE  SER  ALA  THR  SER  THR  ASP  ALA  ...
G G C G A A A T T T C C G C C A C T T C C A C C G A T G C G ...
3010
...  ILE  ASN  GLY  SER  GLN  LEU  TYR  ALA  VAL  ALA
...  A T T A A C G G C A G C C A G T T G T A T G C T G T G G C A
3020
...
3030...
3040
...
3050
3060

LYS  GLY  VAL  THR  ASN  LEU  ALA  GLY  GLN  VAL  ...
A A G G G T A A C A A A T C T T G C T G G A C A A G T G ...
3070
...  ASN  LYS  VAL  GLY  LYS  ARG  ALA  ASP  ALA  GLY
...  A A T A A A G T G G G C A A A C G T G C A G A T G C A G G T
3080
...
3090...
3100
3110
3120

THR  ALA  SER  ALA  LEU  ALA  SER  GLN  LEU  ...
A C A G C A A G T G C A T T A G C A G C T T C A C A G T A ...
3130
3140
3150...

```

FIG.20P

```

... PRO GLN ALA SER MET PRO GLY LYS SER MET
... C C A C A G C C T C T A T G C C A G G T A A A T C A A T G
... 3160 3170 3180

VAL SER ILE ALA GLY SER TYR GLN GLY ...
G T T C T A T T G C G G G A A G T A G T T A T C A A G G T ...
3190 3200 3210...
... GLN ASN GLY LEU ALA ILE GLY VAL SER ARG
... C A A A A T G G T T T A G C T A T C G G G G T A T C A C G A
... 3220 3230 3240

ILE SER ASP ASN GLY LYS VAL ILE ILE ARG ...
A T T C C G A T A A T G G C A A A G T G A T T A T T C G C ...
3250 3260 3270...
... LEU SER GLY THR THR ASN SER GLN GLY LYS
... T T G T C A G G C A C A C C A A T A G C C A A G G T A A A
... 3280 3290 3300

THR GLY VAL ALA ALA GLY VAL GLY TYR GLN ...
A C A G G C G T T G C A G C A G G T G T T G G T T A C C A G ...
3310 3320 3330...
... TRP ***
... T G G T A A T A G A A T T C C G G A T C C G C
... 3340 3350

```

FIG.21A

NTHi-Stamm M4071 Hia

```

      MET ASN LYS ILE PHE ASN VAL...
GCGAATTCAATGAAACAATAATTTTAACGT...
10                                     20 30 ...
      ... ILE TRP ASN VAL MET THR GLN THR TRP ALA
      ...TATTTGGAATGTTATGACTCAAACTTGGGC
40                                     50 60
      ...

      VAL VAL SER GLU LEU THR ARG ALA HIS THR...
TGTCGTATCTGAACCTCACTCGCGCCACAC...
70                                     80 90 ...
      ... LYS ARG ALA SER ALA THR VAL ALA THR ALA
      ...CAACGTCGCCCTCCGCAACCGTGCGCAACCGC
100                                    110 120
      ...

      VAL LEU ALA THR LEU LEU SER THR THR VAL...
CGTATTGGCGACGTTGTGTCTACAAACAGT...
130                                    140 150 ...
      ... GLN ALA THR THR THR GLY GLY THR THR SER
      ...TCAGCGACAACCTACTGGCGGTACGACAAG
160                                    170 180
      ...

      THR ASN GLY LEU LYS ALA TYR GLY SER THR...
TACAACGGTTTGAAAGCTTATGGAGTAC...
190                                    200 210 ...

```

FIG.21B

```

... ASN ASN PRO ASN PHE ASN ALA ALA GLY ASN
...G A A T A T C C G A A T T T C A A T G C T G C A G G T A A
...
220
230
240

SER ALA THR ASP LEU ALA ARG GLN PHE ASP...
C T C T G C A A C T G A T T A G C T A G A C A G T T T G A ...
250
260
270 ...
... GLY ALA TYR ASP GLY LEU LEU ASN LEU ASN
...T G G T G C T T A T G A C G G T T A T T A A T C T A A A
280
290
300

GLU LYS ASP ALA ASN LYS ASN LEU LEU VAL...
T G A A A A G A T G C G A A T A A A A T C T G T T G G T ...
310
320
330 ...
... THR ASP ASP LYS ALA ALA THR VAL GLY ASN
...G A C T G A T G A T A A G G C G G C C G T A G G C A A
340
350
360

LEU ARG LYS LYS LEU GLY TRP VAL LEU SER SER...
T T T G C G T A A A T T G G G T T G G G T A T T G T C T A G ...
370
380
390 ...
... LYS ASN GLY THR ARG ASN GLU LYS SER GLN
...T A A A A C G G C A C A G G A A C G A G A A A G C C A
400
410
420

```


FIG.21C

GLN VAL LYS HIS ALA ASP GLU VAL LEU PHE...
 A C A G T C A A A C A C G C G G A T G A A G T G T T ...
 430
 ... GLU GLY LYS ASP GLY VAL THR VAL THR SER
 ...T G A A G G C A A A G A C G G T G T A A C G G T T A C T T C
 460 470 480
 ...

LYS SER GLU ASN GLY LYS HIS THR VAL THR...
 C A A T C T G A A A C G G T A A A C A C C G T T A C ...
 490 510 ...
 ... PHE THR LEU GLU LYS ASP LEU ASN VAL LYS
 ...T T T T A C C C T T G A G A A G A C C T T A A T G T A A A
 520 530 540
 ...

ASN ALA THR VAL SER ASP LYS LEU SER LEU...
 A A C G C A A C C G T T A G C G A T A A T T A T C G C T ...
 550 570 ...
 ... GLY ALA ASN GLY ASN LYS VAL ASP ILE THR
 ...T G G T G C A A A C G G C A A T A A A G T C G A T A T T A C
 580 590 600
 ...

SER ASP THR ASN GLY LEU LYS PHE ALA LYS...
 C A G T G A T A C A A C G G C T T G A A A T T T G C G A A ...
 610 620 630 ...

FIG.21D

```

... PRO SER THR ASN GLY GLN ASN GLY ASN VAL
...A C C A A G T A C G A A T G G T C A A A A C G G T A A T G T 660
... 640

HIS LEU ASN GLY ILE ALA SER THR LEU THR...
T C A C T T A A A C G G T A T T G C C C T C T A C C T T A C ... 670
680 ...
... ASP THR ILE THR GLY THR LYS SER ALA
...T G A C A C A A T T A C A G G T A C A C A A A A T C T G C 710
... 700 720

THR ASN GLY VAL ASP VAL GLN ASN HIS ASN...
A A C T A A T G G T G T A G A T G T G C A G A A T C A T A A ... 730
740 ...
... ARG ALA ALA SER VAL ALA ASP VAL LEU ASN
...T C G T G C T G C G A G T G T A G C T G A T G T A T G A A 760
... 770 780

ALA GLY TRP ASN ILE GLN GLY ASN GLY ALA...
T G C A G G C T G G A A T A T T C A A G G C A A C G G A G C ... 790
800 ...
... SER VAL ASP PHE VAL ASN THR TYR ASP THR
...G A G C G T T G A T T T T G T C A A T A C T T A C G A C A C 820
... 830 840

```

FIG.21E

```

VAL  ASP  PHE  VAL  ASN  GLY  LEU  ASN  THR  ASN...
AGT  AGA  TTT  GTCA  ATGG  TTAA  ATACCA  ...
850
...  VAL  ASN  VAL  THR  THR  ASP  THR  ALA  HIS  ASN
...TGTGAACGTTACGACTGATACGGCTCAACA
880
...
LYS  LYS  THR  THR  VAL  ARG  VAL  ASP  VAL  THR...
CAAAAGACACCGTCCGTGTGATGTAAC...
910
...  GLY  LEU  PRO  VAL  GLN  TYR  VAL  THR  GLU  ASP
...GGGCTTGCCGGTCCAAATATGTTACGGAGAGA
940
...
GLY  GLU  THR  VAL  VAL  LYS  VAL  GLY  ASN  GLU...
CGCGGAACCGTTGTGAAGAAGTGGCAATGA...
970
...  TYR  TYR  GLU  ALA  LYS  GLN  ASP  GLY  SER  ALA
...GTATTACGAAGCCAGCAAGACGGTTCGGC
1000
...
ASP  MET  ASP  LYS  LYS  VAL  GLU  ASN  GLY  LYS...
GGATATGGATAAATAAGTCCGAATAATGGCA...
1030
1040
1050
1020

```

FIG.21F

```

...      LEU  ALA  LYS  THR  LYS  VAL  LYS  LEU  VAL  SER
...GCTGGCGA A A A C T A A G T T A A A T T G G T A T C      1080
...
1060
...

ALA  ASN  GLY  THR  ASN  PRO  VAL  LYS  ILE  SER...
GGC A A A C G G T A C A A A T C C G G T G A A A A T C A G ...
1090
1100
...      ASN  VAL  ALA  ASP  GLY  THR  GLU  ASN  THR  ASP
...C A A T G T T G C G G A C G G C A C G G A A A A T A C C G A
1120
1130
...

ALA  VAL  SER  PHE  LYS  GLN  LEU  LYS  ALA  LEU...
T G C G G T C A G C T T T A A G C A G T T G A A A G C C T T ...
1150
1160
...      GLN  ASP  LYS  GLN  VAL  THR  LEU  SER  ALA  SER
...G C A A G A C A A A C A G G T T A C G T T A A G T G C G A G
1180
1190
1200
...

ASN  ALA  TYR  ALA  ASN  GLY  GLY  SER  ASP  ALA...
C A A T G C T T A T G C C A A T G G C G G T A G C G A T G C ...
1210
1220
...      ASP  GLY  GLY  LYS  GLY  ILE  GLN  THR  LEU  SER
...C G A C G G C G C A A G G G A A T T C A A A C T T A A G
1240
1250
1260
...
```

FIG.21G

```

ASN  GLY  LEU  ASN  PHE  LYS  PHE  LYS  SER  THR...
C A A T G G T T T G A A T T T A A A T C C A C ...
1270
...      ASP  GLY  GLU  LEU  LEU  ASN  ILE  LYS  ALA  GLU
...A G A C G C G A G T T G A A T A T C A A A G C A G A
1300      1310      1320
...

ASN  ASP  THR  VAL  THR  PHE  THR  PRO  LYS  LYS...
A A A T G A C A C G G T T A C C T T T A C G C C G A A A A ...
1330      1340      1350 ...
...      GLY  SER  VAL  GLN  VAL  GLY  ASP  ASP  GLY  LYS
...A G G T T C G G T G C A G G T T G G C G A T G A T G G T A A
1360      1370      1380
...

ALA  THR  ILE  GLN  ASP  GLY  ALA  LYS  THR  THR...
G G C T A C G A T T C A A G A C G G C G C A A A A C A C ...
1390      1400      1410 ...
...      THR  GLY  LEU  VAL  GLU  ALA  SER  GLU  LEU  VAL
...T A C C G G T T T G G T T G A G G C T T C T G A A T T G G T
1420      1430      1440
...

ASP  SER  LEU  ASN  LYS  LEU  GLY  TRP  LYS  VAL...
T G A C A G C C T G A A C A A A T T G G G T T G G A A A G T ...
1450      1460      1470 ...

```

FIG.21H

... GLY THR GLY THR ASP GLY THR GLY VAL THR
 ...GGGCCACCGGCACCTGTGACGGCACAGGAGTGAC
 1480 1490 1500
 ...

ASP GLY THR HIS THR ASP THR LEU VAL LYS...
 CGATGGCCACGCATACCGACCTTTAGTGA...
 1510 1520 1530 ...
 ... SER GLY ASP LYS VAL THR LEU LYS ALA GLY
 ...GTCGGCGCATTAAGTAACTTTGAAAGCCGG
 1540 1550 1560
 ...

ASP ASN LEU LYS VAL LYS GLU GLY THR...
 CGACAATCTGAAGGTCAACAAGAGGTAC...
 1570 1580 1590 ...
 ... ASN PHE THR TYR ALA LEU LYS ASP GLU LEU
 ...A A C T T C A C T T A T G C G C T C A A A G A T G A A T T
 1600 1610 1620
 ...

THR ASP VAL LYS SER VAL GLU PHE LYS ASP...
 GACGACGTGAAGAGCGTGGAAGTTTAAAGA...
 1630 1640 1650 ...
 ... THR ALA ASN GLY ALA ASN GLY ALA SER THR
 ...C A C G G C G A A T G G T G C A A C G G T G C A A G C A C
 1660 1670 1680
 ...

FIG.21I

```

LYS  ILE  THR  LYS  ASP  GLY  LEU  THR  ILE  THR...
G A G A T T A C C A A A G A C G G C T T G A C C A T T A C ...
1690
...  PRO  ALA  ASN  GLY  ALA  GLY  ALA  ALA  GLY  ALA
...G C C G G C A A A C G G T G C G G T G C G G C A G G T G C
1700
...  1720
...  1730
...  1740

ASN  THR  ALA  ASN  THR  ILE  SER  VAL  THR  LYS...
A A A C A C T G C A A A C A C C A T T A G C G T A C C A A ...
1750
...  ASP  GLY  ILE  SER  ALA  GLY  ASN  LYS  ALA  VAL
...A G A C G G C A T T A G C G C G G G T A A T A A G C A G T
1760
...  1770
...  1780
...  1790
...  1800

LYS  ASN  VAL  VAL  SER  GLY  LEU  LYS  LYS  PHE...
T A A A A C G T T G T G A G C G G A C T G A A G A A T T ...
1810
...  1820
...  1830
...  GLY  ASP  ALA  ASN  PHE  ASP  PRO  LEU  THR  SER
...T G G T G A T G C G A A T T T C G A T C C G C T G A C T A G
1840
...  1850
...  1860

SER  ALA  ASP  ASN  LEU  THR  LYS  GLN  TYR  ASP...
C T C A G C C G A C A C T T A C G A A A C A A T A T G A ...
1870
...  1880
...  1890
...  1900

```

FIG.21J

```

... ASN ALA TYR LYS GLY LEU THR ASN LEU ASP
...C A A T G C C T A T A A G G C T T G A C C A A T C T G G A
...
1900
1910
1920

GLU LYS SER LYS GLY LYS GLN THR PRO THR...
T G A A A A A G T A A A G G C A A G C A A A C T C C G A C ...
1930
1940
1950 ...
... VAL ALA ASP ASN THR ALA ALA THR VAL GLY
...C G T T G C T G A C A A T A C C G C T G C A A C C G T G G G
1960
1970
1980

ASP LEU ARG ARG GLY LEU GLY TRP VAL ILE SER...
C G A T T T G C G C G G C T T G G G C T G G G T C A T T T C ...
1990
2000
2010 ...
... ALA ASP LYS THR LYS GLY GLU LEU ASN LYS
...T G C A G A C A A A A C C A A A G G C G A A C T C A A T A A
2020
2030
2040

GLU TYR ASN ALA GLN VAL ARG ASN ALA ASN...
G G A T A C A A C G C A C A A G T G C G T A A C G C T A A ...
2050
2060
2070 ...
... GLU VAL LYS PHE LYS SER GLY ASN GLY ILE
...T G A A G T G A A A T T C A A G A G C G G C A A C G G T A T
2080
2090
2100

```


FIG.21K

```

ASN  VAL  SER  GLY  LYS  THR  LEU  ASP  ASN  GLY...
C A A T G T T T C C G G T A A A C A T T G G A T A A C G G ...
2110
...      THR  ARG  GLU  ILE  THR  PHE  GLU  LEU  ALA  LYS
...T A C G C G C G A A A T T A C T T T T G A A T T G G C T A A
2120
...      2140
...      2150
...      2160

```

```

ASP  GLU  ASN  ALA  ILE  ALA  PHE  GLY  SER  GLY...
A G A C G A A A A T G C C A T T G C T T T C G G T T C T G G ...
2170
...      SER  LYS  ALA  LEU  ARG  ASP  ASN  THR  VAL  ALA
...C T C A A A A G C C T T G C G C G A T A A C A C G G T G G C
2180
...      2190
...      2200
...      2210
...      2220

```

```

ILE  GLY  THR  GLY  ASN  VAL  VAL  ASN  ALA  GLU...
A A T T G G T A C G G G C A A C G T T G T G A A T G C G G A ...
2230
...      LYS  SER  GLY  ALA  PHE  GLY  ASP  PRO  ASN  TYR
...A A A T C T G G T G C A T T C G G C G A T C C G A A C T A
2240
...      2250
...      2260
...      2270
...      2280

```

```

ILE  GLU  ASP  LYS  ALA  GLY  GLY  SER  TYR  ALA...
C A T C G A A G A T A A G C C G G T G G C A G C T A C G C ...
2290
...      2300
...      2310
...      2320

```

FIG.21L

```

... PHE GLY ASN ASP ASN ARG ILE THR SER LYS
...T T C G G T A A C G A T A A C C G T A T T A C T T C T A A
... 2320 2330 2340

ASN THR PHE VAL LEU GLY ASN SER VAL ASN...
A A C A C T T T G T G T T G G G T A A T A G T G T T A A ...
2350 2360 2370 ...
... ALA LYS ARG ASP ALA ASN GLY ASN VAL LEU
...T G C G A A A C G T G A T G C A A A T G G C A A T G T A C T
... 2380 2390 2400

THR GLU GLU LYS GLU VAL VAL GLY LYS ASP...
G A C C G A A G A A A A G A A G T G G T T G G A A A G A ...
2410 2420 2430 ...
... GLY ALA LYS THR LYS VAL THR VAL PRO GLN
...C G G T G C G A A G A C G A A A G T A A C C G T G C C G C A
... 2440 2450 2460

ALA LEU GLY GLU THR VAL GLU ASN SER VAL...
A G C C T T A G G C G A A A C C G T A G A A A T T C T G T ...
2470 2480 2490 ...
... TYR LEU GLY ASN ALA SER THR ALA THR LYS
...T T A T C T C G G T A A T G C T T C A A C T G C G A C A A
... 2500 2510 2520

```

FIG.21M

```

ASP  LYS  GLY  LYS  ASN  LEU  LYS  SER  ASP  GLY...
A G A T A A G G G T A A A A C C T G A A A T C T G A T G G ...
2530
...      THR  ALA  GLY  ASN  THR  THR  ALA  GLY  ALA
...T A C G G C G G G T A A C A C T A C A C T G C T G G C G C
2540
...      2560      2570      2580

```

```

THR  GLY  THR  VAL  ASN  GLY  PHE  ALA  GLY  ALA...
A A C G G G T A C G G T A A A C G G C T T T G C C G G T G C ...
2590
...      THR  ALA  HIS  GLY  ALA  VAL  SER  VAL  GLY  ALA
...A A C G G C G C A C G G T G C G G T T T C T G T C G G C G C
2600
...      2620      2630      2640

```

```

SER  GLY  GLU  GLU  ARG  ARG  ILE  GLN  ASN  VAL...
A A G T G G C G A A G A A A G A C G T A T C C A A A C G T ...
2650
...      ALA  ALA  GLY  GLU  ILE  SER  ALA  THR  SER  THR
...C G C G G C A G G C G A A T T T C C G C T A C T T C C A C
2660
...      2680      2690      2700

```

```

ASP  ALA  ILE  ASN  GLY  SER  GLN  LEU  TYR  ALA...
A G A T G C G A T T A A C G G T A G C C A G T T G T A T G C ...
2710
...      2720      2730

```

FIG.21N

```

... VAL ALA LYS GLY VAL THR ASN LEU ALA GLY
...TGTGGCAAAAGGGGTAAACAACCTTGCTGG
... 2740 2750 2760

GLN VAL ASN LYS VAL GLY LYS ARG ALA ASP...
ACAGTGATAAGTGGGCAACGTGCA...
2770 2780 2790 ...
... ALA GLY THR ALA SER ALA LEU ALA ALA SER
...TGCAGGTACAGCAAGTGCAATTAGCGGCTTC
... 2800 2810 2820

GLN LEU PRO GLN ALA SER MET PRO GLY LYS...
ACAGTTACCAACAAGCCCTCTATGCCAGGTA...
2830 2840 2850 ...
... SER MET VAL SER ILE ALA GLY SER SER TYR
...ATCAATGGTTTCTATTGCGGGAAGTAGTTA
... 2860 2870 2880

GLN GLY GLN SER GLY LEU ALA ILE GLY VAL...
TCAGGTCAAGTGGTTTAGCTATCGGGT...
2890 2900 2910 ...
... SER ARG ILE SER ASP ASN GLY LYS VAL ILE
...ATCAAGATAATTCCGATAATGGCAAGTGAT
... 2920 2930 2940

```

FIG.210

ILE	ARG	LEU	SER	GLY	THR	THR	ASN	SER	GLN...																																							
T	A	T	T	C	G	C	T	T	G	T	C	A	G	G	C	A	C	C	A	A	T	A	G	C	C	A	...																					
										2950											2960											2970																
										...	GLY	LYS	THR	GLY	VAL	ALA	ALA	GLY	VAL	GLY																												
										...	A	G	G	T	A	A	A	C	A	G	G	C	G	T	T	G	C	A	G	C	A	G	G	T	G	T	G	G										
										...											2980											2990											3000					
TYR	GLN	TRP	***	***	***	ASN	SER	GLY	SER																																							
T	T	A	C	C	A	G	T	G	G	T	A	A	T	A	G	A	A	T	T	C	C	G	G	A	T	C	C	G	C																			
										3010											3020											3030																

FIG.22A

NTHi-Stamm K9 hia-Sequenz

```

MET ASN LYS ILE PHE ASN VAL ILE TRP ASN ...
A T G A A C A A A T T T T A A C G T T A T T T G G A A T ...
10
... VAL MET THR GLN THR TRP ALA VAL VAL SER
... G T T A T G A C T C A A A C T T G G G C T G T C G T A T C T
20
...
40
50
60

GLU LEU THR ARG ALA HIS THR LYS ARG ALA ...
G A C T C A C T C G C G C C C A C A C C A A C G T G C C ...
70
... SER ALA THR VAL ALA THR ALA VAL LEU ALA
... T C C G C A A C C G T G C G C A C C G C C G T A T T G G C G
80
...
100
110
120

THR GLN LEU SER ALA THR ALA GLU ALA ASN ...
A C G C A G T T G T C T G C A A C G G C T G A A G C G A A C ...
130
140
150...
... SER SER ALA SER VAL THR SER ARG LEU ASN
... A G T A G T G C T T C T G T T A C G A G T A G G T T G A A T
160
170
180

VAL TYR GLY ASP THR ASN THR LYS PHE ASN ...
G T T T A T G G C G A T A C G A A T A C T A A A T T C A A T ...
190
200
210...

```

FIG.22B

... ALA ALA ASN ASN SER ILE ALA ASP LEU ASN
 ... GCAGCCAAATAATTCATAAGCAGATTTAAT
 ... 220 230 240

LYS GLN ASN ASP GLY VAL HIS ASP GLY LEU ...
 AACAATAATGATGGTTTCACGATGGTTTA...
 ... 250 260 270...
 ... LEU ASN LEU ASN GLU ASN GLY ALA ASN LYS
 ... TTAATACTGAATGAACACGGTGCGAATAA
 ... 280 290 300

LYS LEU LEU VAL ASP ASP THR ALA ALA ...
 AGCTGTTGGTGGA TGACAAATAC TGCGCG...
 ... 310 320 330...
 ... THR VAL GLY ASP LEU ARG LYS LEU GLY TRP
 ... ACCGTAGGCGAATTACGTAAATTGGCGCTGG
 ... 340 350 360

VAL VAL SER THR LYS ASN GLY LYS GLU ASN ...
 GTCGTATCAACCAAAATGGCAAGGAATA...
 ... 370 380 390...
 ... GLU LYS SER GLN GLN VAL LYS GLN ALA ASP
 ... GAGAAAGCCACAAGTCAAACAGCGGAT
 ... 400 410 420

FIG.22C

```

GLU VAL LEU PHE LYS GLY SER LYS GLY GLY ...
G A G T G T T G T T T A A G G C A G C A A G G C G G T ...
430
... VAL GLN VAL THR SER THR SER GLU ASN GLY
... G T G C A G G T T A C T T C C A C C T C T G A A A C G G C
440
... 460
... 470
... 480

LYS HIS ALA ILE THR PHE ALA LEU ALA LYS ...
A A C A C G C C A T T A C C T T T G C T T T A G C G A A A ...
490
... 500
... ASP LEU ASP MET ARG THR ALA THR VAL SER
... G A C C T T G A T A T G A G A C T G C G A C T G T G A G T
510
... 520
... 530
... 540

ASP THR LEU THR ILE GLY GLY SER THR THR ...
G A T A C C T T A A C G A T T G G C G G T A G T A C T A C T ...
550
... 560
... THR GLY SER ALA THR THR THR PRO LYS VAL ASN
... A C A G G T A G T G C A C A C A C C A A A G T G A A T
570
... 580
... 590
... 600

VAL THR SER THR ALA SER GLY LEU ASN PHE ...
G T G A C T A G C A C G G C A A G C G G C T T G A A C T T ...
610
... 620
... 630
... 640

```


FIG.22D

```

...   ALA   LYS   GLY   ALA   THR   GLY   ALA   ASN   GLY   ASP
...   GCGAAAGGCGCTACAGGTGCTAATGGCGAT
...                                     650
...
...   THR   VAL   HIS   LEU   THR   ASN   ILE   ALA   SER ...
ACTACGGTTCACTTGACTAATAATTGCTTCA...
680                                     690...
...   THR   LEU   GLN   ASP   THR   LEU   LEU   ASN   THR   GLY
...   ACTTTGCAAGATACTCTATTGAAATACTGGG
...                                     710
...                                     720
...
VAL   VAL   SER   LYS   LEU   ASP   GLY   ASN   GLY   ILE ...
GTTGTGAGTAATAATTAGATGGTAATGGTATT...
730                                     740
...   THR   ALA   ASP   GLU   LYS   LYS   ARG   ALA   ALA   SER
...   ACTGCTGACGAGAAATAACGTGGCGCAAGC
...                                     760
...                                     770
...                                     780
...
VAL   GLN   ASP   VAL   LEU   ASN   SER   GLY   TRP   ASN ...
GTTCAAGATGTTTAAATAAGTGTTGGAAAT...
790                                     800
...   ILE   LYS   GLY   VAL   LYS   THR   GLY   ALA   THR   THR
...   ATCAAGGGTGTTAATAACAAGGTGGCGACGACT
...                                     820
...                                     830
...                                     840
...

```

FIG.22E

```

SER  ASP  ASN  VAL  ASP  PHE  VAL  ARG  THR  TYR  ...
TCTGATACGTTGATTTCCTCCGTACTTAC...
850
...  ASP  THR  VAL  GLU  PHE  LEU  SER  GLY  SER  GLU
...  GACACAGTTGAGTTTTCGAGCGGAGTGAA
880
...
GLU  THR  THR  LEU  VAL  THR  VAL  ASP  SER  GLU  ...
GAACCTACCTGGTTACAGTGGA TAGTGA...
910
...  SER  ASN  GLY  LYS  SER  THR  LYS  VAL  LYS  ILE
...  AGTAATGGAAATACTACTAAGTTTAAATC
940
...
GLY  ALA  LYS  THR  SER  VAL  ILE  LYS  GLU  LYS  ...
GGTGGAGAGACCTCTCTTATCAAGAAATA...
970
...  ASP  GLY  LYS  LEU  PHE  THR  GLY  LYS  ALA  ASN
...  GACGGTAAGTTATTCTGGAAGAGCTAAT
1000
...
LYS  ASP  THR  ASN  GLN  VAL  ALA  SER  ASN  ASN  ...
AAGACACAAATCAAGTCGCAAGTAATAAT...
1030
1040
1050...

```

FIG.22F

```

... ALA ALA ASP ASP THR ASP GLU GLY LYS GLY
... G C A G C T G A T G A T A C G G A T G A G G C C A A G G C
... 1060 1070 1080

LEU VAL THR ALA GLU THR VAL ILE ASN ALA ...
T T A G T C A C T G C A G A G A C T G T T A T C A A T G C A ...
1090 1100 1110...
... VAL ASN LYS ALA GLY TRP ARG ILE LYS THR
... G T A A A C A A G G C T G G T T G G A G A A T T A A A C A
... 1120 1130 1140

THR GLY ALA ASN ASN GLN ALA GLY GLN PHE ...
A C G G G T G C T A A T A A T C A A G C T G G T C A G T T ...
1150 1160 1170...
... GLU THR VAL THR SER GLY THR ASN VAL THR
... G A A A C T G T C A C A T C A G G C A C A A T G T A A C C
... 1180 1190 1200

PHE ALA ASP GLY ASN GLY THR THR ALA VAL ...
T T T G C T G A T G G C A A T G G T A C A C T G C A G T C ...
1210 1220 1230...
... VAL THR GLY ASP ALA THR ASN GLY ILE THR
... G T A A C T G G C G A T G C T A C C A A T G G T A T T A C T
... 1240 1250 1260

```

FIG.22G

```

VAL  LYS  TYR  GLU  ALA  LYS  VAL  GLY  ASP  GLY  ...
G T T A A T A C G A A G C G A A A G T T G G C G A C G G C ...
1270
...  LEU  LYS  ILE  GLY  ASN  ASP  GLN  LYS  ILE  THR
...  T T G A A G A T T G G T A A C G A C C A A A A T C A C T
1280
...  1290...
1300
1310
1320

ALA  ASP  THR  THR  ALA  LEU  THR  VAL  THR  GLY  ...
G C A G A T A C G A C C G C A C T T A C T G T G A C G G C ...
1330
...  GLY  LYS  VAL  THR  ALA  PRO  ASP  ALA  THR  ASN
...  G G T A A A G T T A C T G C C C C T G A T G C A C C A A T
1340
...  1350...
1360
1370
1380

GLY  LYS  LYS  LEU  VAL  ASN  ALA  SER  GLY  LEU  ...
G G T A A G A A A C T T G T T A A T G C A A G T G G T T T A ...
1390
...  ALA  ASP  ALA  LEU  ASN  LYS  LEU  SER  TRP  THR
...  G C T G A T G C G T T A A C A A A T T A A G T T G G A C T
1400
...  1410...
1420
1430
1440

ALA  LYS  ALA  GLU  ALA  ASP  THR  ALA  ASN  GLY  ...
G C A A A G C T G A A G C A G A T A C T G C T A A T G G C ...
1450
...  1460
1470...

```

FIG.22H

... GLY GLU LEU ASP GLY THR ALA ASP GLU LYS
 ... G G C G A G C T T G A T G G A A C T G C A G A T G A A A A
 ... 1480 1490 1500

GLU VAL LYS ALA GLY THR VAL THR PHE ...
 G A A G T T A A A G C A G G C G A A A C G G T A A C C T T T ...
 ... 1510 1520 1530...
 ... LYS ALA GLY LYS ASN LEU LYS VAL LYS GLN
 ... A A A G C G G C A A G A A C T T A A A A G T G A A C A A
 ... 1540 1550 1560

ASP GLY ALA ASN PHE THR TYR SER LEU GLN ...
 G A T G G T G C G A A C T T A C T T A T T C A C T G C A A ...
 ... 1570 1580 1590...
 ... ASP ALA LEU THR GLY LEU THR SER ILE THR
 ... G A T G C T T T A A C A G G C T T A A C G A G C A T T A C T
 ... 1600 1610 1620

LEU GLY THR GLY ASN ASN GLY ALA LYS THR ...
 T T A G G T A C A G G A A A T A A T G G T G C G A A A C T ...
 ... 1630 1640 1650...
 ... GLU ILE ASN LYS ASP GLY LEU THR ILE THR
 ... G A A A T C A A C A A A G A C G G C T T A A C C A T C A C A
 ... 1660 1670 1680

FIG.221

```

PRO  ALA  ASN  GLY  ALA  GLY  ALA  ASN  ASN  ALA  ...
CCAGCAATA TGGTGCGGTGCAATAATGCA...
1690                                1700
...  ASN  THR  ILE  SER  VAL  THR  LYS  ASP  GLY  ILE
...  AACACCATCAGCGTAA CCAAGACGGCAT T
...                                1720
...                                1730
...                                1740

SER  ALA  GLY  GLY  GLN  SER  VAL  LYS  ASN  VAL  ...
AGTGCGGCGGTCAGTCGGTTAAACGTT...
1750                                1760
...  VAL  SER  GLY  LEU  LYS  LYS  PHE  GLY  ASP  ALA
...  GTGAGCGGACTGTGAAGAAATTTGGTGATGCG
...                                1780
...                                1790
...                                1800

ASN  PHE  ASP  PRO  LEU  THR  SER  SER  ALA  ASP  ...
AATTTCGATCCGCTGACTAGCTCCGCGAC...
1810                                1820
...  ASN  LEU  THR  LYS  GLN  TYR  ASP  ASP  ALA  TYR
...  AACTTAACGAACAATA TGACGATGCC TAT
...                                1840
...                                1850
...                                1860

LYS  GLY  LEU  THR  ASN  LEU  ASP  GLU  LYS  GLY  ...
AAGGCTTGACCAATTTGGATGAAGAGT...
1870                                1880
...                                1890
...                                1900

```

FIG.22J

```

... ALA ASP LYS GLN THR LEU THR VAL ALA ASP
... GCGGACAAAGCAAACTCTGACTGTGCCGAC
...
1900
1910
1920

ASN THR ALA ALA THR VAL GLY ASP LEU ARG ...
AATAC TGGCCGCAACCGTG GGGCGATT TGGCGC...
1930
1940
1950...
... GLY LEU GLY TRP VAL ILE SER ALA ASP LYS
... GGC TTGGGCTGGGTCTATTCTGCGGACAAA
1960
1970
1980

THR THR GLY GLU LEU ASP LYS GLU TYR ASN ...
ACCAAGGCGCAACTCGATAAGGAATACAAAC...
1990
2000
2010...
... ALA GLN VAL ARG ASN ALA ASN GLU VAL LYS
... GCGCAAGTGCGGTAAACGCCCAATGAGTGAAA
2020
2030
2040

PHE LYS SER GLY ASN GLY ILE ASN VAL SER ...
TTCAAAGCGGCAACGGTATCAATGTTTC...
2050
2060
2070...
... GLY LYS THR VAL ASN GLY ARG ARG GLU ILE
... GGTAACCTGTCAACGGTAGGCGTGAAATT
2080
2090
2100

```

FIG.22K

```

THR  PHE  GLU  LEU  ALA  LYS  GLY  GLU  VAL  VAL  ...
ACTTTGAAATTGGCTAAAGGCCGAAGTGGTT...
2110
...  LYS  SER  ASN  GLU  PHE  THR  VAL  LYS  GLU  THR
...  AATCGAATGAATTCTCTCAAGAAACC
...
2120
...  LYS  SER  ASN  GLU  PHE  THR  VAL  LYS  GLU  THR
...  AATCGAATGAATTCTCTCAAGAAACC
...
2130...
2140
...
2150
...
2160

ASN  GLY  LYS  GLU  THR  SER  LEU  VAL  LYS  VAL  ...
AATGGCAAGGAACGAGCCCTGGTTTAAAGTT...
2170
...  GLY  ASP  LYS  TYR  TYR  SER  LYS  GLU  ASP  ILE
...  GCGATAATAATTACAGCAAGAGGATATT
...
2180
...  LYS  TYR  TYR  SER  LYS  GLU  ASP  ILE
...  GCGATAATAATTACAGCAAGAGGATATT
...
2190...
2200
...
2210
...
2220

ASP  PRO  ALA  THR  GLY  LYS  PRO  LYS  VAL  THR  ...
GACCAAGCAACCGGTAAACCGAAGTTACA...
2230
...  ASN  GLY  ASN  ALA  VAL  ALA  ALA  LYS  TYR  GLN
...  AATGGCAATGCAAGTTGCTGCCGAATAACA
...
2240
...  ASN  GLY  ASN  ALA  VAL  ALA  ALA  LYS  TYR  GLN
...  AATGGCAATGCAAGTTGCTGCCGAATAACA
...
2250...
2260
...
2270
...
2280

ASP  LYS  ASP  GLY  LYS  VAL  VAL  SER  ALA  ASP  ...
GATAAGATGGCAAGTCTGTTCTCTGCTGAC...
2290
...
2300
...
2310...

```


FIG.22L

... GLY SER SER ASN THR ALA VAL THR LEU THR
 ... GGCAGCAATAACCGCTGTTAACCTAACCC
 ... 2320 2330 2340

ASN LYS GLY TYR GLY TYR VAL THR GLY ASN ...
 AACAAAGGTTATGGCTATGTAAACAGGTAAAC...
 2350 2360 2370...
 ... GIN VAL ALA ASP ALA ILE ALA LYS SER GLY
 ... CAGTGGCAGATGCCGATTGGCAATAATCAGGC
 ... 2380 2390 2400

PHE GLU LEU GLY LEU ALA ASP ALA GLU LYS ...
 TTGAGCTTGGTTTGGCTGTGATGCAGAAATAA...
 2410 2420 2430...
 ... ALA LYS ALA ALA PHE GLY ASP GLU THR LYS
 ... GCGAAGCTGCCGTTTGGCGATGAACAACA
 ... 2440 2450 2460

ALA LEU SER SER ASP LYS LEU GLU THR VAL ...
 GCTTGTCTCTGATAATAATGGGAACCCGTA...
 2470 2480 2490...
 ... ASN ALA ASN ASP LYS VAL ARG PHE ALA ASN
 ... AATGCCAACGACAAAGTCCGTTTTCCTAAT
 ... 2500 2510 2520

FIG.22M

GLY LEU ASN THR LYS VAL SER ALA ALA THR ...
 GGT TTA ATA CCA AAG TGA GCG GCA ACG ...
 2530 2540 2550...
 ... VAL GLU SER ILE ASP ALA ASN GLY ASP LYS
 ... GTG GAAGCATCG ATGCAACGGCGATAA
 ... 2560 2570 2580

VAL THR THR THR PHE VAL LYS THR ASP VAL ...
 GTG ACTACAACCTTTGTGAAACCGATGTG...
 2590 2600 2610...
 ... GLU LEU PRO LEU THR GLN ILE TYR ASN THR
 ... GAATTGCCCTTTACGCAATACTACATACC
 ... 2620 2630 2640

ASP ALA ASN GLY LYS LYS ILE VAL LYS ASN ...
 GATGCAACGGTAAAGAAATCGTTAAAT...
 2650 2660 2670...
 ... GLY ASP LYS TRP TYR THR LYS ASP ASP
 ... GGCGATAAATGGTTATTACCAAGAGATGAC
 ... 2680 2690 2700

GLY SER THR ASP MET THR LYS GLU VAL THR ...
 GGCTCAACTGATATGACTAAGAGATTACC...
 2710 2720 2730...

FIG.22N

... LEU GLY ASN VAL ASP SER ASP GLY LYS LYS
 ... C T T G G C A A T G T G G A T T C A G A C G G C A A G A A A
 ... 2740 2750 2760

VAL VAL LYS GLU ASP ASN LYS TRP TYR HIS ...
 G T T G T G A A A G A A G A C A A C A A G T G G T A T C A C ...
 2770 2780 2790...
 ... VAL LYS SER ASP GLY SER THR ASP LYS THR
 ... G T T A A A T C T G A T G G T T C T A C G G A T A A A C A
 ... 2800 2810 2820

GLN VAL VAL GLU GLU ALA LYS VAL SER THR ...
 C A G G T G T C G A A G A A G C T A A A G T T T C T A C C ...
 2830 2840 2850...
 ... ASP GLU LYS HIS VAL VAL SER LEU ASP PRO
 ... G A T G A A A A A C A C G T T G T C A G C C T T G A T C C A
 ... 2860 2870 2880

ASN ASP GLN SER LYS GLY LYS VAL VAL ...
 A A T G A T C A A T C A A A G G T A A A G G C G T G T C ...
 2890 2900 2910...
 ... ILE ASN ASN MET ALA ASN GLY GLU ILE SER
 ... A T T A A C A A T A T G G C T A A T G G C G A A T T T C T
 ... 2920 2930 2940

FIG.22O

```

ALA  THR  SER  THR  ASP  ALA  ILE  ASN  GLY  SER  ...
GCC  ACC  TTC  ACC  GAT  GCC  GAT  TAC  GGA  AGT...
2950
...  GIN  LEU  TYR  ALA  VAL  ALA  LYS  GLY  VAL  THR
...  CAG  TTG  TAT  GCC  GTG  GCA  AAG  GGG  TACA  CA
2960
...
2970...
2980
2990
3000

ASN  LEU  ALA  GLY  GLN  VAL  ASN  ASN  LEU  GLU  ...
AAC  CTT  GCT  GGC  ACA  GAG  TGA  ATA  TCT  TGA  G...
3010
3020
...  GLY  LYS  VAL  ASN  LYS  VAL  GLY  LYS  ARG  ALA
...  GGC  AAA  GAG  TGA  ATA  AAG  TGG  GCA  AAC  GTG  CA
3030...
3040
3050
3060

ASP  ALA  GLY  THR  ALA  SER  ALA  LEU  ALA  ALA  ...
GAT  GCA  GGT  ACT  GCA  AGT  GCA  TTAG  CGG  CT...
3070
3080
...  SER  GLN  LEU  PRO  GLN  ALA  THR  MET  PRO  GLY
...  TCA  CAG  TTA  CCA  CACA  GCC  ACT  ATG  CCA  GGT
3090...
3100
3110
3120

LYS  SER  MET  VAL  SER  ILE  ALA  GLY  SER  ...
AAA  TCA  ATG  GTT  TCT  ATTT  GCG  GGA  AGT  AGT...
3130
3140
3150...

```

FIG.22P

```

... TYR  GLN  GLY  GLN  ASN  GLY  LEU  ALA  ILE  GLY
... TAT  CAA  AGG  TCA  AAA  ATG  GTT  TAG  CTT  ATC  GGG
...                                     3160      3170      3180

VAL  SER  ARG  ILE  SER  ASP  ASN  GLY  LYS  VAL  ...
GTA  TCA  AGA  ATA  TTC  CGA  TAA  TGG  CAA  AGT  G...
3190                                     3200      3210...
... ILE  ILE  ARG  LEU  SER  GLY  THR  THR  ASN  SER
... ATA  TAT  CGC  TTG  TCA  GGC  ACA  CCA  ATA  GT
...                                     3220      3230      3240

GLN  GLY  LYS  THR  GLY  VAL  ALA  ALA  GLY  VAL  ...
CAG  GTA  AAA  ACA  GGC  GTT  GCA  GCA  GGT  GTT...
3250                                     3260      3270...
... GLY  TYR  GLN  TRP  ***
... GGT  TAC  CAG  TGG  TAA  TAG  AAT  TCC  GGA  TCC
...                                     3280      3290      3300

```

FIG.23A

NTHi-Stamm K22 Hia

```

      MET ASN LYS ILE PHE ASN...
GCGAATTCAATGACAAATTTTAA...
10      ... VAL ILE TRP ASN VAL VAL THR GLN THR TRP VAL
      ... CGTTATTGGAATGTTGTGACTCAAACTTGGGT 60
      ... 30 40 50

      VAL VAL SER GLU LEU THR ARG ALA HIS...
TGTCGTATCTGAACCTCACTCGCGCCA...
70      ... THR LYS CYS ALA SER ALA THR VAL ALA VAL ALA
      ... CACCAATGCGCCCTCCGCCACCGTGGCGTGGTGC 120
      ... 90 100 110

      VAL LEU ALA THR ALA LEU SER ALA THR...
CGTATTGGCAACTGCGTTGTCTGCAC...
130      ... ALA GLU ALA ASN ASN THR SER VAL THR ASN
      ... GGCTGAAGCGAACACAATACTCTGTACGAA 180
      ... 150 160 170

```

FIG.23B

GLY LEU ASN ALA TYR GLY ASP THR ASN...
 TGGGTTGAAATGCTTATGGCGATACTAA ...
 190
 ... PHE ASN THR THR ASN ASN SER ILE ALA ASP LEU
 200
 ...TTTAAATACAACAATAATTCGATAGCAGATT
 210
 ... 220
 230
 240

GLU LYS HIS VAL GLN ASP ALA TYR LYS...
 GGAAACAACGTTCAAGATGCTTATAA ...
 250
 ... GLY LEU LEU ASN LEU ASN GLU LYS ASP THR ASN
 260
 ...AGGCTTATTAAATCTGAATGAAAGATACAAA
 270
 ... 280
 290
 300

LYS SER SER PHE LEU VAL ALA ASP ASN...
 TAGTCAAGTTTCTTGGTTGCCGACAA ...
 310
 ... THR ALA ALA THR VAL GLY ASN LEU LYS LEU
 320
 ...TACCGCCGCAACCGTAGGCAATTTCGTAATAATT
 330
 ... 340
 350
 360

GLY TRP VAL LEU SER SER LYS ASN GLY...
 GGGCTGGGTATTGTCTAGCAAAACGG ...
 370
 ... 380
 ...

FIG.23C

```

...   THR   ARG   ASN   GLU   LYS   SER   TYR   GLN   VAL   LYS   GLN
...CACAAGGAACGAGAGAAAGCTATCAAGTAAACA   420
...   390   400

ALA   ASP   GLU   VAL   LEU   PHE   THR   GLY   SER...
AGCTGATGAAGTTCTCTTTTACTGGATC...
430   440   ...
...   GLY   ALA   ALA   THR   VAL   SER   SER   SER   LYS   ASP
...TGGTGCTGCACACGGTTAGTTCAGCTCTAAGA   480
...   450   460   470

GLY   LYS   HIS   THR   ILE   THR   ILE   SER   VAL...
CGGTAAACATACCATTACCATTCTGT...
490   500   ...
...   THR   LYS   GLY   SER   PHE   ALA   GLU   VAL   LYS   THR   ASP
...TACCACAAGGTAGTTTGTGCTGAGGTAAACACTGA   540
...   510   520   530

ALA   THR   THR   GLY   GLY   GLN   VAL   ASN   ALA...
TGCAACTACTGGAGGTCAAGTAAACGC...
550   560   ...
...   ASP   ARG   GLY   LYS   VAL   LYS   ALA   GLU   ASP   GLU   ASN
...CGACCGTGGTAAAGTGAAAGCTGAGGACGAGAA   600
...   570   580   590

```


FIG.23D

```

GLY  ALA  ASP  VAL  ASP  VAL  LYS  LYS  VAL  ALA...
TGGAGCTGATGTTGATAGAAAGTTGC ...
610
...  THR  VAL  LYS  ASP  VAL  ALA  LYS  ALA  ILE  ASN  ASP
...AAC TGTAAAGATGTTGCTAAGGCGATTAA C G A 660
... 630 640 650

ALA  ALA  THR  PHE  VAL  LYS  VAL  GLU  SER...
TGCCCGCAACTTTCGTGAAAGTGGAAG ...
670
...  THR  ASP  ASP  ILE  GLU  ASN  GLY  ALA  ALA  GLY
...CACAGATGATGACATTGAAATA TGGTGCTGCAGG 720
... 690 700 710

LYS  ASN  GLU  THR  THR  ASP  GLN  ALA  LEU...
CAAAATGAATACTACAGACCAAGCTCT ...
730
...  LYS  ALA  GLY  ASP  THR  LEU  THR  LYS  ALA  GLY
...CAAGCAGGCGACACCTTAACCTTAAAGCGGG 780
... 750 760 770

LYS  ASN  LEU  LYS  ALA  LYS  LEU  ASP  GLN...
TAAACCTTAAGCTAAGTTAGACCA ...
790
800
...
```

FIG.23E

... ASN GLY LYS SER VAL THR PHE ALA LEU ALA LYS
 ...A A A T G G T A A A T C A G T A C C T T T G C T T T A G C G A A 840
 ... 810 820 830

ASP LEU ASP VAL THR SER ALA LYS VAL...
 A G A C C T T G A T G T G A C C T C T G C G A A A G T ...
 850 860 ...
 ... SER ASP LYS LYS LEU SER ILE GLY LYS ASP THR ASN
 ...G A G T G A T A A G T T G T C T A T T G G T A A A G A T A C G A A 900
 ... 870 880 890 ;

LYS VAL ASP ILE THR SER ASP ALA ASN...
 T A A A G T T G A T A T T A C C A G T G A T G C A A A ...
 910 920 ...
 ... GLY LEU LYS LYS LEU ALA LYS THR GLY ASN GLY ASN
 ...T G G C T T G A A A T T G G C G A A A C A G G T A A C G G A A A 960
 ... 930 940 950

GLY GLN ASN GLY ASN VAL HIS LEU ASN...
 T G G T C A A A A C G G T A A T G T C C A C T T A A A ...
 970 980 ...
 ... GLY ILE ALA SER THR LEU THR ASP THR ILE THR
 ...T G G T A T T G C T T C G A C T T T G A C C G A T A C C A T T A C 1020
 ... 990 1000 1010

FIG.23F

```

GLY  MET  THR  THR  GLN  ALA  SER  ASN  GLY...
AGGTATGACACACAGCAAGCAATGG ...
1030
... VAL  ALA  VAL  GLN  ASN  HIS  ASN  ARG  ALA  ALA  SER
...CGTGGCTGTGCAGAAATCAATAATCGTGCTGCCGAG
1040
... 1050 1060 1070 1080

VAL  ALA  ASP  VAL  LEU  ASN  ALA  GLY  TRP...
TGTGGCTGATGTATTAAATGCAGGCTG ...
1090
... ASN  ILE  GLN  GLY  ASN  GLY  ALA  SER  VAL  ASP  PHE
...GAATATTCAAGGCAACGGAGCGAGCGTTGATT
1100
... 1110 1120 1130 1140

VAL  ASN  ALA  TYR  ASP  THR  VAL  ASP  PHE...
TGTCAAATGCTTACGACACAGTAGATT ...
1150
... VAL  ASN  GLY  THR  ASN  THR  ASN  VAL  ASN  VAL  THR
...TGTCAAATGGTACAAACACCAATGTGACGTTAC
1160
... 1170 1180 1190 1200

THR  ASP  THR  ALA  HIS  LYS  LYS  THR  THR...
GACTGATACGGCTCACAAAGACAC ...
1210
... 1220

```

FIG.23G

```

... VAL ARG VAL ASP VAL THR GLY LEU PRO VAL GLN
...CGTCCGTGTGGATGTAAACAGGCTTGCCGGTTCA
... 1230 1240 1250 1260

TYR VAL THR GLU ASP GLY LYS THR VAL...
ATA TGT TAC GGA AGA CGG CAA AAC CGT ...
1270 1280
... VAL LYS VAL ASP ASN LYS TYR TYR GLU ALA LYS
...TGTGAAAGTGGAACAATAAGTATTACGAAGCTAA
... 1290 1300 1310 1320

GLN ASP GLY SER ALA ASP MET ASP LYS...
GCAAGACGGTTTCGGCGGATATGGATAA...
1330 1340
... LYS VAL GLU ASN GLY GLU LEU ALA LYS THR LYS
...AAAGTCGAATAATGGCGAGCTGGCGAAACCAA
... 1350 1360 1370 1380

VAL LYS LEU VAL SER ALA SER GLY GLN...
AGTGAAATTTGGTGTCTGGCAGCGGTCA...
1390 1400
... ASN PRO VAL LYS ILE SER ASN VAL ALA GLU GLN
...AATCCGGTGAAATAATCAGCAATGTTGCCGGAAGG
... 1410 1420 1430 1440

```

FIG.23H

```

THR  GLU  GLU  ASN  ASP  ALA  VAL  SER  PHE...
C A C G G A A G A A A C G A T G C G G T C A G C T T ...
1450
...  LYS  GLN  LEU  LYS  ALA  LEU  GLN  GLU  LYS  GLN  VAL
...T A A G C A A T T G A A A G C C T T G C A A G A G A A C A G G T
1460
... 1470 1480 1490 1500

```

```

THR  LEU  THR  ALA  SER  ASN  ALA  TYR  ALA...
T A C T T A A C T G C G A G C A A T G C T T A T G C ...
1510
...  ASN  GLY  GLY  ASN  ASP  ALA  ASP  GLY  GLY  LYS  ALA
...C A A T G G T G G T A A C G A T G C C G A C G G C G G C A A G G C
1520
... 1530 1540 1550 1560

```

```

THR  GLN  THR  LEU  ASN  ASN  GLY  LEU  ASN...
A A C T C A A A C T T T A A A C A A T G G T T T G A A ...
1570
...  PHE  LYS  PHE  LYS  SER  THR  ASP  GLY  GLU  LEU  LEU
...T T T A A A T T T A A A T C C A C A G A C G G C G A G T T G T T
1580
... 1590 1600 1610 1620

```

```

ASN  ILE  LYS  VAL  GLU  ASN  ASP  THR  VAL...
G A A C A T C A A A G T A G A A A A T G A C A C A G T ...
1630
... 1640

```

FIG.23I

```

...   THR  PHE  THR  PRO  LYS  LYS  GLY  SER  VAL  GLN  VAL
...TACCCTTTACGCCCGAA A A A A GGTTCGGTACAGGT
... 1650                                     1670 1680

      GLY  GLU  ASP  GLY  LYS  ALA  THR  ILE  GLN...
TGGCGAAGACGGGTAAAGGCTACGATTCA...
1690                                     1700
...   ASN  GLY  THR  LYS  THR  THR  ASP  GLY  LEU  VAL  GLU
...AATGGTACGAA A A A A CCGACGGTTTGGTTGA
... 1710                                     1720 1730 1740

      ALA  SER  GLU  LEU  VAL  GLU  SER  LEU  ASN...
AGCTTC CGA AATTGGTTGAAGCCTGAA...
1750                                     1760
...   LYS  LEU  GLY  TRP  LYS  VAL  GLY  VAL  ASP  LYS  ASP
...CAAACTGGGCTGGAAAGTGCGCGTTGATAAGAA
... 1770                                     1780 1790 1800

      GLY  SER  GLY  GLU  LEU  ASP  GLY  ALA  SER...
CGGCAGCGGCGAGCTTGATGGTGCA TC...
1810                                     1820
...   ASN  GLU  THR  LEU  VAL  LYS  SER  GLY  ASP  LYS  VAL
...CAATGAACCTTTAGTGAGTCCGGCGGATAAGT
... 1830                                     1840 1850 1860

```

FIG.23J

```

THR  LEU  LYS  ALA  GLY  GLU  ASN  LEU  LYS...
A A C T T T G A A A G C C G G C G A G A A T C T G A A ...
1870
... VAL  LYS  GLN  ASP  THR  THR  ASN  PHE  THR  TYR  ALA
...G G T C A A A C A A G A C G G C A C A A C T T C A C T T A C G C
1880
... 1890 1900 1910 1920

LEU  LYS  ASP  GLU  LEU  THR  GLY  VAL  LYS...
G C T C A A A G A T G A A T T G A C G G C G T G A A ...
1930
... SER  VAL  GLU  PHE  LYS  ASP  THR  ALA  ASN  GLY  SER
...G A G C G T G G A G T T T A A G A C A C G G C G A A T G G T T C
1940
... 1950 1960 1970 1980

ASN  GLY  ALA  SER  THR  LYS  ILE  THR  LYS...
A A A C G G T G C A A G C A C G A A G A T T A C C A A ...
1990
... ASP  GLY  LEU  THR  ILE  THR  SER  ALA  ASN  GLY  ALA
...A G A C G G C T T G A C C A T T A C G T C G G C A A A C G G T G C
2000
... 2010 2020 2030 2040

ASN  GLY  ALA  ALA  ALA  THR  ASP  ALA  ASP...
G A A T G G T G C G G C G G C G A C T G A T G C G G A ...
2050
... 2060

```

FIG.23K

```

...   LYS   ILE   LYS   VAL   ALA   SER   ASP   GLY   ILE   SER   ALA
...CAAGATTAAAGTGGCTTCAAGACGGCATCAGTGC      2100
... 2070      2080      2090

GLY   ASN   LYS   ALA   VAL   LYS   ASN   VAL   VAL...
GGGTAAATAAGCGGTTAATAACGTTGT...
2110      2120      ...
...   SER   GLY   LEU   LYS   LYS   PHE   GLY   ASP   ALA   ASN   PHE
...GAGCGGACTGAAGAAATTTGGTGATGCGAATT      2160
... 2130      2140      2150

ASN   PRO   LEU   THR   SER   SER   ALA   ASP   ASN...
CAATCCACTGACCAAGTTCGCGCGACA...
2170      2180      ...
...   LEU   THR   LYS   GLN   TYR   ASP   ASP   ALA   TYR   LYS   GLY
...CTTAACGAACAATAATGACGATGCCATAAAGG      2220
... 2190      2200      2210

LEU   THR   ASN   LEU   ASP   GLU   LYS   GLY   ALA...
CTGACCAATTGGATGAATAAGGTGC...
2230      2240      ...
...   ASP   LYS   GLN   THR   LEU   THR   VAL   ALA   ASP   ASN   THR
...GGACAAGCAAACTCTGACTGTGCCGACATAC      2280
... 2250      2260      2270

```


FIG.23L

```

ALA  ALA  THR  VAL  GLY  ASP  LEU  ARG  GLY...
TGCCGCAACCGTGCGCGGATTTGCGCGG...
2290                                     2300
...  LEU  GLY  TRP  VAL  ILE  SER  ALA  ASP  LYS  THR  THR
...CTTGGGCTGGGTCAATTCTGCGGACAAACCAAC
... 2310                                     2320 2330 2340

GLY  GLU  LEU  ASN  LYS  GLU  TYR  ASN  ALA...
AGCGCAACTCAATAAGGAATAACAAGC...
2350                                     2360
...  GLN  VAL  ARG  ASN  ALA  ASN  GLU  VAL  LYS  PHE  LYS
...GCAAGTGCGTAACGCCCAATGAAGTGAATTCAA
... 2370                                     2380 2390 2400

SER  GLY  ASN  GLY  ILE  HIS  VAL  SER  GLY...
GAGCGCAACGGTATCCATGTTTCCGG...
2410                                     2420
...  LYS  THR  VAL  ASN  GLY  ARG  ARG  GLU  ILE  THR  PHE
...TAAACGGTCAACGGTAGGCGCGAAATTACTTT
... 2430                                     2440 2450 2460

GLU  LEU  ALA  LYS  ASP  GLU  ASN  ALA  ILE...
TGAAATTGGCTAAAGACGGAATAATGCCAT...
2470                                     2480
...
```

FIG.23M

```

...   ALA   PHE   GLY   TYR   GLY   SER   LYS   ALA   LEU   ARG   ASP
...TGC TTT CGG TTA TGG CTTCAA AAG CCTTG CGCGGA      2510 2520
... 2490      2500

ASN THR VAL ALA ILE GLY THR GLY ASN...
TAACACGGTGGCAATTGGTACGGGCA A ...
2530      2540
... VAL VAL ASN ALA ALA GLU LYS SER GLY ALA PHE GLY
...CGTTGTGAATGCGGGA A AATCTGGTGCA TTCGG      2570 2580
... 2550      2560

ASP PRO ASN TYR ILE GLU ASP LYS ALA...
CGATCCGAAC TACATCGAAGATAAGC ...
2590      2600
... GLY GLY SER TYR ALA PHE GLY ASN ASP ASN ARG
...CGGTGGCAGCTACGCTTTCGGTAACGATAACCG      2630 2640
... 2610      2620

ILE THR SER LYS ASN THR PHE VAL LEU...
TATTA CTCTA A A A C A C T T T T G T G T T ...
2650      2660
... GLY ASN GLY VAL ASN ALA LYS TYR LYS ALA ASN
...GGGTAAATGGAGTTAATGCGGAATAATAAGCCAA      2690 2700
... 2670      2680

```

FIG.23N

```

GLY  ASP  VAL  ASP  THR  GLU  THR  VAL  THR...
T G G A G A T G T T G A T A C G G A A A C C G T A A C ...
2710
... VAL  LYS  ASP  LYS  ASP  GLY  LYS  GLU  THR  THR  VAL
...C G T T A A G G A C A A A G A C G G T A A A G A C T A C C G T
2720
... 2730 2740 2750 2760

```

```

THR  VAL  PRO  LYS  ALA  LEU  GLY  ALA  THR...
T A C T G T T C C T A A A G C G T T A G G G G C T A C ...
2770
... VAL  GLU  ASN  SER  VAL  TYR  LEU  GLY  ASN  LYS  SER
...G G T T G A A A A C T C C G T T T A T T G G G T A A T A A T C
2780
... 2790 2800 2810 2820

```

```

THR  ALA  THR  LYS  ASP  LYS  GLY  LYS  ASN...
G A C T G C G A C A A A A G A T A A G G G T A A A A ...
2830
... LEU  LYS  SER  ASP  GLY  THR  ALA  GLY  ASN  THR  THR
...C C T G A A A T C T G A T G G T A C G G C G G T A A C A C T A C
2840
... 2850 2860 2870 2880

```

```

THR  ALA  GLY  THR  THR  GLY  THR  VAL  ASN...
A A C T G C T G G C A C A A C G G G T A C G G T A A A ...
2890
...

```

FIG.230

```

...   GLY  PHE  ALA  GLY  ALA  THR  ALA  HIS  GLY  ALA  VAL
...CGGCCTTTGGCCGGTGCAACGGCGCACGGTGCGGT      2940
... 2910      2920      2930

SER  VAL  GLY  ALA  SER  GLY  GLU  GLU  ARG...
TTC TGT C G G C G C A A G C G G C G A A G A A G ...
2950      2960
...   ARG  ILE  GLN  ASN  VAL  ALA  GLY  GLU  ILE  SER
...ACGTATCCAAACAACGTCGCGCGCAAGCGAATTTCT      3000
... 2970      2980      2990

ALA  THR  SER  THR  ASP  ALA  ILE  ASN  GLY...
CGCCACTTCCACCGATGCGATTAAACGG ...
3010      3020
...   SER  GLN  LEU  TYR  ALA  VAL  ALA  LYS  GLY  VAL  THR
...CAGCCAGTTGTATGCTGTGGCAAAAGGGGTAAAC      3060
... 3030      3040      3050

ASN  LEU  ALA  GLY  GLN  VAL  ASN  LYS  VAL...
AATCTTGTGCGACAAAGTGAAATAAAGT ...
3070      3080
...   GLY  LYS  ARG  ALA  ASP  ALA  GLY  THR  ALA  SER  ALA
...GGGCAACCGTGCAAGATGCAGGTACAGCAAGTGC      3120
... 3090      3100      3110

```

FIG.23P

```

LEU  ALA  ALA  SER  GLN  LEU  PRO  GLN  ALA...
A  T  A  G  C  A  G  C  T  T  C  A  C  A  G  T  T  A  C  C  A  C  A  G  C  ...
3130
...  SER  MET  PRO  GLY  LYS  SER  MET  VAL  SER  ILE  ALA
...C  T  C  T  A  T  G  C  C  A  G  G  T  A  A  A  T  C  A  A  T  G  G  T  T  C  T  A  T  T  G  C
... 3150 3160 3170 3180

GLY  SER  SER  TYR  GLN  GLY  GLN  ASN  GLY...
G  G  G  A  A  G  T  A  G  T  T  A  T  C  A  A  G  G  T  C  A  A  A  T  G  G  ...
3190
...  LEU  ALA  ILE  GLY  VAL  SER  ARG  ILE  SER  ASP  ASN
...T  T  A  G  C  T  A  T  C  G  G  G  G  T  A  T  C  A  C  G  A  A  T  T  C  C  G  A  T  A  A
... 3210 3220 3230 3240

GLY  LYS  VAL  ILE  ILE  ARG  LEU  SER  GLY...
T  G  G  C  A  A  A  G  T  G  A  T  T  A  T  T  C  G  C  T  T  G  T  C  A  G  G  ...
3250
...  THR  THR  ASN  SER  GLN  GLY  LYS  THR  GLY  VAL  ALA
...C  A  C  A  A  C  C  A  A  T  A  G  C  C  A  A  G  G  T  A  A  A  C  A  G  C  G  T  T  G  C
... 3270 3280 3290 3300

ALA  GLY  VAL  GLY  TYR  GLN  TRP  ***
A  G  C  A  G  G  T  G  T  T  G  G  T  T  A  C  C  A  G  T  G  G  T  A  A  T  A  ...
3310
...
...G  A  A  T  T  G  A  T  C  C  G  C
... 3330

```

FIG.24A

H. influenzae Typ c Stamm API *hla*-Sequenz

```

MET  ASN  LYS  ILE  PHE  ASN  VAL  ILE  TRP  ASN  ...
A T G A A C A A A T T T T A A C G T T A T T T G G A A T ...
      10                               20      30...
      ... VAL MET THR GLN THR TRP VAL VAL VAL SER
      ... G T T A T G A C T C A A A C T T G G G T T G T C G T A T C T
      ...                               40      50      60

```

```

GLU  LEU  THR  ARG  THR  HIS  THR  LYS  ARG  ALA  ...
G A A C T C A C T C G C A C C C A C A C A A C G C G C C ...
      70                               80      90...
      ... SER ALA THR VAL GLU THR ALA VAL LEU ALA
      ... T C C G C A A C C G T G G A G A C C G C G T A T T G C C G
      ...                               100      110      120

```

```

THR  LEU  LEU  PHE  ALA  THR  VAL  GLN  ALA  ASN  ...
A C A C T G T T G T T T G C A A C G G T T C A G G C G A A T ...
      130                               140      150...
      ... ALA THR ASP GLU ASP GLU GLU LEU ASP PRO
      ... G C T A C C G A T G A A G A T G A A G A G T A G A C C C C
      ...                               160      170      180

```

```

VAL  VAL  ARG  THR  ALA  PRO  VAL  LEU  SER  PHE  ...
G T A G T A C G C A C T G C T C C C G T G T T G A G C T T C ...
      190                               200      210...

```

FIG.24B

```

... HIS SER ASP LYS GLU GLY THR GLY GLU LYS
... C A T T C C G A T A A G A G G C A C G G G A G A A A A
... 220 230 240

GLU VAL THR GLU ASN SER ASN TRP GLY ILE ...
G A G T T A C A G A A A T T C A A T T G G G A A T A ...
250 260 270...
... TYR PHE HIS ASN LYS GLY VAL LEU LYS ALA
... T A T T C C A C A A T A A G G A G T A C T A A A G C C
... 280 290 300

GLY ALA ILE THR LEU LYS ALA GLY ASP ASN ...
G G A G C A A T C A C C C T C A A A G C C G G C A C A C ...
310 320 330...
... LEU LYS ILE LYS GLN SER THR ASN ALA SER
... C T G A A A A T C A A C A A A G C C A C C A T G C C A G T
... 340 350 360

SER PHE THR TYR SER LEU LYS LYS ASP LEU ...
A G C T T C A C C T A C T C G C T G A A A A G A C C T C ...
370 380 390...
... THR ASP LEU THR SER VAL ALA THR GLU LYS
... A C A G A T C T G A C C A G T G T G C A C T G A A A A
... 400 410 420

```

FIG.24C

```

LEU  SER  PHE  GLY  ALA  ASN  GLY  ASP  LYS  VAL  ...
T T A T C G T T T G G C G C A A A C G G C G A T A A G T T ...
430
... ASP  ILE  THR  SER  ASP  ALA  ASN  GLY  LEU  LYS
... G A T A T T A C C A G T G A T G C A A A T G G C T T G A A A
440
...
460
470
480

LEU  ALA  LYS  THR  GLY  ASN  GLY  ASN  VAL  HIS  ...
T T G G C G A A A A C A G G T A A C G G A A A T G T T C A T ...
490
... LEU  ASN  GLY  LEU  ASP  SER  THR  LEU  PRO  ASP
... T T G A A T G G T T T G G A T T C A A C T T T G C C T G A T
500
...
520
530
540

ALA  VAL  THR  ASN  THR  GLY  VAL  LEU  SER  SER  ...
G C G G T A A C G A A T A C A G G T G T G T T A A G T T C A ...
550
... SER  SER  PHE  THR  PRO  ASN  ASP  VAL  GLU  LYS
... T C A A G T T T T A C A C C T A A T G A T G T T G A A A A
560
...
580
590
600

THR  ARG  ALA  ALA  THR  VAL  LYS  ASP  VAL  LEU  ...
A C A G A G C T G C A A C T G T T A A A G A T G T T T A ...
610
620
630...

```


FIG.24D

```

... ASN ALA GLY TRP ASN ILE LYS GLY ALA LYS
... AATGCCAGGTTGGAAACATTAAAGGTGCTAAA 660
... 640
... 650

THR ALA GLY GLY ASN VAL SER VAL ASP ...
ACTGCTGGAGGTAATGTTGAGAGTGTGAT...
670 680 690...
... LEU VAL SER ALA TYR ASN ASN VAL GLU PHE
... TTAGTGTCCTTATAATAATGTTGAATT 720
... 700 710

ILE THR GLY ASP LYS ASN THR LEU ASP VAL ...
ATTACAGGCGATAAACAACGCTTGATGT...
730 740 750...
... VAL LEU THR ALA LYS GLU ASN GLY LYS THR
... GTATTACAGCTAAAGAAACGGTAACA 780
... 760 770

THR GLU VAL LYS PHE THR PRO LYS THR SER ...
ACCGAAGTGAAATTCACACCGAACAACCTCT...
790 800 810...
... VAL ILE LYS GLU LYS ASP GLY LYS LEU PHE
... GTATCAAGAAAGACGGTAAGTTATT 840
... 820 830
...

```

FIG.24E

```

THR  GLY  LYS  GLU  ASN  ASN  THR  ASN  LYS  ...
A C T G G A A A A G A G A A T A A C G A C A C A A T A A A ...
850
... VAL  THR  SER  ASN  THR  ALA  THR  ASP  ASN  THR
... G T T A C A A G T A A C A C G G C G A C T G A T A A T A C A
880
...
890
900

ASP  GLU  GLY  ASN  GLY  LEU  VAL  THR  ALA  LYS  ...
G A T G A G G G T A A T G G C T T A G T C A C T G C A A A A ...
910
... ALA  VAL  ILE  ASP  ALA  VAL  ASN  LYS  ALA  GLY
... G C T G T G A T T G A T G C T G T G A A C A A G G C T G G T
940
...
950
960

TRP  ARG  VAL  LYS  THR  THR  THR  ALA  ASN  GLY  ...
T G G A G A G T T A A A C A C T A C T G C T A A T G G T ...
970
... GLN  ASN  GLY  ASP  PHE  ALA  THR  VAL  ALA  SER
... C A A A A T G G C G A C T T C G C A A C T G T T G C G T C A
1000
...
1010
1020

GLY  THR  ASN  VAL  THR  PHE  GLU  SER  GLY  ASP  ...
G G C A C A A A T G T A A C C T T T G A A A G T G G C G A T ...
1030
1040
1050...

```

FIG.24F

```

... GLY THR THR ALA SER VAL THR LYS ASP THR
... GGTACAACAGCGTCAGTAACATAAGATACT
... 1060 1070 1080

ASN GLY ASN GLY ILE THR VAL LYS TYR ASP ...
ACGGCAATGGCATCACTGTTAAGTACGAC...
1090 1100 1110...
... ALA LYS VAL GLY ASP GLY LEU LYS PHE ASP
... GCGAAGTTGGCGACGGCTTGAAATTGTGAT
... 1120 1130 1140

SER ASP LYS LYS ILE VAL ALA ASP THR ...
AGCGATAAATAAATCGTTGCAGATACGACC...
1150 1160 1170...
... ALA LEU THR VAL THR GLY LYS VAL ALA
... GCACTTACTGTGACAGGTGGTAAGGTAGCT
... 1180 1190 1200

GLU ILE ALA LYS GLU ASP LYS LYS ...
GAATTGCTAAGAGAGATGACAGAAATAA...
1210 1220 1230...
... LEU VAL ASN ALA GLY ASP LEU VAL THR PHE
... CTTGTTAATGCCAGGCCGATTGTGTAACAGCT
... 1240 1250 1260

```

FIG.24G

```

LEU  GLY  ASN  LEU  SER  TRP  LYS  ALA  LYS  ALA  ...
T T A G G T A A T C T A A G T T G G A A A G C A A A G C T ...
1270                                     1280
...  GLU  ALA  ASP  THR  ASP  THR  ASP  GLY  ALA  LEU
...  G A G G C T G A T A C T G A T A C T G A T G G T G C G C T T
1290...                                     1300
...                                     1310
...                                     1320

GLU  GLY  ILE  SER  LYS  LYS  ASP  GLN  GLU  VAL  LYS  ...
G A G G G A T T T C A A A A G A C C A A G A A G T C A A A ...
1330                                     1340
...  ALA  GLY  GLU  THR  VAL  THR  PHE  LYS  ALA  GLY
...  G C A G G C G A A A C G G T A A C C T T T A A A G C G G C
1350...                                     1360
...                                     1370
...                                     1380

LYS  ASN  LEU  LYS  VAL  LYS  GLN  ASP  GLY  ALA  ...
A A G A A C T T A A A A G T G A A A C A G G A T G G T G C G ...
1390                                     1400
...  ASN  PHE  THR  TYR  SER  LEU  GLN  ASP  ALA  LEU
...  A A C T T A C T T A T T C A C T G C A A G A T G C T T T A
1410...                                     1420
...                                     1430
...                                     1440

THR  GLY  LEU  THR  SER  ILE  THR  LEU  GLY  GLY  ...
A C G G G T T T A A C G A G C A T T A C T T T A G G T G G T ...
1450                                     1460
...                                     1470...

```

FIG.24H

```

... THR THR ASN GLY GLY ASN ASP ALA LYS THR
... A C A C T A A T G G C G G A A A T G A T G C G A A A C C
... 1480 1490 1500

VAL ILE ASN LYS ASP GLY LEU THR ILE THR ...
G T C A T C A A C A A G A C G G T T T A C C A T C A C G...
1510 1520 1530...
... PRO ALA GLY ASN GLY GLY THR THR GLY THR
... C C A G C A G G T A A T G G C G G T A C G A C A G G T A C A
... 1540 1550 1560

ASN THR ILE SER VAL THR LYS ASP GLY ILE ...
A A C A C C A T C A G C G T A C C C A A A G A T G G C A T T...
1570 1580 1590...
... LYS ALA GLY ASN LYS ALA ILE THR ASN VAL
... A A G C A G G T A A T A A G C T A T T A C T A A T G T T
... 1600 1610 1620

ALA SER GLY LEU ARG ALA TYR ASP ASP ALA ...
G C G A G T G G T T T A G A G C T T A T G A C G A T G C G...
1630 1640 1650...
... ASN PHE ASP VAL LEU ASN ASN SER ALA THR
... A A T T T G A T G T T T T A A T A A C T C T G C A A C T
... 1660 1670 1680

```

FIG.24I

```

ASP  LEU  ASN  ARG  HIS  VAL  GLU  ASP  ALA  TYR  ...
G A T T A A A T A G A C A C G T T G A A G A T G C T T A T ...
1690
...  LYS  GLY  LEU  LEU  LEU  ASN  LEU  ASN  GLU  LYS  ASN
...  A A A G G T T T A T T A A A T C T A A A T G A A A A A T
1700
...  .
1720
1730
1740

ALA  ASN  LYS  GLN  PRO  LEU  VAL  THR  ASP  SER  ...
G C A A A T A A A C A A C C G T T G G T G A C T G A C A G C ...
1750
...  THR  ALA  ALA  THR  VAL  GLY  ASP  LEU  ARG  LYS
...  A C G G C G G C G A C T G T A G G C G A T T T A C G T A A A
1760
...
1780
1790
1800

LEU  GLY  TRP  VAL  VAL  SER  THR  LYS  ASN  GLY  ...
T T G G G T T G G T A G T A T C A A C C A A A A C G G T ...
1810
...  THR  LYS  GLU  GLU  SER  ASN  GLN  VAL  LYS  GLN
...  A C G A A A G A A G A A G C A A T C A A G T T A A C A A
1820
...
1840
1850
1860

ALA  ASP  GLU  VAL  LEU  PHE  THR  GLY  ALA  GLY  ...
G C T G A T G A A G T C C T C T T T A C C G G A G C C G T ...
1870
1880
1890

```

FIG.24J

```

... ALA ALA THR VAL THR SER LYS SER GLU ASN
... GCTGCTACGGTTACTTCCTCTGAAAC
...
1900
1910
1920

GLY LYS HIS THR ILE THR VAL SER VAL ALA ...
GGTAACATACTATTAACCGTTAGTGCT...
1930
1940
1950...
... GLU THR LYS ALA ASP SER GLY LEU GLU LYS
... GAACATAAGCGGATAGCGGCTCTGAAAA
1960
1970
1980

ASP GLY ASP THR ILE LYS LEU LYS VAL ASP ...
GATGGGATACTATTAAGCTCAAGTGGA T...
1990
2000
2010...
... ASN GLN ASN THR ASP ASN VAL LEU THR VAL
... ATCAAAACACTGATATA TGT TTA CTGT
2020
2030
2040

GLY ASN ASN GLY THR ALA VAL THR LYS GLY ...
GGTATAATGGTACTGCTGTCACTAAAGGT...
2050
2060
2070...
... GLY PHE GLU THR VAL LYS THR GLY ALA THR
... GGCTTTGAACCTGTTAACCTGGAGCGACT
2080
2090
2100

```

FIG.24K

```

ASP  ALA  ASP  ARG  GLY  LYS  VAL  THR  VAL  LYS  ...
GATGCA GATCGCGGTAAAGTAAC TGTA A A A...
2110
...  ASP  ALA  THR  ALA  ASN  ASP  ALA  ASP  LYS  LYS
...  GATGCTACTGCTAATGACGCTGATAGAA A
2140
...  2150
2160

VAL  ALA  THR  VAL  LYS  ASP  VAL  ALA  THR  ALA  ...
GTCGCA ACTGTAAAGATGTGCAACCGCA...
2170
...  ILE  ASN  SER  ALA  ALA  THR  PHE  VAL  LYS  THR
...  ATTAATAGTGGCGCACTTTGTGA A A C A
2200
...  2210
2220

GLU  ASN  LEU  THR  THR  SER  ILE  ASP  GLU  ASP  ...
GAGAA TTAACTACCTCTATTGTAGAT...
2230
...  ASN  PRO  THR  ASP  ASN  GLY  LYS  ASP  ASP  ALA
...  ATCC TACAGATACGGCA AAGATGACGCA
2260
...  2270
2280

LEU  LYS  ALA  GLY  ASP  THR  LEU  THR  PHE  LYS  ...
CTTAAAGCGGCGGATACCTTAACCTTTAA A A...
2290
...  2300
2310...
```


FIG.24L

```

... ALA GLY LYS ASN LEU LYS VAL LYS ARG ASP
... GCAGGTAA A A A C C T G A A A G T T A A A C G T G A T
... 2320 2330 2340

GLY LYS ASN ILE THR PHE ASP LEU ALA LYS ...
G G A A A A T A T T A C T T T G A C T T G G C G A A A ...
2350 2360 2370...
... ASN LEU GLU VAL LYS THR ALA LYS VAL SER
... A A C C T T G A G G T G A A A C T G C G A A A G T G A G T
... 2380 2390 2400

ASP THR LEU THR ILE GLY GLY ASN THR PRO ...
G A T A C T T A A C G A T T G G C G G G A A T A C A C C T ...
2410 2420 2430...
... THR GLY GLY THR THR ALA THR PRO LYS VAL
... A C A G G T G G C A C T A C T G C G A C G C C A A A G T G
... 2440 2450 2460

ASN ILE THR SER THR ALA ASP GLY LEU ASN ...
A A T A T T A C T A G C A C G G C T G A T G G T T T G A A T ...
2470 2480 2490...
... PHE ALA LYS GLU THR ALA ASP ALA SER GLY
... T T T G C A A A A G A A A C A G C C G A T G C C T C G G G T
... 2500 2510 2520

```

FIG.24M

```

SER  LYS  ASN  VAL  TYR  LEU  LYS  GLY  ILE  ALA  ...
T C T A A G A A T G T T T A T T G A A A G G T A T T G C G ...
2530
...  THR  THR  LEU  THR  GLU  PRO  SER  ALA  GLY  ALA
...  A C A A C T T T A A C T G A G C C A A G C G C G G A G C G
2540
...  2560
...  2570
...  2580

LYS  SER  SER  HIS  VAL  ASP  LEU  ASN  VAL  ASP  ...
A G T C T T C A C A C G T T G A T T T A A A T G T G G A T ...
2590
...  ALA  THR  LYS  LYS  SER  ASN  ALA  ALA  SER  ILE
...  G C G A C G A A A A A T C C A A T G C A G C A G T A T T
2600
...  2620
...  2630
...  2640

GLU  ASP  VAL  LEU  ARG  ALA  GLY  TRP  ASN  ILE  ...
G A G A T G T A T T G C C G C G C A G G T T G G A A T A T T ...
2650
...  GLN  GLY  ASN  GLY  ASN  ASN  VAL  ASP  TYR  VAL
...  C A A G G T A A T G G T A A T A A T G T T G A T T A T G T A
2660
...  2680
...  2690
...  2700

ALA  THR  TYR  ASP  THR  VAL  ASN  PHE  THR  ASP  ...
G C G A C G T A T G A C A C A G T A A A C T T T A C C G A T ...
2710
...  2720
...  2730

```

FIG.24N

```

... ASP SER THR GLY THR THR VAL THR VAL
... G A C A G C A C A G G T A C A A C G G T A A C C G T A 2760
... 2740 2750

THR GLN LYS ALA ASP GLY LYS GLY ALA ASP ...
A C C C A A A A G C A G A T G G C A A A G G T G C T G A C ...
2770 2780 2790...
... VAL LYS ILE GLY ALA LYS THR SER VAL ILE
... G T T A A A A T C G G T G C G A A A A C T T C T G T T A T C 2820
... 2800 2810

LYS ASP HIS ASN GLY LYS LEU PHE THR GLY ...
A A G A C C A C A C G G C A A A C T G T T T A C A G G C ...
2830 2840 2850...
... LYS ASP LEU LYS ASP ALA ASN ASN GLY ALA
... A A G A C C T G A A A G A T G C G A A T A A T G G T G C A 2880
... 2860 2870

THR VAL SER GLU ASP ASP GLY LYS ASP THR ...
A C C G T T A G T G A A G A T G A T G G C A A A G A C A C C ...
2890 2900 2910...
... GLY THR GLY LEU VAL THR ALA LYS THR VAL
... G G C A C A G G C T T A G T T A C T G C A A A A C T G T G 2940
... 2920 2930

```

FIG.240

```

ILE  ASP  ALA  VAL  ASN  LYS  SER  GLY  TRP  ARG  ...
A  T  G  A  T  G  C  A  G  T  A  A  T  A  A  A  A  G  C  G  G  T  T  G  G  A  G  G  ...
2950
...  VAL  THR  GLY  GLU  GLY  ALA  THR  ALA  GLU  THR
...  G  T  A  A  C  C  G  G  T  G  A  G  G  G  C  C  G  A  C  T  G  C  C  G  A  A  A  C  C
2980
...
GLY  ALA  THR  ALA  VAL  ASN  ALA  GLY  ASN  ALA  ...
G  G  T  G  C  A  A  C  C  G  C  C  G  T  G  A  A  T  G  C  G  G  G  T  A  A  C  G  C  T  ...
3010
...  GLU  THR  VAL  THR  SER  GLY  THR  SER  VAL  ASN
...  G  A  A  A  C  C  G  T  T  A  C  A  T  C  A  G  G  C  A  C  G  A  G  C  G  T  G  A  A  C
3040
...
PHE  LYS  ASN  GLY  ASN  ALA  THR  THR  ALA  THR  ...
T  T  C  A  A  A  A  C  G  G  C  A  A  T  G  C  G  A  C  C  A  C  A  G  C  G  A  C  C  ...
3070
...  VAL  SER  LYS  ASP  ASN  GLY  ASN  ILE  ASN  VAL
...  G  T  A  A  G  C  A  A  A  G  A  T  A  A  T  G  G  C  A  A  C  A  T  C  A  A  T  G  T  C
3100
...
LYS  TYR  ASP  VAL  ASN  VAL  GLY  ASP  GLY  LEU  ...
A  A  A  T  A  C  G  A  T  G  T  A  A  A  T  G  T  T  G  G  T  G  A  C  G  G  C  T  G  ...
3130
...

```

FIG.24P

```

...   LYS   ILE   GLY   ASP   ASP   LYS   LYS   ILE   VAL   ALA
...   A A G A T T G G C G A T G A C A A A A A A A T C G T T G C A
...                                     3160      3170      3180
...                                     ...

ASP   THR   THR   THR   LEU   THR   VAL   THR   GLY   GLY   ...
G A C A C G A C C A C T T A C T G T A A C A G G T G G T ...
3190                                     3200      3210...
...   LYS   VAL   SER   VAL   PRO   ALA   GLY   ALA   ASN   SER
...   A A G G T G T C T G T T C C T G C T G G T G C T A A T A G T
...                                     3220      3230      3240
...                                     ...

VAL   ASN   ASN   ASN   LYS   LYS   LEU   VAL   ASN   ALA   ...
G T T A T A C A A T A A G A A A C T T G T T A A T G C A ...
3250                                     3260      3270...
...   GLU   GLY   LEU   ALA   THR   ALA   LEU   ASN   ASN   LEU
...   G A G G G T T T A G C G A C T G C T T T A A C A C C T A
...                                     3280      3290      3300
...                                     ...

SER   TRP   THR   ALA   LYS   ALA   ASP   LYS   TYR   ALA   ...
A G C T G G A C G G C A A A A G C C G A T A A T A T G C A ...
3310                                     3320      3330...
...   ASP   GLY   GLU   SER   GLU   GLY   GLY   GLU   THR   ASP   GLN
...   G A T G G C G A G T C A G A G G G C G A A A C C G A C C A A
...                                     3340      3350      3360
...                                     ...

```

FIG.24Q

```

GLU VAL LYS ALA GLY ASP LYS VAL THR PHE ...
G A G T C A A A G C A G G C G A C A A A G T A A C C T T T ...
3370
... LYS ALA GLY LYS ASN LEU LYS VAL LYS GLN
... A A A G C A G G C A A G A A C T T A A A A G T G A A C A G
3380
... 3400
... 3410
... 3420

SER GLU LYS LYS ASP PHE THR TYR SER LEU GLN ...
T C T G A A A A G A C T T T A C T T A T T C A C T G C A A ...
3430
... ASP THR LEU THR GLY LEU THR SER ILE THR
... G A C A C T T T A A C A G G C T T A A C G A G C A T T A C T
3440
... 3450
... 3460
... 3470
... 3480

LEU GLY GLY THR ALA ASN GLY ARG ASN ASP ...
T T A G G T G G T A C A G C T A A T G G C A G A A A T G A T ...
3490
... THR GLY THR VAL ILE ASN LYS ASP GLY LEU
... A C G G G A A C C G T C A T C A A C A A A G A C G G C T T A
3500
... 3510
... 3520
... 3530
... 3540

THR ILE THR LEU ALA ASN GLY ALA ALA ALA ...
A C C A T C A C G C T G G C A A A T G G T G C T G C G G C A ...
3550
... 3560
... 3570

```

FIG.24R

```

... GLY THR ASP ALA SER ASN GLY ASN THR ILE
... GGCACAGATGCCGTCTAACGGAAACACCATC
... 3580 3590 3600

SER VAL THR LYS ASP GLY ILE SER ALA GLY ...
AGTGTAACCAAGACGGCATTAGTGGGGT...
3610 3620 3630...
... ASN LYS GLU ILE THR ASN VAL LYS SER ALA
... AATAAGAAATTACCATAAGAGTGCT
... 3640 3650 3660

LEU LYS THR TYR LYS ASP THR GLN ASN THR ...
TTAAACCTATTAAAGATACCTCAAAACACT...
3670 3680 3690...
... ALA GLY ALA THR GLN PRO ALA ALA ASN THR
... GCAGGTGCAACTCAACCTGGCGCTAATACA
... 3700 3710 3720

ALA GLU VAL ALA LYS GLN ASP LEU VAL ASP ...
GCTGAAGTAGCCCAAACAAGACTTGGTTGAT...
3730 3740 3750...
... LEU THR LYS PRO ALA THR GLY ALA ALA GLY
... TTAACATAAACCTGGCGACAGGTGCGCTGGA
... 3760 3770 3780

```

FIG.24S

```

ASN  GLY  ALA  ASP  ALA  LYS  ALA  PRO  ASP  THR  ...
A A T G G T G C A G A T G C A A A A G C T C C C G A T A C C ...
3790                                     3800
...  THR  ALA  ALA  THR  VAL  GLY  ASP  LEU  ARG  GLY
...  A C A G C T G C A A C C G T A G G C G A C T T G C G T G G T
3810...                                     3820
...                                     3830
...                                     3840

LEU  GLY  TRP  VAL  LEU  SER  ALA  LYS  LYS  THR  ...
T T G G C C T G G G T G C T T T C A G C T A A G A A A A C T ...
3850                                     3860
...  ALA  ASP  GLU  THR  GLN  ASP  LYS  GLU  PHE  HIS
...  G C A G A T G A A A C A C A A G A T A A A G A G T T C C A C
3870...                                     3880
...                                     3890
...                                     3900

ALA  ALA  VAL  LYS  ASN  ALA  ASN  GLU  VAL  GLU  ...
G C C G C C G T T A A A A A C G C A A A T G A A G T T G A G ...
3910                                     3920
...  PHE  VAL  GLY  LYS  ASN  GLY  ASP  GLY  ALA  THR  VAL  SER
...  T T C G T G G G T A A A A A C G G T G C A A C C G T G T C T
3930...                                     3940
...                                     3950
...                                     3960

ALA  LYS  THR  ASP  ASN  ASN  GLY  LYS  HIS  THR  ...
G C A A A A C T G A T A A C A A C G G A A A A C A T A C T ...
3970                                     3980
...                                     3990

```


FIG.24T

```

... VAL THR ILE ASP VAL ALA GLU ALA LYS VAL
... GTACGATTGATTGCGAGAGCCAAAGTT
... 4000 4010 4020

GLY ASP GLY LEU GLU LYS ASP THR ASP GLY ...
GGTGA TGGTCTTGAAAGATACTGACGGC...
4030 4040 4050...
... LYS ILE LYS LEU LYS VAL ASP ASN THR ASP
... AAGATTAAACCTCAAGTAGATATACAGAT
... 4060 4070 4080

GLY ASN ASN LEU LEU THR VAL ASP ALA THR ...
GGGAATAATCTATTATACCGTTTGATGCACA...
4090 4100 4110...
... LYS GLY ALA SER VAL ALA LYS GLY GLU PHE
... AAGGTGCATCCGTTGCCAAGGGCGAGTTT
... 4120 4130 4140

ASN ALA VAL THR THR ASP ALA THR THR ALA ...
ATGCCGTAAACAACAGATGCACACAGCC...
4150 4160 4170...
... GLN GLY THR ASN ALA ASN GLU ARG GLY LYS
... CAGGCACAAATGCCCAATGAGCGCGGTAA
... 4180 4190 4200

```

FIG.24U

VAL VAL VAL LYS GLY SER ASN GLY ALA THR ...
 GTGGTTGTCAAGGGTTCAATAATGGTGCAACT...
 4210 4220 4230...
 ... ALA THR GLU THR ASP LYS LYS VAL ALA
 ... GCTACCGAACA CTGACCAAGAAAGTGCA
 4240 4250 4260
 ...

THR VAL GLY ASP VAL ALA LYS ALA ILE ASN ...
 ACTGTTGGCGACGTTGCTTAAAGCGATTAC...
 4270 4280 4290...
 ... ASP ALA ALA THR PHE VAL LYS VAL GLU ASN
 ... GACGCAGCAACTTTCGTGAAAGTGGAATA
 4300 4310 4320
 ...

ASP ASP SER ALA THR ILE ASP ASP SER PRO ...
 GACGACAGTGCTACGATTGTGATAGCCCA...
 4330 4340 4350...
 ... THR ASP ASP GLY ALA ASN ASP ALA LEU LYS
 ... ACAGATGATGGCGCAATAATGATGCTCTCAA
 4360 4370 4380
 ...

ALA GLY ASP THR LEU THR LEU LYS ALA GLY ...
 GCAGCGACACCTTGACCTTAAGCGGGT...
 4390 4400 4410...

FIG.24V

```

... LYS ASN LEU LYS VAL LYS ARG ASP GLY LYS
... A A A A C T T A A A G T T A A A C G T G A T G G T A A A
...
4420
4430
4440

ASN ILE THR PHE ALA LEU ALA ASN ASP LEU ...
A A T A T T A C T T T G C C C T T G C G A A C G A C C T T ...
4450
4460
... SER VAL LYS SER ALA THR VAL SER ASP LYS
... A G T G T A A A A G C G C A A C C G T T A G C G A T A A A
...
4480
4490
4500

LEU SER LEU GLY THR ASN GLY ASN LYS VAL ...
T T A T C G C T T G G T A C A A C G G C A A T A A A G T C ...
4510
4520
... ASN ILE THR SER ASP THR LYS GLY LEU ASN
... A A T A T C A C A A G C G A C A C C A A A G G C T T G A A C
...
4540
4550
4560

PHE ALA LYS ASP SER LYS THR GLY ASP ASP ...
T T C G C T A A A G A T A G T A A G A C A G G C G A T G A T ...
4570
4580
... ALA ASN ILE HIS LEU ASN GLY ILE ALA SER
... G C T A A T A T T C A C T T A A A T G G C A T T G C T T C A
...
4600
4610
4620

```

FIG.24W

THR LEU THR ASP THR LEU LEU ASN SER GLY ...
 A C T T A A C T G A T A C A T T G T T A A A T A G T G G T ...
 4630 4640 4650...
 ... ALA THR THR ASN LEU GLY ASN GLY ILE
 ... G C G A C A C C A A T T A G G T G G T A A T G G T A T T
 ... 4660 4670 4680

THR ASP ASN GLU LYS LYS ARG ALA ALA SER ...
 A C T G A T A A C G A G A A A A C G C G C G C G A G C ...
 4690 4700 4710...
 ... VAL LYS ASP VAL LEU ASN ALA GLY TRP ASN
 ... G T T A A A G A T G T C T T G A A T G C G G G T T G G A A T
 ... 4720 4730 4740

VAL ARG GLY VAL LYS PRO ALA SER ALA ASN ...
 G T T C G T G G T G T T A A C C G G C A T C T G C A A A T ...
 4750 4760 4770...
 ... ASN GLN VAL GLU ASN ILE ASP PHE VAL ALA
 ... A A T C A A G T G G A G A A T A T C G A C T T T G T A G C A
 ... 4780 4790 4800

THR TYR ASP THR VAL ASP PHE VAL SER GLY ...
 A C C T A C G A C A C A G T G G A C T T T G T T A G T G G A ...
 4810 4820 4830...

FIG.24X

```

... ASP LYS ASP THR THR SER VAL THR VAL GLU
... G A T A A G A C A C C A C G A G T G T A A C T G T T G A A
...
4840
4850
4860

SER LYS ASP ASN GLY LYS ARG THR GLU VAL ...
A G T A A G A T A A T G G C A A G A G A A G T T ...
4870
4880
... LYS ILE GLY ALA LYS THR SER VAL ILE LYS
... A A A T C G G T G C G A A G A C T T C T G T T A T C A A A
4900
4910
4920

ASP HIS ASN GLY LYS LEU PHE THR GLY LYS ...
G A C C A C A C G G C A A A C T G T T T A C A G G C A A A ...
4930
4940
4950...
... GLU LEU LYS ASP ALA ASN ASN GLY VAL
... G A G C T G A A G G A T G C T A A C A A T A A T G G C G T A
4960
4970
4980

THR VAL THR GLU THR ASP GLY LYS ASP GLU ...
A C T G T T A C C G A A A C C G A C G G C A A G A C G A G ...
4990
5000
5010...
... GLY ASN GLY LEU VAL THR ALA LYS ALA VAL
... G G T A A T G G T T T A G T G A C T G C A A A A G C T G T G
5020
5030
5040

```

```

ILE ASP ALA VAL ASN LYS ALA GLY TRP ARG ...
ATTGATGCCGTGAATAAGGC TTGGAGA...
5050                               5060       5070...
... VAL LYS THR THR GLY THR ALA ASN GLN ASN ASN
... GTTAAACAACAGGTGCTAATGGTCAGAAAT
...                               5080       5090       5100

ASP ASP PHE ALA THR VAL ALA SER GLY THR ...
GATGACTTCGCAACTGTTCAGGCACA...
5110                               5120       5130...
... ASN VAL THR PHE ALA ASP GLY ASN GLY THR
... AATGTAAACCCTTTGCTGATGGTAATGGCACA
...                               5140       5150       5160

THR ALA GLU VAL THR LYS ALA ASN ASP GLY ...
ACTGCCGAAGTAAC TAAGCAAACGCGGT...
5170                               5180       5190...
... SER ILE THR VAL LYS TYR ASN VAL LYS VAL
... AGTATTACTGT TAAATACAATGTTAAAGTG
...                               5200       5210       5220

ALA ASP GLY LEU LYS LEU ASP GLY ASP LYS ...
GCTGATGGCTTAACAAC TAGACGGCATAA...
5230                               5240       5250...

```

FIG.24Z

```

... ILE VAL ALA ASP THR THR VAL LEU THR VAL
... ATCGTTGCCAGACACGACCGTACTTACTGTG
... 5260 5270 5280

ALA ASP GLY LYS VAL THR ALA PRO ASN ASN ...
GCA GATGGTAAGAATTACAGCTCCGAATAAT...
5290 5300 5310...
... GLY ASP GLY LYS LYS PHE VAL ASP ALA SER
... GGCGATGGTAAGAAATTGTTGATGCCAGT
... 5320 5330 5340

GLY LEU ALA ASP ALA LEU ASN LYS LEU SER ...
GGTTTAGCGGATGCGTTATAATAATTAGC...
5350 5360 5370...
... TRP THR ALA THR ALA GLY LYS GLU GLY THR
... TGGACGGCAACTGCTGGTAAAGAGGCACT
... 5380 5390 5400

GLY GLU VAL ASP PRO ALA ASN SER ALA GLY ...
GGTAGAGTTGATCCCTGCCAATAATCAGCAGG...
5410 5420 5430...
... GLN GLU VAL LYS ALA GLY ASP LYS VAL THR
... CAGAAGTCAAGCGGGCGGACAAAGTACC
... 5440 5450 5460

```

FIG.24A'

```

PHE  LYS  ALA  GLY  ASP  ASN  LEU  LYS  ILE  LYS  ...
T T T A A G C C G G C G A C A A C C T G A A A A T C A A A ...
5470                               5480                               5490...
...  GLN  SER  GLY  LYS  ASP  PHE  THR  TYR  SER  LEU
...  C A A G C G G C A A A G A C T T T A C C T A C T C G C T G
...                               5500                               5510                               5520
...

LYS  LYS  GLU  LEU  LYS  ASP  LEU  THR  SER  VAL  ...
A A A A G A G C T G A A A G A C C T G A C C A G C G T A ...
5530                               5540                               5550...
...  GLU  PHE  LYS  ASP  ALA  ASN  GLY  GLY  THR  GLY
...  G A G T T C A A A G A C G C A A A C G G C G G T A C A G G C
...                               5560                               5570                               5580
...

SER  GLU  SER  THR  LYS  ILE  THR  LYS  ASP  GLY  ...
A G T G A A A G C A C C A A G A T T A C C A A A G A C G G C ...
5590                               5600                               5610...
...  LEU  THR  ILE  THR  PRO  ALA  ASN  GLY  ALA  GLY
...  T T G A C C A T T A C G C C G G C A A A C G G T G C G G T
...                               5620                               5630                               5640
...

ALA  ALA  GLY  ALA  ASN  THR  ALA  ASN  THR  ILE  ...
G C G G C A G G T G C A A A C A C T G C A A A C A C C A T ...
5650                               5660                               5670...

```


FIG.24B'

```

... SER VAL THR LYS ASP GLY ILE SER ALA GLY
... AGCGTAACCAAGATGGCATTAGCGCGGT
... 5680 5690 5700

ASN LYS ALA VAL THR ASN VAL SER GLY ...
AATAAGCAGTTACAAACGTTGTGAGCGGA...
5710 5720 5730...
... LEU LYS LYS PHE GLY ASP GLY HIS THR LEU
... CTGAAGAAATTGTGGTGATGGTCATACGTTG
... 5740 5750 5760

ALA ASN GLY THR VAL ALA ASP PHE GLU LYS ...
GC A A A T G G C A C T G T T G C T G A T T T G A A A G ...
5770 5780 5790...
... HIS TYR ASP ASN ALA TYR LYS ASP LEU THR
... C A T T A T G A C A A T G C C T A T A A G A C T T G A C C
... 5800 5810 5820

ASN LEU ASP GLU LYS GLY ALA ASP ASN ASN ...
A A T T G G A T G A A A A G G C C G G A T A A T A T ...
5830 5840 5850...
... PRO THR VAL ALA ASP ASN THR ALA ALA THR
... C C G A C T G T T G C C G A C A A T A C C G C T G C A C C
... 5860 5870 5880

```

FIG.24C'

```

VAL  GLY  ASP  LEU  ARG  GLY  LEU  GLY  TRP  VAL  ...
G T G G C G A T T T G C G C G G C T T G G C T G G G T C ...
5890                                     5900
... ILE SER ALA ASP LYS THR THR GLY GLU PRO
... A T T C T G C G A C A A A C C A C A G G C G A C C C
5910...                                     5920
...                                     5930
...                                     5940

ASN  GLN  GLU  TYR  ASN  ALA  GLN  VAL  ARG  ASN  ...
A A T C A G G A A T A C A A C G C G C A A G T G C G T A A C ...
5950                                     5960
... ALA ASN GLU VAL LYS PHE LYS SER GLY ASN
... G C C A A T G A A G T G A A A T T C A A G A G C G G C A A C
5970...                                     5980
...                                     5990
...                                     6000

GLY  ILE  ASN  VAL  SER  GLY  LYS  THR  LEU  ASN  ...
G G T A T C A A T G T T T C C G G T A A A C A T T G A A C ...
6010                                     6020
... GLY THR ARG VAL ILE THR PHE GLU LEU ALA
... G G T A C G C G C G T G A T T A C C T T T G A A T T G G C T
6030...                                     6040
...                                     6050
...                                     6060

LYS  GLY  GLU  VAL  VAL  LYS  SER  ASN  GLU  PHE  ...
A A A G G C G A A G T G G T T A A A T C G A A T G A A T T T ...
6070                                     6080
...                                     6090
...                                     6100

```

FIG.24D'

```

... THR VAL LYS ASN ALA ASP GLY SER GLU THR
... ACCGTTAAGAAATGCCGATGGTTTCGGAAACG
...
6100 6110 6120

ASN LEU VAL LYS VAL GLY ASP MET TYR ...
A A C T T G G T T A A G T T G G C G A T A T G T A T T A C ...
6130 6140 6150...
... SER LYS GLU ASP ILE ASP PRO ALA THR SER
... A G C A A A G A G G A T A T T G A C C C G G C A A C C A G T
...
6160 6170 6180

LYS PRO MET THR GLY LYS THR GLU LYS TYR ...
A A C C G A T G A C A G G T A A A A C T G A A A A T A T ...
6190 6200 6210...
... LYS VAL GLU ASN GLY LYS VAL VAL SER ALA
... A A G G T T G A A A A C G G C A A A G T C G T T C T G C T
...
6220 6230 6240

ASN GLY SER LYS THR GLU THR LEU THR ...
A A C G G C A G C A A G A C C G A A G T T A C C C T A A C C ...
6250 6260 6270...
... ASN LYS GLY SER GLY TYR VAL THR GLY ASN
... A A C A A A G G T T C C G G C T A T G T A A C A G G T A A C
...
6280 6290 6300

```

FIG.24E'

```

G L N   V A L   A L A   A S P   A L A   I L E   A L A   L Y S   S E R   G L Y   ...
C A A G T G G C T G A T G C G A T T G C G A A A T C A G G C ...
6310
... P H E   G L U   L E U   G L Y   L E U   A L A   A S P   A L A   A L A   G L U
... T T T G A G C T T G G T T T G G C T G A T G C G C A G A A
6320
...
6330
...
6340
6350
6360

A L A   G L U   L Y S   A L A   P H E   A L A   G L U   S E R   A L A   L Y S   ...
G C T G A A A A G C C T T T G C A G A A A G C G C A A A A ...
6370
... A S P   L Y S   G L N   L E U   S E R   L Y S   A S P   L Y S   A L A   G L U
... G A C A A G C A A T T G T C T A A A G A T A A A G C G G A A
6380
...
6390
6400
6410
6420

T H R   V A L   A S N   A L A   H I S   A S P   L Y S   V A L   A R G   P H E   ...
A C T G T A A A T G C C C A C G A T A A A G T C C G T T T ...
6430
... A L A   A S N   G L Y   L E U   A S N   T H R   L Y S   V A L   S E R   A L A
... G C T A A T G G T T T A A A T A C C A A A G T G A G C G C G
6440
...
6450
6460
6470
6480

A L A   T H R   V A L   G L U   S E R   T H R   A S P   A L A   A S N   G L Y   ...
G C A A C G G T G G A A A G C A C T G A T G C A A A C G G C ...
6490
6500
6510

```

FIG.24F'

```

...   ASP   LYS   VAL   THR   THR   THR   PHE   VAL   LYS   THR
...   G A T A A G T G A C C A C C T T T G T G A A A C C
...                                     6520      6530      6540

ASP   VAL   GLU   LEU   PRO   LEU   THR   GLN   ILE   TYR ...
G A T G T G G A A T T G C C T T A A C G C A A A T C T A C ...
6550                                     6560      6570...

...   ASN   THR   ASP   ALA   ASN   GLY   ASN   LYS   ILE   VAL
...   A A T A C C G A T G C A A A C G G T A A T A A G A T C G T T
...                                     6580      6590      6600

LYS   LYS   ALA   ASP   GLY   LYS   TRP   TYR   GLU   LEU ...
A A A A A G C T G A C G G A A A T G G T A T G A A C T G ...
6610                                     6620      6630...

...   ASN   ALA   ASP   GLY   THR   ALA   SER   ASN   LYS   GLU
...   A A T G C T G A T G G T A C G G C G A G T A A C A A A G A A
...                                     6640      6650      6660

VAL   THR   LEU   GLY   ASN   VAL   ASP   ALA   ASN   GLY ...
G T G A C A C T T G G T A A C G T G G A T G C A A A C G G T ...
6670                                     6680      6690...

...   LYS   LYS   VAL   VAL   LYS   VAL   THR   GLU   ASN   GLY
...   A A G A A A G T T G T G A A A G T A A C C G A A A T G G T
...                                     6700      6710      6720

```

FIG.24G'

```

ALA  ASP  LYS  TRP  TYR  THR  ASN  ALA  ASP  ...
GCGGATAGTGGTTATTAACCAATGCTGAC...
6730
...  GLY  ALA  ALA  ASP  LYS  THR  LYS  GLY  GLU  VAL
...  GGTGCTGCGGATATAACCAAGGCGAAGTG
6760
...
SER  ASN  ASP  LYS  VAL  SER  THR  ASP  GLU  LYS  ...
AGCAATGATAAAGTTTCTACCGATGAATAA...
6790
...  HIS  VAL  VAL  ARG  LEU  ASP  PRO  ASN  ASN  GLN
...  CACGTTGTCCGCCCTTGATCCGAACAATCAA
6820
...
SER  ASN  GLY  LYS  GLY  VAL  VAL  ILE  ASP  ASN  ...
TCGAACGGCAAGAGCGGTGGTCAATTGACAT...
6850
...  VAL  ALA  ASN  GLY  GLU  ILE  SER  ALA  THR  SER
...  GTGGCTAATGGCGGAATTTCTGCCCACTTCC
6880
...
THR  ASP  ALA  ILE  ASN  GLY  SER  GLN  LEU  TYR  ...
ACCGATGCCGATTATACGGGAAGTCAGTTGTAT...
6910
...

```

FIG.24H'

```

... ALA VAL ALA LYS GLY VAL THR ASN LEU ALA
... GCCGTGGCAAAAGGGGTAAACAACCTTGCT
...
6940
...
GLY GLN VAL ASN ASN LEU GLU GLY LYS VAL ...
GGACAAGTGAATAATCTTGAGGGCAAGTG...
6970
6980
... ASN LYS VAL GLY LYS ARG ALA ASP ALA GLY
... AATAAGTGGCAAAACGTGCAGATGCAGGT
7000
7010
...
THR ALA SER ALA LEU ALA SER GLN LEU ...
ACAGCAAGTGCAATTAGCGGCTTCACAGTTA...
7030
7040
... PRO GLN ALA THR MET PRO GLY LYS SER MET
... CCACAAGCCACTATGCCACAGGTAAATCAATG
7060
7070
...
VAL ALA ILE ALA GLY SER SER TYR GLN GLY ...
GTTGCTATTGCGGGAAGTAGTTATCAAGGT...
7090
7100
... GLN ASN GLY LEU ALA ILE GLY VAL SER ARG
... CAAATGGTTTAGCTATCGGGGTATCAGAA
7120
7130
...
7140

```

FIG.24I'

```

ILE SER ASP ASN GLY LYS VAL ILE ILE ARG ...
A T T C C G A T A A T G G C A A A G T G A T T A T T C G C ...
7150
... LEU SER GLY THR THR ASN SER GLN GLY LYS
... T T G T C A G G C A C A C C A A T A G T C A A G G T A A A
7160
...
7170...
7180
...
7190
7200

THR GLY VAL ALA ALA GLY VAL GLY TYR GLN ...
A C A G G C G T T G C A G C A G G T G T T G G T T A C C A G ...
7210
7220
... TRP ***
... T G G T A A T A G A A T T C C G G A T C C G C
7230...
7240
7250

```


FIG.25A

NTHi-Stamm 12 *hia*-Locus

```

      TYR  TYR  HIS  TRP  ***  PRO  THR  PRO  ...
G A A T T C T A T T A C C A C T G G T A A C C A A C A C C T ...
      10                                20      30 ...
      ... ALA  ALA  THR  PRO  GLU  THR  ALA  GLN  GLN  ILE
      ...G C T G C A A C G C C A G A A A C A G C A C A A A T T
      ...                                40      50      60

      HIS  TRP  LEU  HIS  GLN  PHE  THR  LYS  ALA  ARG  ...
C A C T G G C T A C A T C A A T T A C C A A A G C T C G C ...
      70                                80      90 ...
      ... ILE  GLN  TRP  ARG  LYS  THR  HIS  SER  LEU  PHE
      ...A T T C A A T G G C G C A A A A C C C A T T C C T T A T T C
      ...                                100      110      120

      PHE  LYS  GLU  LYS  PRO  ASP  TYR  ALA  PHE  VAL  ...
T T T A A A G A A A A A C C C G A T T A T G C C T T T G T G ...
      130                                140      150 ...
      ... LEU  ALA  GLU  ASN  GLY  LYS  VAL  GLN  GLU  ILE
      ...C T G G C A G A A A A C G G C A A A G T G C A A G A A T C
      ...                                160      170      180

      LYS  ALA  GLU  TYR  ARG  ARG  ILE  ALA  ASN  GLN  ...
A A A G C A G A A T A T C G C C G C A T T G C C A A T C A A ...
      190                                200      210 ...

```

FIG.25B

... ILE VAL GLU GLU ALA MET ILE ILE ALA ASN
 ...ATTGTGGAAGAGCAATGATTATTGCCAAC
 ... 220 230 240

ILE CYS ALA ALA GLN PHE LEU HIS GLU GLN ...
 ATCTGCCGCCCAATTTTTACACGACAG ...
 250 260 270 ...
 ... ALA LYS THR GLY ILE PHE ASN ALA HIS SER
 ...GCCAAACAGGCATTTTTCAACGCCCAAGC
 ... 280 290 300

GLY PHE ASP LYS LYS TYR LEU GLU ASN ALA ...
 GGT TTTGATAAATACTTAGAAATGCG ...
 310 320 330 ...
 ... HIS HIS PHE LEU MET ALA ASN LEU ALA ASN
 ... 6431.SL (...
 ...CACCA TTTCTTAATGGCAATAATTAGCCAAAT
 ... 340 350 360

GLU GLN ASN GLN THR GLU LEU ALA GLU ARG ...
 GACAAATCAACCTGAACCTGGCAGAACGT ...
 370 380 390 ...
 ... TYR SER VAL GLU ASN LEU ALA THR LEU ASN
 ...TATTCAGTAGAAACCTTAGCCACCTTAAC
 ... 400 410 420

FIG.25C

```

GLY  TYR  CYS  GLN  MET  ARG  HIS  ASP  ILE  GLU  ...
GGCTATTGCCCAAATGCCGTCA CGATATTGAA ...
430
... PRO  ILE  GLU  SER  ASP  TYR  LEU  GLU  LEU  ARG
...CCCATCCGAAAGCGATTATTAGAAC TGCGT
440
...
450
...
460
470
480

LEU  ARG  ARG  TYR  LEU  THR  PHE  ALA  GLU  PHE  ...
TTACGCCGTATTATACTTTCGCCGAATT ...
490
... LYS  SER  GLU  LEU  ALA  PRO  HIS  PHE  GLY  LEU
...AATCAGAAATTAGCACCGCACCTTGGTCTT
500
...
510
520
530
540

GLY  LEU  GLU  GLY  TYR  ALA  THR  TRP  THR  SER  ...
GGTTAGAAAGGCTATGCCCACTTGGA C ATCG ...
550
... PRO  ILE  ARG  LYS  TYR  SER  ASP  MET  VAL  ASN
...CCCATCCGCAATAATTCTCAGATATGGTTAAT
560
...
570
580
590
600

HIS  ARG  LEU  ILE  LYS  ALA  VAL  LEU  ALA  LYS  ...
...
CATCGCTTAATCAAGCCGTGCTGGCAAA ...
610
620
630
...
```

FIG.25D

```

... GLN PRO TYR GLU LYS PRO GLN ASN ASP VAL
...
...CAGCCCTTATGA A A A C C A A A T G A C G T G
... 640 650 660

LEU ALA ARG LEU GLN GLU SER ARG ARG GLN ...
6432.SL (
TTGGCACGTTTGGCAAGAGTCTCGCCGCA A ...
670 680 690 ...
... ASN ARG LEU VAL GLU ARG ASP ILE ALA ASP
...
...ATCGCCCTAGTGGACCGTGATATTGCCGAT
... 700 710 720

TRP LEU TYR CYS ARG TYR LEU ALA ASP LYS ...
TTGGCTATATTGGCCGTTTATCTTGCTGACAA A ...
730 740 750 ...
... VAL ALA GLU ASN VAL GLU PHE ASN ALA GLU
...GTGGCTGAA A A A TGTGGAATT T A A T G C A G A A
... 760 770 780

VAL GLN ASP VAL MET ARG ALA GLY LEU ARG ...
GTGCAAGATGTATA T G C G T G C A G G C T T A C G C ...
790 800 810 ...

```

FIG.25E

... VAL GLN LEU LEU GLU ASN GLY ALA SER LEU
 ...GTACAACTGCTCGAATAATGGTGCAATCGCTA
 ... 820 830 840

PHE ILE PRO ALA ALA THR LEU HIS ASN ASN ...
 TTTATTCTCCCGCCACGTTTGCAACAAC...
 850 860 870 ...

... LYS GLU GLU ILE GLN LEU ASN PRO ASP GLU
 ...AAGAGAAATAACAGCTAACCCTGACGAA
 ... 880 890 900

LEU ALA LEU TYR ILE LYS GLY GLU ARG THR ...
 CTCGCCCTCTATTATAAAGGCCGACGCACT...
 910 920 930 ...
 ... TYR LYS ILE GLY ASP ILE VAL LYS VAL LYS
 ...TACAAATAAGGCCGACATTGTGAAGTGAA
 ... 940 950 960

LEU THR GLU VAL LYS GLU ALA THR ARG SER ...
 CTCACAGAGTGAAAGAGACACTCGCAGT...
 970 980 990 ...
 ... ILE VAL GLY GLU ILE LEU GIN *** LEU PRO
 ...ATTGTGGCGGAATACTTCAATAATAATGCC
 ... 1000 1010 1020

FIG.25F

```

PHE  GLN  TYR  VAL  THR  GLU  ASP  GLY  LYS  THR...
G T T C C A A T A T G T T A C G G A A G A C G G C A A A C ...
1030                                     1040
... VAL VAL LYS VAL GLY ASN GLU TYR TYR GLU
...C G T T G T G A A A G T G G G C A A T G A G T A T T A C G A
...                                     1060
                                     1070
                                     1080

ALA  LYS  GLN  ASP  GLY  SER  ALA  ASP  MET  ASP...
A G C C A A G C A A G A C G G T T C G G C G G A T A T G G A ...
( 6295.SL
1090                                     1100
... LYS LYS VAL LYS ASN GLY GLU LEU VAL LYS
...T A A A A A G T C A A A A T G G C G A G C T G G T G A A
...
...                                     1120
                                     1130
                                     1140

THR  LYS  VAL  LYS  LEU  VAL  SER  ALA  ASN  GLY...
A A C T A A A G T G A A A T T G G T A T C G G C A A A C G G ...
1150                                     1160
... THR ASN PRO VAL LYS ILE SER ASN VAL ALA
...T A C A A A T C C G G T G A A A T C A G C A A T G T T G C
...                                     1180
                                     1190
                                     1200

GLU  GLY  THR  GLU  ASP  THR  ASP  ALA  VAL  SER...
G G A A G G C A C G G A A G A T A C C G A T G C G T C A G ...
1210                                     1220
                                     1230

```

FIG.25G

```

... PHE LYS GLN LEU LYS ALA LEU GLN ASN LYS
...C T T A G C A G T T G A A A G C C T T G C A A A C A A
...
1240
1250
1260

Gln Val Thr Leu Ser Ala Ser Asn Ala Tyr...
A C A G G T T A C G T T A A G C G C G A G C A A T G C T T A ...
1270
1280
... ALA ASN GLY GLY SER ASP ALA ASP VAL GLY
...T G C C A A T G G C G G T A G C G A T G C C G A C G T C G G
1290 ...
1300
1310
1320

Lys Val Thr Gln Thr Leu Ser Asn Gly Leu...
C A G G T A A C T C A A C T T T A G C C A A T G G T T T ...
1330
1340
... ASN PHE LYS PHE LYS SER THR ASP GLY GLU
...G A A T T T A A A T T T A A A T C C C A C A G A C G G C G A
...
1350 ...
1360
1370
1380

Leu Leu Asn Ile Lys Ala Asp Lys Asp Thr...
G T T G T T G A A C A T C A A A G C A G A C A G G A C A C ...
1390
1400
... Val Thr Ile Thr Arg Ala Ser Gly Ala Asn
...G G T T A C C A T T A C G C G G G C A A G C G G T G C G A A
...
1410 ...
1420
1430
1440

```

FIG.25H

```

GLY  ALA  ALA  ALA  THR  ASP  ALA  ASP  LYS  ILE...
TGG  TGG  GCG  GCG  GCG  ACT  GAT  GCG  CAC  AAG  AT ...
1450                                     1460 1470 ...
...  LYS  VAL  ALA  SER  ASP  GLY  ILE  SER  ALA  GLY
...TAA  AGT  GGC  TTC  AGA  CGG  CAT  TAG  CGCG  GG
...                                     1480 1490 1500

ASN  LYS  ALA  VAL  LYS  ASN  VAL  ALA  ALA  GLY...
TAA  TAA  AGC  AGT  TAA  A  AAC  GTC  GCG  CAG  G ...
1510                                     1520 1530 ...
...  GLU  ILE  SER  ALA  THR  SER  THR  ASP  ALA  ILE
...CGA  AAT  TTC  CGCC  ACT  TCC  ACC  GAT  GCG  AT
...                                     ( 6271.SL
...                                     1540 1550 1560

ASN  GLY  SER  GLN  LEU  TYR  ALA  VAL  ALA  LYS...
TAC  GGC  AGT  CAG  TTGT  GAT  GCG  CGT  GCG  CAA  A ...
1570                                     1580 1590 ...
...  GLY  VAL  THR  ASN  LEU  ALA  GLY  GLN  VAL  ASN
...GGG  GGT  AAC  AAC  CACT  TGC  TGG  AAC  AGT  GAA
...                                     1600 1610 1620

LYS  VAL  GLY  LYS  ARG  ALA  ASP  ALA  GLY  THR...
TAA  AGT  GGC  AAC  AAC  GTGC  AGAT  GCG  AGGT  AC ...
1630                                     1640 1650 ...

```


FIG.25I

... ALA SER ALA LEU ALA ALA SER GLN LEU PRO
 ...A G C A A G T G C A T T A G C G G C T T C A C A G T T A C C
 ... 1660 1670 1680

GLN ALA SER MET PRO GLY LYS SER MET VAL...
 A C A A G C C T C T A T G C C G G G T A A A T C A A T G G T ...
 1690 1700 1710 ...
 ... SER ILE ALA GLY SER SER TYR GLN GLY GLN
 ...T T C T A T T G C G G G A A G T A G T T A T C A A G G T C A
 ... 1720 1730 1740

SER GLY LEU ALA ILE GLY VAL SER ARG ILE...
 A A G T G G T T A G C C T A T C G G G G T A T C A A G A A T ...
 1750 1760 1770 ...
 ... SER ASP ASN GLY LYS LEU ILE ILE ARG LEU
 ...T T C C G A T A A T G G C A A A T T G A T T A T C G C C T T
 ... 1780 1790 1800

SER GLY THR THR ASN SER GLN GLY LYS THR...
 G T C A G G C A C A C C C A A T A G C C C A A G G T A A A C ...
 1810 1820 1830 ...
 ... GLY VAL ALA ALA GLY VAL GLY TYR GLN TRP
 ...A G G C G T T G C A G C A G G T G T T G G T T A C C A G T G
 ... 1840 1850 1860

*** ***
 G T A A T A G A A T T C
 1870

FIG.26A

ATG AAC AAA ATT TTT AAC GGT ATT TGG AAT GGT GTG ACT CAA ACT TGG Met Asn Lys Ile Phe Asn Val Ile Trp Asn Val Val Thr Gln Thr Trp 2130 2135 2140	48
GGT GTC GTA TCT GAA CTC ACT CCG ACC CAC ACC AAA TGC GCC TCC GCC Val Val Val Ser Glu Leu Thr Arg Thr His Thr Lys Cys Ala Ser Ala 2145 2150 2155	96
ACC GTG GCG GGT GGC GTA TIG GCA ACC CTG TIG TCC GCA ACG GGT GAG Thr Val Ala Val Ala Val Leu Ala Thr Leu Leu Ser Ala Thr Val Glu 2160 2165 2170 2175	144
GCG AAC AAC AAT ACT CCT GGT ACT ACG AAT AAG TIG AAG GCT TAT GCG GAT Ala Asn Asn Asn Thr Pro Val Thr Asn Lys Leu Lys Ala Tyr Gly Asp 2180 2185 2190	192
GCG AAT TTT AAT TTC ACT AAT AAT TCG ATA GCA GAT GCA GAA AAA CAA Ala Asn Phe Asn Phe Thr Asn Asn Ser Ile Ala Asp Ala Glu Lys Gln 2195 2200 2205	240
GGT CAA GAG GCT TAT AAA GGT TTA TTA AAT CTA AAT CAA AAA AAT GCG Val Gln Glu Ala Tyr Lys Gly Leu Leu Asn Leu Asn Glu Lys Asn Ala 2210 2215 2220	288

FIG.26B

AGT GAT AAA CTG TTG GTG GAG GAC AAT ACT GCG GCG ACC GTA GGC AAT	336
Ser Asp Lys Leu Val Glu Asp Asn Thr Ala Ala Thr Val Gly Asn	
2225 2230 2235	
TTG CGT AAA TTG GGC TGG GTA TTG TCT AGC AAA AAC GGC ACA AGG AAC	384
Leu Arg Lys Leu Gly Trp Val Leu Ser Ser Lys Asn Gly Thr Arg Asn	
2240 2245 2250 2255	
GAG AAA AGC CAA CAA GTC AAA CAT CCG GAT GAA GTG TTG TTT GAA GGC	432
Glu Lys Ser Gln Val Lys His Ala Asp Glu Val Leu Phe Glu Gly	
2260 2265 2270	
AAA GGC GGT GTG CAG GTT ACT TCC ACC TCT GAA AAC GGC AAA CAC ACC	480
Lys Gly Gly Val Gln Val Thr Ser Thr Ser Glu Asn Gly Lys His Thr	
2275 2280 2285	
ATT ACC TTT GCT TTA GCG AAA GAC CTT GGT GTG AAA ACT GCG ACT GTG	528
Ile Thr Phe Ala Leu Ala Lys Asp Leu Gly Val Lys Thr Ala Thr Val	
2290 2295 2300	
AGT GAT ACC TTA ACG ATT GGC GGT GGT GCT GCT GCA GGT GCT ACA ACA	576
Ser Asp Thr Leu Thr Ile Gly Gly Gly Ala Ala Ala Gly Ala Thr Thr	
2305 2310 2315	

FIG.26C

ACA CCG AAA GTG AAT GTA ACT AGT ACA ACT GAT GGC TTG AAG TTC GCT	624
Thr Pro Lys Val Asn Val Thr Ser Thr Thr Asp Gly Leu Lys Phe Ala	
2320 2325 2330 2335	
AAA GAT GCT GCG GGT GCT AAT GGC GAT ACT ACG GTT CAC TTG AAT GGT	672
Lys Asp Ala Ala Gly Ala Asn Gly Asp Thr Thr Val His Leu Asn Gly	
2340 2345 2350	
ATT GGT TCA ACC TTG ACA GAC ACG CTT GTG GGT TCT OCT GCT ACT CAT	720
Ile Gly Ser Thr Leu Thr Asp Thr Leu Val Gly Ser Pro Ala Thr His	
2355 2360 2365	
ATT GAC GGA GGA GAT CAA AGT ACG CAT TAC ACT CGT GCA GCA AGT ATC	768
Ile Asp Gly Gly Asp Gln Ser Thr His Tyr Thr Arg Ala Ala Ser Ile	
2370 2375 2380	
AAG GAT GTC TTG AAT GCG GGT TGG AAT ATC AAG GGT GTT AAA GCT GGC	816
Lys Asp Val Leu Asn Ala Gly Trp Asn Ile Lys Gly Val Lys Ala Gly	
2385 2390 2395	
TCA ACA ACT GGT CAA TCA GAA AAT GTC GAT TTT GTT CAT ACT TAC GAT	864
Ser Thr Thr Gly Gln Ser Glu Asn Val Asp Phe Val His Thr Tyr Asp	
2400 2405 2410 2415	

FIG.26D

ACT GTT GAG TTC TTG AGT GCG GAT ACA GAG ACC ACG ACT GTT ACT GTA	912
Thr Val Glu Phe Leu Ser Ala Asp Thr Glu Thr Thr Thr Val Thr Val	
2420 2425 2430	
GAT AGC AAA GAA AAC GGT AAG AGA ACC GAA GTT AAA ATC GGT GCG AAG	960
Asp Ser Lys Glu Asn Gly Lys Arg Thr Glu Val Lys Ile Gly Ala Lys	
2435 2440 2445	
ACT TCT GTT ATC AAA GAA AAA GAC GGT AAG TTA TTT ACT CGA AAA GCT	1008
Thr Ser Val Ile Lys Lys Glu Lys Asp Gly Lys Leu Phe Thr Gly Lys Ala	
2450 2455 2460	
AAC AAA GAG ACA AAT AAA GTT GAT GGT GCT AAC GCG ACT GAA GAT GCA	1056
Asn Lys Glu Thr Asn Lys Val Asp Gly Ala Asn Ala Thr Glu Asp Ala	
2465 2470 2475	
GAC GAA GGC AAA GGC TTA GTG ACT GCG AAA GAT GIG ATT GAC GCA GIG	1104
Asp Glu Gly Lys Gly Leu Val Thr Ala Lys Asp Val Ile Asp Ala Val	
2480 2485 2490 2495	
AAT AAG ACT GGT TGG AGA ATT AAA ACA ACC GAT GCT AAT GGT CAA AAT	1152
Asn Lys Thr Gly Thr Arg Ile Lys Thr Thr Asp Ala Asn Gly Gln Asn	
2500 2505 2510	

FIG.26E

GCC GAC TTC GCA ACT GTT GCA TCA GGC ACA AAT GTA ACC TTT GCT AGT Gly Asp Phe Ala Thr Val Ala Ser Gly Thr Asn Val Thr Phe Ala Ser	1200
2515 2520 2525	
GGT AAT GGT ACA ACT GCG ACT GTA ACT AAT GGC ACC GAT GGT ATT ACC Gly Asn Gly Thr Thr Ala Thr Val Thr Asn Gly Thr Asp Gly Ile Thr	1248
2530 2535 2540	
GTT AAG TAT GAT GCG AAA GTT GGC GAC GGC TTA AAA CTA GAT GGC GAT Val Lys Tyr Asp Ala Lys Val Gly Asp Gly Leu Lys Leu Asp Gly Asp	1296
2545 2550 2555	
AAA ATC GCT GCA GAT ACG ACC GCA CTT ACT GTG AAT GAT GGT AAG AAC Lys Ile Ala Ala Asp Thr Thr Thr Ala Leu Thr Val Asn Asp Gly Lys Asn	1344
2560 2565 2570 2575	
GCT AAT AAT CCG AAA GGT AAA GTG GCT GAT GTT GCT TCA ACT GAC GAG Ala Asn Asn Pro Lys Gly Lys Val Ala Asp Val Ala Ser Thr Asp Glu	1392
2580 2585 2590	
AAG AAA TIG GTT ACA GCA AAA GGT TTA GTA ACA GCC TTA AAC AGT CTA Lys Lys Leu Val Thr Ala Lys Gly Leu Val Thr Ala Leu Asn Ser Leu	1440
2595 2600 2605	

FIG.26F

AGC TGG ACT ACA ACT GCT GCT GAG GCG GAC GGT GGT ACG CTT GAT GGA	1488
Ser Trp Thr Thr Thr Ala Ala Glu Ala Asp Gly Thr Leu Asp Gly	
2610 2615 2620	
AAT GCA AGT GAG CAA GAA GTT AAA GCG GCG GAT AAA GTA ACC TTT AAA	1536
Asn Ala Ser Glu Gln Glu Val Lys Ala Ala Gly Asp Lys Val Thr Phe Lys	
2625 2630 2635	
GCA GCG AAG AAC TTA AAA GTG AAA CAA GAG GGT GCG AAC TTT ACT TAT	1584
Ala Gly Lys Asn Leu Lys Val Lys Gln Glu Gly Ala Asn Phe Thr Tyr	
2640 2645 2650 2655	
TCA CTG CAA GAT GCT TTA ACA GCG TTA ACG AGC ATT ACT TTA GGT ACA	1632
Ser Leu Gln Asp Ala Leu Thr Gly Leu Thr Ser Ile Thr Leu Gly Thr	
2660 2665 2670	
GGA AAT AAT GGT GCG AAA ACT GAA ATC AAC AAA GAC GGC TTA ACC ATC	1680
Gly Asn Asn Gly Ala Lys Thr Glu Ile Asn Lys Asp Gly Leu Thr Ile	
2675 2680 2685	
ACA CCA GCA AAT GGT GCG GGT GCA AAT AAT GCA AAC ACC ATC AGC GTA	1728
Thr Pro Ala Asn Gly Ala Gly Ala Asn Asn Ala Asn Thr Ile Ser Val	
2690 2695 2700	

FIG.26G

ACC AAA GAC GGC ATT AGT GCG GCG GGT CAG TCG GTT AAA AAC GTT GTC	1776
Thr Lys Asp Gly Ile Ser Ala Gly Gln Ser Val Lys Asn Val Val	
2705 2710 2715	
AGC GGA CTG AAG AAA TTT GGT GAT GCG AAT TTC GAT CCG CTG ACT AGC	1824
Ser Gly Leu Lys Lys Phe Gly Asp Ala Asn Phe Asp Pro Leu Thr Ser	
2720 2725 2730 2735	
TCC GCC GAC AAC TTA ACG AAA CAA AAT GAC GAT GCC TAT AAA GGC TTG	1872
Ser Ala Asp Asn Leu Thr Lys Lys Gln Asn Asp Asp Ala Tyr Lys Gly Leu	
2740 2745 2750	
ACC AAT TTG GAT GAA AAA GGT ACA GAC AAG CAA ACT CCA GTT GTT GCC	1920
Thr Asn Leu Asp Glu Lys Lys Gly Thr Asp Lys Gln Thr Pro Val Val Ala	
2755 2760 2765	
GAC AAT ACC GCC GCA ACC GTC GCG GAT TTG CCG GCC TTG GCG TGG GTC	1968
Asp Asn Thr Ala Ala Thr Val Gly Asp Leu Arg Gly Leu Gly Trp Val	
2770 2775 2780	
ATT TCT GCG GAC AAA ACC ACA GCG GCG TCA ACG GAA TAT CAC GAT CAA	2016
Ile Ser Ala Asp Lys Thr Thr Gly Gly Ser Thr Glu Tyr His Asp Gln	
2785 2790 2795	

FIG.26H

GTT CCG AAT GCG AAC GAA GTG AAA TTC AAA AGC GGC AAC GGT ATC AAT Val Arg Asn Ala Asn Glu Val Lys Phe Lys Ser Gly Asn Gly Ile Asn 2800 2805 2810 2815	2064
GTT TOC GGT AAA ACG GTC AAC GGT AGG CGT GAA ATT ACT TTT GAA TTG Val Ser Gly Lys Thr Val Asn Gly Arg Arg Glu Ile Thr Phe Glu Leu 2820 2825 2830	2112
GCT AAA GGT GAA GTG GTT AAA TCG AAT GAA TTT ACC GTC AAA GAA ACC Ala Lys Gly Glu Val Val Lys Ser Asn Glu Phe Thr Val Lys Glu Thr 2835 2840 2845	2160
AAT GGA AAG GAA ACG AGC CTG GTT AAA GTT GGC GAT AAA TAT TAC AGC Asn Gly Lys Glu Thr Ser Leu Val Lys Val Gly Asp Lys Tyr Tyr Ser 2850 2855 2860	2208
AAA GAG GAT ATT GAC TTA ACA ACA GGT CAG CCT AAA TTA AAA GAT GGC Lys Glu Asp Ile Asp Leu Thr Thr Gly Gln Pro Lys Leu Lys Asp Gly 2865 2870 2875	2256
AAT ACA GTT GCT GCG AAA TAT CAA GAT AAA GGT GGC AAA GTC GTT TCT Asn Thr Val Ala Ala Lys Tyr Gln Asp Lys Gly Gly Lys Val Val Ser 2880 2885 2890 2895	2304

FIG.26I

GTA ACG GAT AAT ACT GAA GCT ACC ATA ACC AAC AAA GGT TCT GGC TAT Val Thr Asp Asn Thr Glu Ala Thr Ile Thr Asn Lys Gly Ser Gly Tyr	2352
	2900 2905 2910
GTA ACA GGT AAC CAA GTG GCA GAT GCG ATT GCG AAA TCA GGC TTT GAG Val Thr Gly Asn Gln Val Ala Asp Ala Ile Ala Lys Ser Gly Phe Glu	2400
	2915 2920 2925
CTT GGC TTG GCT GAT GAA GCT GAT GCG AAA CCG GCG TTT GAT GAT AAG Leu Gly Leu Ala Asp Glu Ala Asp Ala Lys Arg Ala Phe Asp Asp Lys	2448
	2930 2935 2940
ACA AAA GCC TTA TCT GCT GGT ACA ACG GAA ATT GTA AAT GCC CAC GAT Thr Lys Ala Leu Ser Ala Gly Thr Thr Glu Ile Val Asn Ala His Asp	2496
	2945 2950 2955
AAA GTC CGT TTT GCT AAT GGT TTA AAT ACC AAA GTG ACG GCG GCA ACG Lys Val Arg Phe Ala Asn Gly Leu Asn Thr Lys Val Ser Ala Ala Thr	2544
	2960 2965 2970 2975
GTG GAA AGC ACC GAT GCA AAC GCG GAT AAA GTG ACC ACA ACC TTT GTG Val Glu Ser Thr Asp Ala Asn Gly Asp Lys Val Thr Thr Thr Phe Val	2592
	2980 2985 2990

FIG.26J

AAA ACC GAT GIG GAA TTG CCT TTA ACG CAA ATC TAC AAT ACC GAT GCA 2640
 Lys Thr Asp Val Glu Leu Pro Leu Thr Gln Ile Tyr Asn Thr Asp Ala
 2995 3000 3005

AAC GGT AAG AAA ATC ACT AAA GTT GTC AAA GAT GGG CAA ACT AAA TGG 2688
 Asn Gly Lys Lys Ile Thr Lys Val Val Lys Asp Gly Gln Thr Lys Trp
 3010 3015 3020

TAT GAA CTG AAT GCT GAC GGT ACG GCT GAT ATG ACC AAA GAA GTT ACC 2736
 Tyr Glu Leu Asn Ala Asp Gly Thr Ala Asp Met Thr Lys Glu Val Thr
 3025 3030 3035

CTC GGT AAC GIG GAT TCA GAC GGC AAG AAA GTT GTG AAA GAC AAC GAT 2784
 Leu Gly Asn Val Asp Ser Asp Gly Lys Lys Val Val Lys Asp Asn Asp
 3040 3045 3050 3055

GGC AAG TGG TAT CAC GCC AAA GCT GAC GGT ACT GCG GAT AAA ACC AAA 2832
 Gly Lys Trp Tyr His Ala Lys Ala Asp Gly Thr Ala Asp Lys Thr Lys
 3060 3065 3070

GGC GAA GIG AGC AAT GAT AAA GTT TCT ACC GAT GAA AAA CAC GTT GTC 2880
 Gly Glu Val Ser Asn Asp Lys Val Ser Thr Asp Glu Lys His Val Val
 3075 3080 3085

FIG.26K

AGC CTT GAT CCA AAT GAT CAA TCA AAA GGT AAA GGT GTC GTC ATT GAC 2928
 Ser Leu Asp Pro Asn Asp Gln Ser Lys Gly Lys Gly Val Val Ile Asp
 3090 3095 3100
 AAT GIG GCT AAT GGC GAT ATT TCT GCC ACT TCC ACC GAT GCG ATT AAC 2976
 Asn Val Ala Asn Gly Asp Ile Ser Ala Thr Ser Thr Asp Ala Ile Asn
 3105 3110 3115
 GCA AGT CAG TTG TAT GCT GIG GCA AAA GGG GTA ACA AAC CTT GCT GCA 3024
 Gly Ser Gln Leu Tyr Ala Val Ala Lys Gly Val Thr Asn Leu Ala Gly
 3120 3125 3130 3135
 CAA GIG AAT AAT CTT GAG GGC AAA GIG AAT AAA GIG GGC AAA CGT GCA 3072
 Gln Val Asn Asn Leu Glu Gly Lys Val Asn Lys Val Gly Lys Arg Ala
 3140 3145 3150
 GAT GCA GGT ACA GCA AGT GCA TTA GCG GCT TCA CAG TTA CCA CAA GCC 3120
 Asp Ala Gly Thr Ala Ser Ala Leu Ala Ala Ser Gln Leu Pro Gln Ala
 3155 3160 3165
 ACT ATG CCA GGT AAA TCA ATG GTT GCT ATT GCG GGA AGT AGT TAT CAA 3168
 Thr Met Pro Gly Lys Ser Met Val Ala Ile Ala Gly Ser Ser Tyr Gln
 3170 3175 3180

FIG.26L

GGT CAA AAT GGT TTA GCT ATC GCG GTA TCA AGA ATT TCC GAT AAT GGC	3216
Gly Gln Asn Gly Leu Ala Ile Gly Val Ser Arg Ile Ser Asp Asn Gly	
3185 3190 3195	
AAA GTG ATT ATT CCG TIG TCA GCG ACA ACC AAT AGT CAA GGT AAA ACA	3264
Lys Val Ile Ile Arg Leu Ser Gly Thr Thr Asn Ser Gln Gly Lys Thr	
3200 3205 3210 3215	
GGC GTT GCA GCA GGT GTT GGT TAC CAG TGG	3294
Gly Val Ala Ala Gly Val Gly Tyr Gln Trp	
3220 3225	

FIG.27A

Ausrichtung des 5'-ORFs von NTHi-Stamm 12 mit HI1733 aus H. influenzae-Stamm Rd

X	10	20	30	40	50	60	70
	PTPAATPETAQQIHWLHQFTKARIQWRKTHSLFFKEKPDYAFVLAENCKVQEIKAERYRRIANQIVEEFAMIIA						
	AWQPEMPETAQQIHWLHQFTKARIQWRKTHSLFFKEKPDYAFVLAENCKVQEIKAERYRRIANQIVEEFAMIIA						
330	340	350	360	370	380	390	400
80	90	100	110	120	130	140	
	NICAQFLHEQAKTIGIFNAHSGFDKKYLENAHHFLMANLANEQNQTELAERYSVENLATLNGYCQMRHDIEP						
	NICAQFLHEQAKTIGIFNTHSGFDKKFLFNAHFLMANLANEQNQTELAERYSVENLATLNGYCQMRHDIEP						
410	420	430	440	450	460	470	
150	160	170	180	190	200	210	
	IESDYLETLRRLRYLTFAEFKSEIAPHFGLGEGYAIWISPIRKYSIDMNHRLIKAVLAKQPYEKPQNDVLAR						
	IESDYLETLRRLRYLTFAEFKSEIAPHFGLGEGYAIWISPIRKYSIDMNHRLIKAVLAKQPYEKPQNDVLAR						
480	490	500	510	520	530	540	
220	230	240	250	260	270	280	
	LQESRRQNRLVERDIADWLYCRYLADKVAENVEFNAEVQDMRAGLRVQLLENGASLFIIPAATLHNNKEETIQ						
	LQFARRQNRLVERDIADWLYCRYLADKVASNAEFAEVQDMRAGLRVQLLENGASLFIIPAATLHNNKEETIQ						
550	560	570	580	590	600	610	

FIG.27B

290 300 310 320 330
LNPDELALYIKGERTYKIGDIVKVKLTVEKEATRSIVGEILQ
|||||
LNPDELALYIKGERTYKIGDIVKVKLTVEKEATRSIVGEILQ
620 630 640 650 X

; ## Querverweise GB:L42023; TIGR:HI1733
; ## Anmerkung benannt als Homologes zu einem Protein aus Escherichia coli
; ZUSAMMENFASSUNG #Länge 659 #Molekulargewicht 75782 #Checksum 8365

A64139
MFQDNPLLAQLKQIHDSEKEQVEGVVKSTDKAYGFLECDKKTTFIAPPSMKKVMHGDKIKATIEKQGDKE
QAEPEALIEFMLTRFLAKVRFNKDKKLQVLVDHPSINQPIGAQQAKSVKEELQEGDWVAVNLKTHPLRDD
RFFYATINQLICRADDELAPWMTLARHEQSRYPVRCGAEPYEMLDQKTRENLTALHFVTIDSESTMDMD
ALYTIEPIAQNSTIQMKLVVAIADPTAYIALDSQIEQFAKQRCFTINILPGFNILMLPRELSDELCSLIAN
ETRPALVCYTEITDLTGNTITAKPHFVSAYVQSKAKLAYNKVSDYLEQADNAWQPEMPETAQQIHMLHQFTK
ARIQWRKTHSLFFKEKPDYAFVLAENGKVQETKAEYRRIANQIVVEEAMI IANICAAQFLHEQAKTIGIFNT
HSGFDKKFL ENAHNFTMANLANEQNQTELAERYSVENLATLNGYCQMRHDIEPIESDYLETILRRYLTFE
EFKSELAPHFGGLGEGYATWTSPIRKYSIDMNNRLIKAVLAKQPYEKPQNDV/LARLQFAPRQNRLVERDI
ADWLCRYLADKVASNAEFAEVQDVMPAGLRVQLLENGASLFI PAATTIHNKKEETQINPDEI ALYIKGE
RTYKIGDMVKVLTVEKEATRSIVGEILQ

FIG. 28B

```

... ..RTAPVLSFHSDKEGTGEKEVTENSNGIYFFHNKGVLKA-----
... ..GTFVKVQSTEDIEDSAATKDDNKNNQALKAGDTLTLKA-----
... ..GSAYAQQKDKTHIAIGEONQPPRRSGTAKADGDRAIAGENANAQGG
... ..GSAYAQN-NSK-AIFGTTGNNDN---ASASNEASIAIGSLAKAHAN
... ..      *      *      *      *      *      *

```

.....

GAITLKAGDNLKIKQNIDESTINASSFTYSLKKDLTDLTSVATEKLSFGANGKVDITS DANG...
 GAITLKAGDNLKIKQ-----STINASSFTYSLKKDLTDLTSVATEKLSFGANGKVDITS DANG...
 GKN-LKAKLDQCGKSVTFALAKDLVDKTAKVSDTLTIGENTPAAGGATP---KVSITSTADG...
 QALAI GSSNKT VNG-SSLDKIGTDTIQESIAIGGDVKASGDASIAIGSDDLHLLDQHGNPK...
 QALAI GGSKPDPRNQANQKAGSHAKGESIAIGGDVLAEGDASIAIGSDDL YLDRNSTNSK...

•
•
•

*
*
*
*
*
*
*
*

*
*

33

32

29

K22

M4071

11

[illegible]

FIG. 28D

M4071
11
K9
HSF
API
Rd
4223
LES-1

[illegible]

FIG. 28E

```

32  ---
29  ---
K22 ---
M4071 ---
11  ---
K9  ---
HSF ..KDINGNGITIVKYD-ALVGDGLKFDSDKKIVADTTALTIVTG----
API ..KDINGNGITIVKYD-ALVGDGLKFDSDKKIVADTTALTIVTG----
Rd  ...
Rd  ...NSTD--GITIVKYE-ALVGDGLKIDGDQKIVADTTALTIVTG----
4223 ...AETNA--LITDNN-IGVVKEADNSGLKVKLAKTINNLTETVTTTL
LES-1 ...-ETQADKLTDDNNNIGVVTD-NNIGLKVKLAKNLSGLETIVSTKNL
... **** * * * * * ** *
-----
-----
-----
-----
-----
-----
-----
-----
GKVAEIAKEDDKKKLVNAGDLVVTALGNLSWKAKAEADTD--GALEGISKDQEVKAGEIVTFK...
GKVAEIAKEDDKKKLVNAGDLVVTALGNLSWKAKAEADTDTDGALEGISKDQEVKAGEIVTFK...
GKVAEIAKEDDKKKLVNAGDLVVTALGNLSWKAKAEADTD--GALEGTSKDQEVKAGEIVTFK...
NATTTIVKVGSSSITAEILLSDSLTFTQPNIGSQSTSKTIVGVNGVKFTINNAEITTAIGTT-R...

```

*

TASEKVTIVGSGNN-TAELOSGGLTFT-PTINA-STDKTVYGIDGLKFTDNSN-TALEDTT-R...
 * * * * * * * * * * * * * *
 ...-----
 ...-----
 ...-----
 ...-----
 ...-----
 ...-----
 ...-----
 ...-----
 ...-----
 ...AGKNLKVQDGANFTYSLQDALIGLTSATLGGTTINGENDA
 ...AGKNLKVQDGANFTYSLQDALIGLTSATLGGTTINGENDA
 ...AGKNLKVQDGANFTYSLQDALIGLTSATLGGTTINGENDA
 ...ITRDKIGEARD-GDWDE-----
 ...ITKDKIGFSENKAGTVDENKPYLDKDKLKVENSTLNNGGLT
 *** * * * * *
 ...

60 70 ...

-----AGSTTGINSINVYCK-NNSNFNSANNISIA...

 ---NM--SV.G.A.D--T..TT.....
 ---N--SV.G.A.D--T..TT.....
 ---T.G.S.G.KA.ST.P...A.G.AT...
 ---N--PV.K.KA.D-A.F.--T.....
 ---ASV.CR....D-T.TK.A.....
 KTVINKDGLTITPACNGGTGTGNTISBTKDGIK.NKAI.VASGLRAYDDA.DVL...AT...

FIG.28G

```

KTVINKDGLTITPAGNGGTIGTINTISBTKDGIK..NKAI..VASGLRAYDDA..DVL...AT...
KTVINKDGLTITPAGNGGTIGTINTISBTKDGIK..NKAI..VASGLRAYDDA..DVL...AT...
-----
-----VNNITIGGSNKQIQVGADGIKFDVNNVNSNAKFGTTRITTEEEIGFAD....
** *** * ** ***** * * * * *
...80      90      100      110      120
...DLNKQNDVYDGLLNINEKGTDKSKFLVADETTATVGNLRKL-----
...-----ATDENED..EELEPVQRSV.-----
...E.HVQDA.K.....D.N..S.....N.A.....
...E.HVQDA.K.....D.N..S.....N.A.....
...AR.F.GA.....DAN.N-L..T.DKA.....
...AE..VOEA.K.....NAS-D.L..E.N.A....D.....
...G.H.....N.AN..-L..D.N.A....D.....
...RHVEDA.K.....NAN.QP-.T.S.A....D.....
...RHVEDA.K.....NAN.QP-.TDS.A....D.....
...RHVEDA.K.....NAN.QP-.....S.A....D.....
...-----KQAP.LDKKQ.KVGSVAITIDNGL.AGNKKIS..A.GSSANDA
...GKVDKK.P.LDKKQ.QVG.VKIT.DSGINAGDQKISNVKDATDDTDA
...*      *      **      *      *
...
130      140      150      160      ...
GwVSTKNSTKEE-SNQVKQADEVLFEFEG-KDGVITVTSKSENGKHTVT-----
R.SFKSAKEGTG.QEGTTEV-----
...L.S..G.RN.K.Y.....T.-SGAA..S.S.KD....I.-----
...L.S..G.RN.K.Y.....T.-SGAA..S.S.KD....I.-----

```

33
32
29
K22
M4071
11
K9
HSF
API
Rd
4223
LFS-1

```
L.S.G.RN.K.Q..H.....-.....  
L.S.G.RN.K.Q..H.....-G.Q.T....I.-..  
GKEN.K.Q.....K.S.G.Q.T....AI.....  
G.....T.-AGAA.....I.VSVAETKADOGLEKD.  
G.....T.-AGAA.....I.VSVAETKADSGLERD.  
G.....T.-AGAA.....I.VSVAETKADSGLERD.  
VTIEQL.AAKPTINAGAGISVTPTEISVDAKSEN..APTY.IGVKT.ELNSDGTSDKFSVG...  
VTYKQL.....
```

```

...
33  -----
32  -----
29  -----
K22 -----
M4071 -----
11  -----
K9  -----
HSF  ....GDTIKLKVDNQNTDNLTVGNNGTAVTKGGFEIVTKGATDADRGKVT
API  ....GDTIKLKVDNQNTDNLTVGNNGTAVTKGGFEIVTKGATDADRGKVT
Rd   ....GDTIKLKVDNQNTDNLTVGNNGTAVTKGGFEIVTKGATDADRGKVT
4223 ....SGTNSLVTAEHLASYLNEVNRTPADSAŁSF-TVKEED-DDDANAİT
LES-1 ....-----QVQQDADGALQSF-SİRDEK-GQEFİTİSN
...  * * * * *
...  * * * * *
-----

```

FIG.28I

-----	...	
-----	...	
-----	...	
-----	...	
-----	...	
-----	...	
-----	...	
VKDATANDADKKVATVKDVATAINSAATFVKTENLTTSIDEDNPTDNGKDDALKAGDILTFK...		
VKDATANDADKKVATVKDVATAINSAATFVKTENLTTSIDEDNPTDNGKDDALKAGDILTFK...		
VKDATANDADKKVATVKDVATAINSAATFVKTENLTTLDEADAKDQG-DDALKAGDILTFK...		
VAKDTTKNAGAVSILKLGKNGLTVATKGD-GIVTFGLSQDSGLTIGKSTLNNDGLTVKDN...		
LYSNGVTPNPFETITFA-GENGISISNDIAKGKVKVGIDPINGLITTPKLTJVGSDKDKGTQLV...		
* * *	** * *	...
...	-----	33
...	-----	32
...	-----	29
...	-----	K22
...	-----	M4071
...	-----	11
...	-----	K9
...	-----	HSF
...AGKNLKVKRDGKNITFDLAKNLEVKTKVSDTLTIGCVTPGGTTAT--		API
...AGKNLKVKRDGKNITFDLAKNLEVKTKVSDTLTIGCVTPGGTTAT--		Rd
...AGKNLKVKRDGKNITFDLAKNLEVKTKATFSDRLTIG-----		4223
...-EQIQVGANGI.FINWNGSNPGIGIANITARITRDKIGFAGSDGAVDINK		LES-1
...IEQVASGN-.T...IR-----		
...	* * *	*
	* * *	*
	* * *	*
	* * *	*

[illegible]

PYLDQKLOQVNWKITNTGINAGGKAITGLSPTLPSIADQSS-RNIELGNTI-QDKDKSNAA...
 -----GLSPTL[SITNAGGVRTTEQGNITITSDEKSKAA...

*
*
*
*
*

*

*
*
*

*
*

*
*
*
*

*
*
*

*
*

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

•
•
•

.....

.....

A

•
•
•

...

.....

STEDVI BACITNOG CENNI DVATVETI NEI DOSTG

STEFAN LUTHE

...STEDVTRAGWLNQNGENNDYVATYDVTNFIIDDSIG

.....

STNDT I NTCENT KNNNSNS/GEV/STVNTVDFEIDGNAT

...:JINDHINIGI WITVANS, G. D. H. N. DE LEEUW
CROON, T. VAN G. ICHTJE, H. DEIDAM

...SIGDILNIGFNLKNNNSVGFVSIIYNI VDF IJGNAI

33
32
29
K22
M4071
11
K9
HSF
API
Rd
4223
LES-1

FIG. 28K

... ** *** ***** * * * **** * * * *

KGADV KIGAKTSV I KD HNGKLF T GDLKDAN NGATV SEDDGK DGTGLVTAKTV IDAVNKS G...

KGADV KIGAKTSV I KD HNGKLF T GDLKDAN NGATV SEDDGK DGTGLVTAKTV IDAVNKS G...

TSKV VYDVNVDDTTIHLTGTDNK - KLGVKTTLNKT DAN GTATNFVNSSDE DALVNAKD...

TSKV TVD VNVD EKTIELJGDNGKINKIGVKTPTLTITINANGK - AINF - SITINDALVNAKDI...

* * * * * *

33 -----

32 -----

29 -----

K22 -----

M4071 -----

11 -----

K9 -----

HSF WRVTGEGATAETGCATAVNACNAETVTSGTSVNFKNGNATTATVSKDN GNIN

API WRVTGEGATAETGCATAVNACNAETVTSGTSVNFKNGNATTATVSKDN GNIN

Rd -----

FIG. 28L

[illegible]

FIG.28M

...WTAKADKYADGESEGETDQEVKAGDKVTF-KAGKNLKVQSEKDFYSLQD	API
...	Rd
...TTSGLKAGKST-LNDGGLSIKNPTGSEQIQVGADG	4223
...TQSGLKAGDSTTLNKDGKSIKNPASNEQIQVGADG	LES-1
... * **** * * * * *	

TLTGLTSITLGGTANGRNDTGIVINKDGLTITLANGAAAGTDA SNGT----ISVTKDGISA...	
TLTGLTSITLGGTANGRNDTGIVINKDGLTITLANGAAAGTDA SNGT----ISVTKDGISA...	

VKFAKVNNGVVGAGIDGTTRI TRDEIGFTGTCNGSLDKSKPHL-----SLDGINA...	
VKFAKVDK-CNSSTGIDGTSRI TKDQIGFTGANGSLDTTKPHLTKDKLKVGEVEITNTGINA...	
* ** * * * * * * * * *	
...	33
...	32
...	29
...	K22
...	M4071
...	11

FIG. 28N

[illegible]

FIG. 280

[illegible]

FIG.28P

...	DAITGGQVNAD-RGKVK-----AEDENGADVDDKKV-----	K22
...	-----	M4071
...	-----	11
...	-----	K9
...	DAITTAQGTNANERGVVKGSGNGATATETDDKKV-----	HSF
...	DAITTAQGTNANERGVVKGSGNGATATETDDKKV-----	API
...	QNGQNTITGLSNITLANVINDKGSVRTTEQGNIIKDEDKTRA	Rd
...	KDGQNTITGLSNITLANVINDGAGHSLS-QGLAN-DTDKTRA	LES-1
...	-----*	
...	-----	
...	INLNTDSSGNAVGSSTITFKAGNLIKQSCN...	
ATVGDVAKAINDAAITFVKVESTDDDIENGAAGKNETTDQALKAGDTLTLKAGKNLKAQLDQN...		
ATVGDVAKAINDAAITFVKVESTDDDIENGAAGKNETTDQALKAGDTLTLKAGKNLKAQLDQN...		

ATVGDVAKAINDAAITFVKVEN-DDSATIDDSPTDDGANDALKAGDTLTLKAGKNLKVRRDG-...		
ATVGDVAKAINDAAITFVKVEN-DDSATIDDSPTDDGANDALKAGDTLTLKAGKNLKVRRDG-...		

ASIVDVLSAGFNLCNGEAVDFVSTYDITVNFADGNATTAKVITYDDTSKTSKWYDVNVDDTT...		
ASIGDVINAGFNLCNGEAVDFVSTYDITVDFIDGNATTAKVITYDDTSKTSKWYDVNVDDNKT...		
*** *	* *	
...	170 180 190 200	

FIG.28Q

```

...-----FALANDLVKNATVSDKLSLGANGKKVDITSDANG-----
...D--FTYS.KKE.KNLTSEITE...F...N.....
...GKSVT...K.D.TS.K.....I.KDIN.....
...GKSVT...K.D.TS.K.....I.KDIN.....
...-----T.EK.....N.....T.....
...-----K.G.T...T.TI.GGAAAGAT.TPKVNVSTTDG
...-----K..SMRT...T.TI.GSITIGSA.TPKVNVSTASG
...-KNIT...S..S.....T.N.N...TK.-----
...-KNIT...S..S.....T.N.N...TK.-----
...-----N.....T.....
...IEVK-DKKLGVKTTTTLTSTGTGANKFALSNOATGDALVKASDIVA--
...IEVTSDDKKLGVKTTTLTKTSANGNATKFSA-ADGDALVKASDIAT--
...
...      * * * * *
210      220      230      240      250      ...
LKFAKQGT-NGQNGN--VHLNGIASTLDDPRVGGKTAHLTKEISDTERN--RAASVGDVLNA...
..L..T.NG....S.--.....T.TLA.T.G.VDIN.DAVNYH--.....Q....S...
..L..T.NG....--.....T.TIT.MT.QASNGVAVQ-NH--.....A.....
..L..T.NG....--.....T.TIT.MT.QASNGVAVQ-NH--.....A.....
.....PS.-.....T.TIT.TTKSATNGVDVQNH-...A.....
.....DAA--A..DIT.....G...T.TK..SPAT.IDGDOQS.HYT--...IK.....
.V...GA.GANGDIT--...TN....Q.TLLATGVVSKLDGNGITADEKK....Q....S...
.....DSKT-.DDA.--I.....T.TLLNSGATTNLGGGITTNEKK....K.....
..x..DSKT-.DDA.--I.....T.TLLNSGATTNLGGGITTNEKK....K.....
.....P.-.....--.....

```

33
32
29
K22
M4071
11
K9
HSF
API
Rd
4223
LES-1

FIG.28R

```

-----...TLSGDIQTAKGASQANNSAGYVDADGNKVIYDSTDNKYYQA...
-----...TLSGDIQTAKGASQASSASYVDADGNKVIYDSTDNKYYQV...
*  ***  **  ***  ***  *  **  *  ...
...260      270      280      290      300
...GMNIRGAK--TIGG-TVDNVDFVSTYDIVEFASGANANVSVTIDN--
...Q.NGNVDFVR.Y.T...N-----A.TAH-
...Q.NGASVDFVNAY.T...N-----T.T.N...TAH-
...Q.NGASVDFVNAY.T...N-----T.T.N...TAH-
...Q.NGAS-----N...D.VN.L.T.N...TAHN
...K.V.AGSTIT-GQSE...H...L.-.DTEITTV.V.S--
...K.V.TGAT---S...R...L..SEETTL.V.S---
...V.V.PASANNQ-.E.I..A...D.V.DKDTT...VES---
...V.V.PASANNQ-.E.I..A...D.V.DKDTT...VES---
...--
...KNDGTVD.TKEVAKDKLVAQAQTPDGTILAQMNVKSVI.KEQVN.A.--
...NDKGQVD.NKEVAKDKLVAQAQTPDGTILAQMNVKSVI.KEQVN.A.--
...

```

```

310      320      330      340      350      360 ...
KKTIVRVDVTIGLFPVQYVTEDSKTVVKVGNIEYEAQDGSADMDKKV-ENGKLAKTKVKLVSA...
...G...K.D...NQ...E...
...G...D.K...E...
...G...D.K...E...
...GE...
...F...G...-K.E.V...

```

... ENGK. TE. KIGAKTS. IKEKDGKLF.T. KANK. TNKVDG. NATEDA-DE.. GLV. AKDVID.
 ... ESNGKSTK. KIGAKTSGIKEKDGKLF.T. KANKDTN. VASNNAADDT-DE.. GLV. AETVIN.
DNGK. TE. KIGAKTS. IK. HNGKLF.T. K. LKD. NNN. VTVTETDGKDE. NGLV. AKAVID.
DNGK. TE. KIGAKTS. IK. HNGKLF.T. K. LKD. NNN. VTVTETDGKDE. NGLV. AKAVID.

QGINEDNAFVKGLEKASDNKTKNAAVTVGDLNAVAQTPLTFAG-DT. TT. . KLGETILT.I.
QGINEDNAFIKGLENAAKDTKTKNAAVTVGDLNAVAQTPLTFAG-DT. TT. . KLGETILT.I.
 ... ** * ...
 ... 370 380 390 400
NGINPFVKISNVADGTEDTDAVSFKQLKALQDKQVTLAS
 ... S. T.
 ... S.Q. E.. EN. E. T..
 ... S.Q. E.. EN. E. T..
 N.
 ... E. N.
 ... VNKTGWR. KTTDANGQNG. ---FATVASGINVTF---..
 ... VNKAGWR. KTTGANNQAGQ---FETVTSGINVTF---.D
 ... VNKAGWRVKTGTGANGQND. ---FATVASGINVTF---.D
 ... VNKAGWRVKTGTGANGQND. ---FATVASGINVTF---.D

 ... KGGQTDINKLTNNIGVVAGTDGFTV. LAK. LTNLN. VN
 ... KGGQTDINKLTNNIGVVAGTDGFTV. LAK. LTNLN. VN
 *

410	420	430	440	450	460	...
-----	-----	-----	-----	-----	-----	-----

33
32
29
K22
M4071
12
11
K9
HSF
API
Rd
4223
LESS-1

FIG.28T

NAYANGSDADGGKATQTLGNDLNFKFKSTDSSELLNIKAAGDTVTFIPKKGSVQVGDDGKAT...	
.....T.N.....S.G.....S.G...K.S.T.....S.....S.....	33
.....N.....N.G.....G.....VEN.....E.....	32
.....N.....N.G.....G.....VEN.....E.....	29
.....G.I...S.G.....G.....EN.....	K22
.....V..V...S.G.....G.....DK...I.....	M4071
GNGTTATVING-TDGI TVKYDAKVGDGLKLDGD-KIAADTTALT VNDCKNANNPKGVADVA...	12
GNGTTAAVVTGDATNGITVKYEAKVGDGLKIGNDQKITADTTALTJTGGK-----VTAPD...	11
GNGTTAEVTKANDGSITVKYNKVADGLKLDGD-KIVADTTVLTVADGK-----VTAPN...	K9
GNGTTAEVTKANDGSITVKYNKVADGLKLDGD-KIVADTTVLTVADGK-----VTAPN...	HSF
.....G.....G.....S.G.....G.....EN.....	API
AGGTFKIDDKGVSF-----	
AGGTRIDEKISFVDANGQAKANTPVL SANGLDLGKRISNIGAAVDDNDNAVNFQFNEVAK...	
.....	...
... 470 480 490 500	
... IQDCAKTTTGLVEASELVDLSNKLGMKVGVGKDGTC---AT	
... SK..N..E.....E.....E.V.S.---EL	
... N.T...D.....E.....D..S.---EL	
... N.T...D.....E.....D..S.---E.	
... T.T...V.---	
... STDEKK-----T.KG..TA..S.S.TTTAAEADG.---TL	
... ATNGKK-----N..G.A.A...S.TAK-AEADTANGEL	
... NGDGKK-----F.D..G.A.A...S.TATA..E.---EV	
... NGDGKK-----F.D..G.A.A...S.TATA..E.---EV	

FIG. 28U

Rd
4223
LES-1

.....
.....
...TVNNLNQNSGASLPFVVTDANGKPIN.TDCKPQAUKGA
....
 * *

510 520 530 540 550 560 ...
DGTHTD-TLVKSGDKVTIKAGDNLKVKEGTNFTYVLRDELITGVKSVEFKDTENGASTK...
SKE-.....A.K.....A..
ASNE-.....E.....D.....A.K.....A.S..
ASNE-.....E.....D.....A.K.....A.S..
.....A.K.....D.....A..

..NASE-QE.A...F.K...A...S.Q.A...LT.ITLGTGN..K---E..
ADE-KE.A.ET.F.K...A...S.Q.A...LT.ITLGTGN..K---E..
PANSAQE.A...F...I.S.KD..S.KK..KDLT.....ANG.TGSE..
PANSAQE.A...F...I.S.KD..S.KK..KDLT.....ANG.TGSE..

..KYH-----ANGVP..
 * * * ..
... 570 580 590 600
...ITKDGLTITPAND-ANGAAATDADKIK---VASDGISAGNKAV
.....L.G....TV..
.....S.G-....
.....S.G-....

FIG.28V

.....G-.GA.G.NT.NT.S---.TK.....	M4071
-----R.SG-----	12
.....N.....G---G.NN.NT.S---.TK.....DQS.	11
.....N.....G---G.NN.NT.S---.TK.....DQS.	K9
.....G-.GA.G.NT.NT.S---.TK.....	HSF
.....G-.GA.G.NT.NT.S---.TK.....	API
.....F.....G-----R.....	Rd
-----	4223
...VD...KP..D.DKL..L..HGKPLDAGHQV...L.-GNSD-.I	LES-1
...* *** ** * * ** ** * * * *	

610	620	630	640	650	660	...
KNVVSGLKKFGDANFNPLTSSADNLTKQYDNAYKGLINLDEKSKGQTPTVADNPTAATVGDL...						

.....D.....D.....GAD...L.....						
.....D.....D.....GAD...L.....						
.....D.....						

.....D.....N.D.....GTD...V.....						
.....D.....D.....GAD...L.....						
T.....GHTLANGTV..FE-.H.....D.....GADNN-						
T.....GHTLANGTV..FE-.H.....XD.....GADNN-						

-----TLTNIKSTLP.I.TPNT.NA.AGQAQSLPSLSAAQSN..S.K.V...						

[illegible]

33
32
29
K22
M4071
12
11
K9
HSF
API
Rd
4223
LES-1

[illegible]

✱
✱
✱

✱

33
32
29
K22
M4071
12
11
K9
HSF

FIG. 28A,

... VESTDANGDKVTTTFVKTDVVELPLTQIVNTDANGNKI ---V	API
... -----	Rd
... .NEQGIRFFHVNDCNQEPVWQGRNGIDSSASGKHSVAIGFQ-	4223
... .NEQGIRFFHVNDCNQEPVWQGRNGIDSSASGKHSVAIGFQ-	LES-1
... * * * * *	

KDGQTKWYELNADGTPADMKEVTILGNVDSDEKKVKNDG-- --KWYHAKADGTADKTIKGEVD...
KNKD-KWYYTKDDGSTIDMTIKEVTILGNVSDSGKKWKEDN- - --KWYGVKSDGSTDKTQWVEE...
KKADGCKWYELNADGTASN-KEVTILGNVDANGKKVKVTENGADKWWYTINADGAADKTIKGEVS...
KKRADGCKWYELNADGTASN-KEVTILGNVDANGKKVKVTENGADKWWYTINADGAADKTIKGEVS...

AKADGEAAVAIGRQTQAAGNQSAIGDNAQATGDQSIAIGTGNVVACKHSIGAIDPSTVKADN...
AKADGEAAVAIGRQTQAAGNQSAIGDNAQATGDQSIAIGTGNVVTEKHSGAIGDPSTVKADN...
* * * * * *

33
32
29
K22
M4071

K22
M4071
12
11
K9
HSF
API
Rd
4223
LES-1

[illegible]

FIG. 28D,

...	960	970	980	990	1000	
...	GSSYQGSGLAIGVSRISDNKVIIRLSGTTNSQKKTGVAAGVGYQW*					33
...	N.	N.	N.	N.	N.	32
...	N.	N.	N.	N.	N.	29
...	N.	N.	N.	N.	N.	K22
...	N.	N.	N.	N.	N.	M4071
...	L.	L.	L.	L.	L.	12
...	N.	N.	N.	N.	N.	11
...	N.	N.	N.	N.	N.	K9
...	N.	N.	N.	N.	N.	HSF
...	N.	N.	N.	N.	N.	API
...	N.	N.	N.	N.	N.	Rd
...	IATHN..GAV.V.L.KL...	QWVFKN.SADT..HV.A.V.A.FHF*				4223
...	IATHN..GAV.V.L.KL...	QWVFKN.SADT..HV.A.V.A.FHF*				LES-1
...	*****	*****	*****	*****	*****	

FIG. 29

Oligonukleotidprimer, um das trunkierte S44-hia-Gen des Stamms 11 mit PCR zu amplifizieren.

5'	GGGAATTCATATGTCGCAACGGTGTGAGCGCAACAATACT	3'	6817.SL	SEQ ID NO:56
3'	<p>Not I</p> <p>M S44A T V E A N N N T</p> <p>Sty I</p> <p>H T I T F A L A K D L G</p> <p>CACACCATTAACCTTTCCTTTAGCGAAACCTTGGT</p> <p>GTGTGGTAATCGAAACGAAATCGCTTTCTCGAACCACCTAGGGC</p>	5'	6818.SL	SEQ ID NO:57 SEQ ID NO:58 SEQ ID NO:59

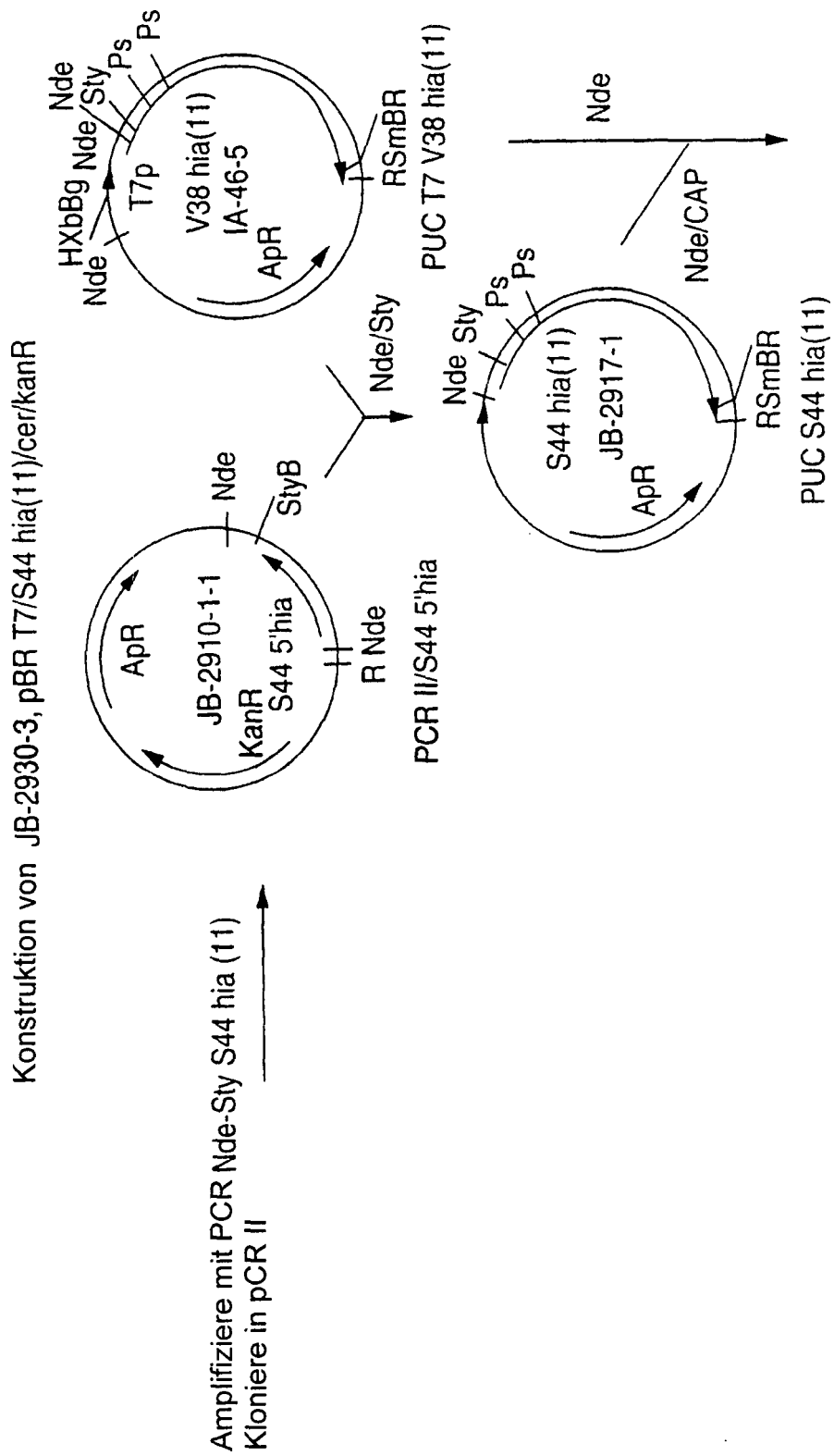


FIG.30A

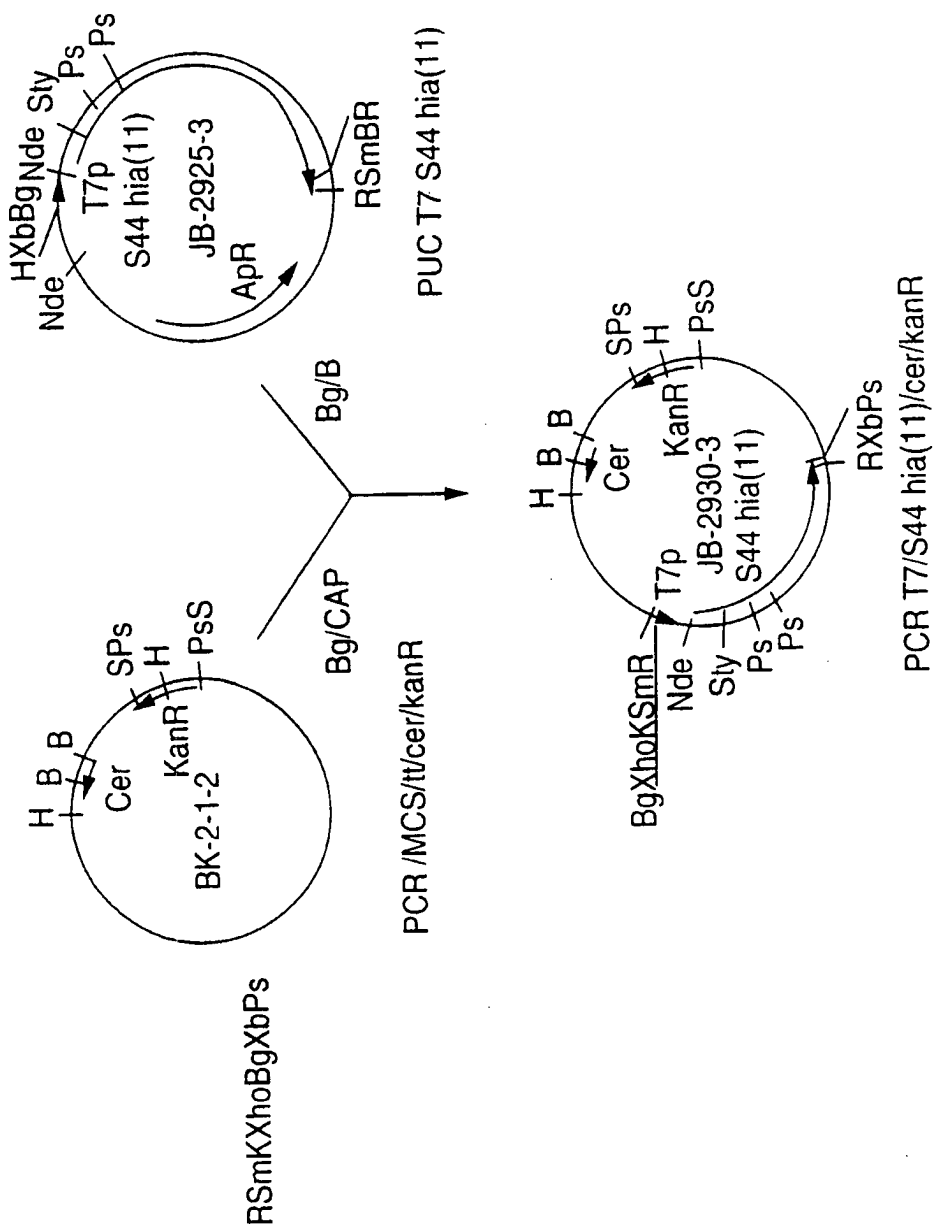


FIG.30B

Produktion von S44-rHia aus verschiedenen Vektoren

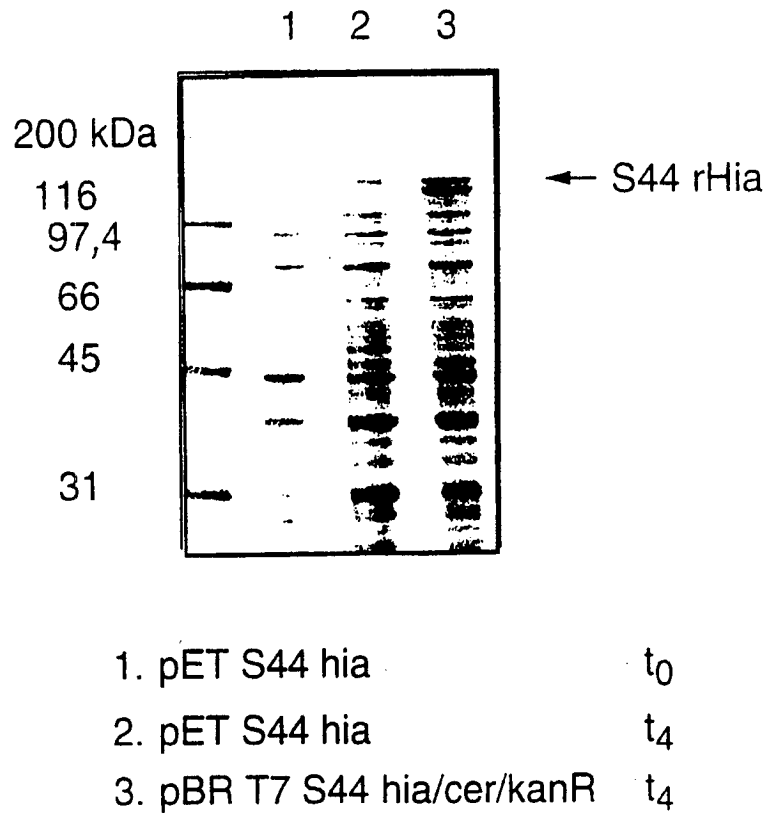


FIG.31

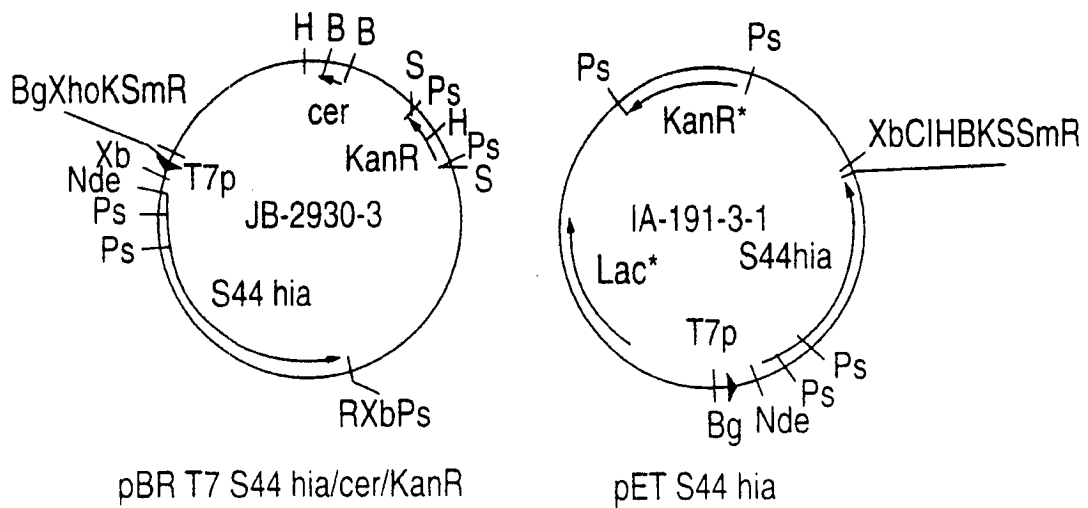


FIG.32