



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial

(21) PI 0713904-7 A2



(22) Data de Depósito: 29/06/2007
(43) Data da Publicação: 25/06/2013
(RPI 2216)

(51) Int.Cl.:
C07K 14/22

(54) Título: POLIPEPTÍDEOS A PARTIR DE NEISSERIA MENINGITIDIS

(30) Prioridade Unionista: 29/06/2006 US 60/817.950

(73) Titular(es): J.GRAIG VENTER INSTITUTE INC, NOVARTIS AG

(72) Inventor(es): Claire Fraser, Davide Serruto, Hervé Tetelin, Maria Scarselli, Mariagrazia Pizza, Rino Rappuoli

(74) Procurador(es): ORLANDO DE SOUZA

(86) Pedido Internacional: PCT IB2007002838 de 29/06/2007

(87) Publicação Internacional: WO 2008/001224de 03/01/2008

(57) Resumo: POLIPEPTÍDEOS A PARTIR DE NEISSERIA MENINGITIDIS. São reveladas várias proteínas meningocócicas específicas. A invenção provê polipeptídeos relacionados, ácidos nucléicos, anticorpos e métodos. Todos esses podem ser utilizados na medicina para tratar ou prevenir doença e/ou infecção causada por meningococo, como meningite bacteriana.

POLIPEPTÍDEOS A PARTIR DE NEISSERIA MENINGITIDIS

Todos os documentos citados aqui são incorporados a título de referência na íntegra.

Campo Técnico

5 A presente invenção refere-se ao campo de Neisseria meningitidis.

Fundamentos da Técnica

Neisseria meningitidis (meningococo) é um diplococo Gram negativo não-móvel que é patogênico em seres
10 humanos. Coloniza a faringe e causa meningite (e ocasionalmente septicemia na ausência de meningite).

Todos os meningococos patogênicos têm uma cápsula de polissacarídeo. Esses polissacarídeos formam a base de vacinas disponíveis contra sorogrupos de meningococo A,C,
15 W135 e Y, porém não são apropriados para uso contra sorogrupo B. Tem havido uma grande quantidade de pesquisa para identificar antígenos alternativos para imunizar contra sorogrupo B. Tais alternativas incluíram proteínas, o lipopolissacarídeo, e vesículas de membrana externa.

20 As referências 1 a 7 revelam vários polipeptídeos derivados da sequência de genoma de um meningococo de sorogrupo B, e selecionam sequências específicas para uso em vacinas. A sequência de genoma para uma cepa de sorogrupo A é revelada na referência 8.

25 Um objetivo da invenção é o de fornecer polipeptídeos adicionais para uso no desenvolvimento de vacinas para evitar e/ou tratar infecções meningocócicas. Em particular, é um objetivo fornecer polipeptídeos para uso em vacinas aperfeiçoadas para evitar e/ou tratar
30 meningite meningocócica. Os polipeptídeos também podem ser

úteis para fins de diagnóstico, e como alvos para antibióticos.

REVELAÇÃO DA INVENÇÃO

Polipeptídeos

5 A invenção provê polipeptídeos compreendendo as sequências de aminoácidos meningocócicos revelados nos exemplos. Essas sequências de aminoácidos são as SEQ ID NOS, pares, entre 2 e 78. Há desse modo 39 sequências de aminoácidos, e essas são mencionadas como B269_nn, onde nn
10 é um número entre 01 e 50 (há onze B269__nn números que não têm sequência: 02, 03, 04, 05, 06, 07, 08, 09, 10, 12 & 40). Duas sequências preferidas são B269_32 e B269_37.

A invenção também provê polipeptídeos compreendendo sequências de aminoácidos que têm identidade
15 de sequência para as sequências de aminoácidos meningocócicos reveladas nos exemplos. Dependendo da sequência específica, o grau de identidade de sequência é preferivelmente maior do que 50% (por exemplo, 60%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%,
20 99% ou mais). Esses polipeptídeos incluem homólogos, ortólogos, variantes alélicos e mutantes funcionais. Tipicamente, 50% ou mais de identidade entre duas sequências de polipeptídeo se considera como sendo uma indicação de equivalência funcional. Para qualquer SEQ ID
25 específica, o grau de identidade de sequência é preferivelmente maior do que os dois valores nas colunas (B) e (A) da Tabela II da presente invenção, e é mais preferivelmente maior do que todos os valores nas colunas (C), (B) e (A) para aquela SEQ ID.

30 Esses polipeptídeos podem, comparados com as

sequências meningocócicas dos exemplos, incluir uma ou mais (por exemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, etc.) substituições de aminoácidos conservativas, isto é, substituições de um aminoácido com outro que tem uma cadeia lateral relacionada. Aminoácidos geneticamente codificados são genericamente divididos em quatro famílias: (1) acídica, isto é, aspartato, glutamato; (2) básica, isto é, lisina, arginina, histidina; (3) não polar, isto é, alanina, valina, leucina, isoleucina, prolina, fenil alanina, metionina, triptofano; e (4) polar inalterado, isto é, glicina, asparagina, glutamina, cistina, serina, treonina, tirosina. Fenil alanina, triptofano e tirosina são às vezes classificadas conjuntamente como aminoácidos aromáticos. Em geral, a substituição de aminoácidos únicos nessas famílias não tem um grande efeito sobre a atividade biológica. Os polipeptídeos podem incluir também uma ou mais (por exemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, etc.) deleções de aminoácidos únicos em relação às sequências meningocócicas dos exemplos. Os polipeptídeos podem incluir também uma ou mais inserções (por exemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, etc.) (por exemplo, cada um de 1, 2, 3, 4 ou 5 aminoácidos) em relação às sequências meningocócicas dos exemplos.

A invenção provê ainda polipeptídeos compreendendo fragmentos das sequências de aminoácidos meningocócicas reveladas nos exemplos. Os fragmentos devem compreender pelo menos n aminoácidos consecutivos a partir das sequências e, dependendo da sequência específica, n é 7 ou mais (por exemplo, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100 ou mais).

O fragmento pode compreender pelo menos uma célula-T ou, preferivelmente, um epítopo de célula-B da sequência. Epítopos de célula T e B podem ser identificados empiricamente (por exemplo, utilizando PEPSCAN [9,10] ou métodos similares), ou podem ser previstos (por exemplo, utilizando o índice antigênico Jameson-Wolf [11], abordagens à base de matriz [12], TEPITOPE [13], redes neurais [14], OptiMer & EpiMer [15, 16], ADEPT [17], Tsites [18], hidrofiliçidade [19], índice antigênico [20] ou os métodos revelados na referência 21, etc.). Outros fragmentos preferidos são (a) os peptídeos de sinal de terminal-N dos polipeptídeos meningocócicos da invenção, (b) os polipeptídeos meningocócicos, porém sem seus peptídeos de sinais de terminal-N, (c) os polipeptídeos meningocócicos, porém sem seu resíduo de aminoácido terminal-N.

Polipeptídeos da invenção podem ser preparados de muitas maneiras, por exemplo, por síntese química (totalmente ou em parte), por digestão de polipeptídeos mais longos utilizando proteases, por tradução de RNA, por purificação a partir da cultura de células (por exemplo, a partir da expressão recombinante), a partir do próprio organismo (por exemplo, após cultura bacteriana, ou diretamente a partir de pacientes), etc. Um método preferido para a produção de peptídeos <40 aminoácidos de comprimento envolve síntese química in vitro [22, 23]. Síntese de peptídeo de fase sólida é particularmente preferida, como métodos baseados na química tBoc ou Fmoc [24]. A síntese enzimática [25] pode ser também utilizada em parte ou totalmente. Como alternativa para a síntese

química, a síntese biológica pode ser utilizada, por exemplo, os polipeptídeos podem ser produzidos por tradução. Isso pode ser realizado in vitro ou in vivo. Métodos biológicos são em geral restritos à produção de polipeptídeos baseados em L-aminoácidos, porém a manipulação de maquinaria de tradução (por exemplo, de moléculas de tRNA de aminoacila) pode ser utilizada para permitir a introdução de D-aminoácidos (ou de outros aminoácidos não naturais, como iodotirosina ou metil fenil alanina, azido homoalanina, etc.) [26]. Onde D-aminoácidos são incluídos, entretanto, prefere-se utilizar síntese química. Polipeptídeos da invenção podem ter modificações covalentes na extremidade-C e/ou extremidade-N.

Polipeptídeos da invenção podem assumir várias formas (por exemplo, nativa, fusões, glicosilada, não glicosilada, lipidada, não lipidada, fosforilada, não fosforilada, miristoilada, não miristoilada, monomérica, multimérica, em partículas, desnaturada, etc.).

Polipeptídeos da invenção são preferivelmente fornecidos em forma purificada ou substancialmente purificada, isto é, substancialmente livre de outros polipeptídeos (por exemplo, livre de polipeptídeos de ocorrência natural), particularmente a partir de outros polipeptídeos meningocócicos ou de células hospedeiras, e são genericamente pelo menos aproximadamente 50% puro (em peso), e normalmente pelo menos aproximadamente 90% puro, isto é menos de aproximadamente 50% e mais preferivelmente menos de aproximadamente 10% (por exemplo, 5%) de uma composição é composta de outros polipeptídeos expressos. Polipeptídeos da invenção são preferivelmente polipeptídeos

meningocócicos. Polipeptídeos da invenção têm, preferivelmente, a função indicada na Tabela 1 para a sequência relevante.

5 Polipeptídeos da invenção podem ser fixados em um suporte sólido. Polipeptídeos da invenção podem compreender um rótulo detectável (por exemplo, um rótulo radioativo ou fluorescente, ou um rótulo de biotina).

O termo "polipeptídeo" se refere a polímeros de aminoácido de qualquer comprimento. O polímero pode ser
10 linear ou ramificado, pode compreender aminoácidos modificados, e pode ser interrompido por não-aminoácidos. Os termos também abrangem um polímero de aminoácido que foi modificado naturalmente ou por intervenção; por exemplo, formação de ligação de dissulfeto, glicosilação, lipidação,
15 acetilação, fosforilação ou qualquer outra manipulação ou modificação, como conjugação com um componente de rotulação. São também incluídos na definição, por exemplo, polipeptídeos contendo um ou mais análogos de um aminoácido (incluindo, por exemplo, aminoácidos não naturais, etc.),
20 bem como outras modificações conhecidas na técnica. Polipeptídeos podem ocorrer como cadeias únicas ou cadeias associadas. Polipeptídeos da invenção podem ser glicosilados naturalmente ou não naturalmente (isto é, o polipeptídeo tem um padrão de glicosilação que difere do
25 padrão de glicosilação encontrado no polipeptídeo de ocorrência natural, correspondente).

A invenção provê polipeptídeos compreendendo uma sequência -X-Y- ou -Y-X-, onde: -X- é uma sequência de aminoácidos como definido acima e -Y- não é uma sequência
30 como definido acima, isto é, a invenção provê proteínas de

5 fusão. Onde o códon de extremidade-N de uma sequência de codificação de polipeptídeo não é ATG então aquele códon será traduzido como o aminoácido padrão para aquele códon em vez de como Met, que ocorre quando o códon é um códon de partida.

A invenção provê um processo para produzir polipeptídeos da invenção, compreendendo a etapa de cultivar uma célula hospedeira da invenção sob condições que induzem expressão de polipeptídeo.

10 A invenção provê um processo para a produção de um polipeptídeo da invenção, onde o polipeptídeo é sintetizado em parte ou totalmente utilizando meio químico.

A invenção provê uma composição compreendendo dois ou mais polipeptídeos da invenção.

15 A invenção também provê um polipeptídeo híbrido representado pela fórmula $\text{NH}_2\text{-A-}[-\text{X-L-}]_n\text{-B-COOH}$, onde X é um polipeptídeo da invenção como definido acima, L é uma sequência de aminoácido ligador opcional, A é uma sequência de aminoácido terminal-N opcional, B é uma sequência de aminoácido terminal-C opcional, e n é um número inteiro maior do que 1. O valor de n está entre 2 e x, e o valor de x é tipicamente 3, 4, 5, 6, 7, 8, 9 ou 10. Preferivelmente n é 2; 3 ou 4; mais preferivelmente 2 ou 3; mais preferivelmente, n = 2. Para cada caso de n, -X- pode ser
20 igual ou diferente. Para cada caso de n de $[-\text{X-L-}]$, a sequência de aminoácido ligador -L- pode estar presente ou ausente. Por exemplo, quando n=2 o híbrido pode ser $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-L}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-L}_1\text{-X}_2\text{-COOH}$, $\text{NH}_2\text{-X}_1\text{-X}_2\text{-L}_2\text{COOH}$, etc. Sequência(s) de aminoácido ligador -L- será,
25 tipicamente curta (por exemplo, 20 ou menos aminoácidos,
30

isto é, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Os exemplos incluem sequências de peptídeo curtas que facilitam clonagem, ligadores de poli-glicina (isto é, Gly_n onde n = 2, 3, 4, 5, 6, 7, 8, 9, 10 ou mais), e marcadores de histidina (isto é, His_n onde n = 3, 4, 5, 6, 7, 8, 9, 10 ou mais). Outras sequências de aminoácido ligador apropriadas serão evidentes para aqueles versados na técnica. -A- e -B- são sequências opcionais que serão tipicamente curtas (por exemplo, 40 ou menos aminoácidos, isto é, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 2, 1). Os exemplos incluem sequências líderes para orientar tráfego de polipeptídeos, ou sequências de peptídeo curtas que facilitam clonagem ou purificação (por exemplo, marcadores de histidina, isto é, His_n, onde n = 3, 4, 5, 6, 7, 8, 9, 10 ou mais). Outras sequências de aminoácido terminal-N e terminal-C apropriadas serão evidentes para aqueles versados na técnica.

Vários testes podem ser utilizados para avaliar a imunogenicidade in vivo de polipeptídeos da invenção. Por exemplo, polipeptídeos podem ser expressos de forma recombinante e utilizados para classificar soros de pacientes por imunoblot. Uma reação positiva entre o polipeptídeo e soro de paciente indica que o paciente montou previamente uma resposta imune à proteína em questão, isto é, a proteína é um imunógeno. Esse método também pode ser utilizado para identificar proteínas imunodominantes.

A invenção provê anticorpos que se ligam a polipeptídeos da invenção. Esses podem ser policlonais ou monoclonais e podem ser produzidos por qualquer meio apropriado (por exemplo, por expressão recombinante). Para
5 aumentar a compatibilidade com o sistema imune humano, os anticorpos podem ser quiméricos ou humanizados (por exemplo, refs. 27 & 28), ou anticorpos totalmente humanos podem ser utilizados. Os anticorpos podem incluir um rótulo detectável (por exemplo, para ensaios de diagnóstico).
10 Anticorpos da invenção podem ser fixados em um suporte sólido. Anticorpos da invenção são preferivelmente anticorpos de neutralização.

Anticorpos monoclonais são particularmente úteis em identificação e purificação dos polipeptídeos
15 individuais contra os quais são dirigidos. Anticorpos monoclonais da invenção também podem ser empregados como reagentes em imunoenaios, radioimunoenaios (RIA) ou ensaios imunosorventes ligados por enzima (ELISA), etc. Nessas aplicações, os anticorpos podem ser rotulados com um
20 reagente analiticamente detectável como radioisótopo, uma molécula fluorescente ou uma enzima. Os anticorpos monoclonais produzidos pelo método acima podem ser também utilizados para a identificação e caracterização molecular (mapeamento de epítomos) de polipeptídeos da invenção.

25 Os anticorpos da invenção são preferivelmente fornecidos em forma purificada ou substancialmente purificada. Tipicamente, o anticorpo estará presente em uma composição que é substancialmente livre de outros polipeptídeos, por exemplo, onde menos de 90% (em peso),
30 normalmente menos de 60% e mais normalmente menos de 50% da

composição é composta de outros polipeptídeos.

Os anticorpos da invenção podem ser de qualquer isotipo (por exemplo, IgA, IgG, IgM, isto é, mas serão genericamente IgG. No isotipo IgG, anticorpos podem ser
5 subclasse IgG1, IgG2, IgG3 ou IgG4. Os anticorpos da invenção

Os anticorpos da invenção podem assumir várias formas, incluindo anticorpos inteiros, fragmentos de anticorpo como fragmentos $F(ab')_2$ e $F(ab)$, fragmentos Fv
10 (heterodímeros não covalentes), anticorpos de cadeia única como moléculas Fv de cadeia única (scFv), minicorpos, oligocorpos, etc. O termo "anticorpo" não indica nenhuma origem específica, e inclui anticorpos obtidos através de processos não convencionais, como exibição de fagos.

15 A invenção provê um processo para detectar polipeptídeos da invenção, compreendendo as etapas de: (a) contatar um anticorpo da invenção com uma amostra biológica sob condições apropriadas para a formação de um complexo de antígeno-anticorpo; e (b) detectar os complexos.

20 A invenção provê um processo para detectar anticorpos da invenção, compreendendo as etapas de: (a) contatar um polipeptídeo da invenção com uma amostra biológica (por exemplo, uma amostra de sangue ou soro) sob condições apropriadas para a formação de um complexo de
25 anticorpo-antígeno; e (b) detectar os complexos.

Ácidos nucleicos

A invenção provê ácido nucleico compreendendo as sequências de nucleotídeos meningocócicas reveladas nos exemplos. Essas sequências de ácido nucleico são as SEQ ID
30 Nos ímpares entre 1 e 77.

A invenção também provê ácido nucléico compreendendo sequências de nucleotídeos tendo identidade de sequência com as sequências de nucleotídeos meningocócicas reveladas nos exemplos.

5 A invenção também provê ácido nucléico que pode hibridizar ao ácido nucléico meningocócico revelado nos exemplos. Reações de hibridização podem ser executadas sob condições de diferente "severidade". Condições que aumentam severidade de uma reação de hibridização amplamente conhecida e publicada na técnica [por exemplo, página 7.52
10 de referência 29]. Os exemplos de condições relevantes incluem (em ordem de severidade crescente): temperaturas de incubação de 25°C, 37°C, 50°C, 55°C e 68°C; concentrações de tampão de 10 x SSC, 6 x SSC, 1 x SSC, 0,1 x SSC (onde
15 SSC é 0.15 M NaCl e 15 mM de tampão de citrato) e seus equivalentes utilizando outros sistemas de tampão; concentrações de formamida de 0%, 25%, 50% e 75%; tempos de incubação de 5 minutos a 24 horas; 1, 2 ou mais etapas de lavagem; tempos de incubação de lavagem de 1, 2 ou 15
20 minutos, e soluções de lavagem de 6 x SSC, 1 X SSC, 0,1 X SSC ou água deionizada. Técnicas de hibridização e sua otimização são bem conhecidas na arte [por exemplo, vide as referências 29-32, etc.].

Em algumas modalidades, o ácido nucléico da
25 invenção hibridiza com um alvo da invenção sob condições de baixa severidade; em outras modalidades hibridiza sob condições de severidade intermediária; em modalidades preferidas, hibridiza sob condições de severidade elevada. Um conjunto exemplar de condições de hibridização de baixa
30 severidade é 50°C e 10 x SSC. Um conjunto exemplar de

condições de hibridização de severidade intermediária é 55°C e 1 x SSC. Um conjunto exemplar de condições de hibridização de severidade elevada é 68°C e 0,1 x SSC.

Fragmentos compreendendo ácido nucléico dessas seqüências são também fornecidos. Esses devem compreender pelo menos n nucleotídeos consecutivos a partir das seqüências meningocócicas e, dependendo da seqüência específica, n é 10 ou mais (por exemplo, 12, 14, 15, 18, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200 ou mais).

A invenção provê ácido nucléico da fórmula 5'-X-Y-Z-3', onde: -X- é uma seqüência de nucleotídeos consistindo em x nucleotídeos; -Z- é uma seqüência de nucleotídeos que consiste em z nucleotídeos; -Y- é uma seqüência de nucleotídeos consistindo em (a) um fragmento de uma das SEQ ID NOS: 1 a 77, de número ímpar, ou (b) o complemento de (a); e o ácido nucléico 5'-X-Y-Z-3' não é (i) um fragmento de uma das SEQ ID NOS: 1 a 77 de número ímpar nem (ii) o complemento de (i). As frações -X- e/ou -Z- podem compreender uma seqüência promotora (ou seu complemento).

A invenção também provê ácido nucléico codificando os polipeptídeo e fragmentos de polipeptídeo da invenção.

A invenção inclui ácido nucléico compreendendo seqüências complementares às seqüências reveladas na listagem de seqüência (por exemplo, para anti-sentido ou sondagem, ou para uso como iniciadores), bem como as seqüências na orientação efetivamente mostrada.

Ácidos nucléicos da invenção podem ser utilizados

em reações de hibridização (por exemplo, Northern ou Southern blots, ou em microconjuntos de ácido nucléico ou 'chips de gene') e reações de amplificação (por exemplo, PCR, SDA, SSSR, LCR, TMA, NASBA, etc.) e outras técnicas de
5 ácido nucléico.

Ácido nucléico, de acordo com a invenção pode assumir várias formas (por exemplo, de fita única, de fita dupla, vetores, iniciadores, sondas, rotulado, etc.). Ácidos nucléicos da invenção podem ser circulares ou
10 ramificados, porém serão genericamente lineares. A menos que de outro modo especificado ou exigido, qualquer modalidade da invenção que utiliza um ácido nucléico pode utilizar a forma de fita dupla e cada uma de duas formas de fita única, complementares, que compõem a forma de fita
15 dupla. Iniciadores e sondas são genericamente de fita única, como são ácidos nucléicos de anti-sentido.

Ácidos nucléicos da invenção são preferivelmente fornecidos em forma purificada ou substancialmente purificada, isto é, substancialmente livre de outros ácidos
20 nucléicos (por exemplo, livre de ácidos nucléicos de ocorrência natural), particularmente de outros Haemophilus ou ácidos nucléicos de célula hospedeira, genericamente sendo pelo menos aproximadamente 50% puro (em peso), e normalmente pelo menos aproximadamente 90% puro. Ácidos
25 nucléicos da invenção são preferivelmente ácidos nucléicos H. influenzae.

Ácidos nucléicos da invenção podem ser preparados de muitas maneiras, por exemplo, por síntese química (por exemplo, síntese de fosforamidite de DNA) integralmente ou
30 em parte, por digestão de ácidos nucléicos mais longos

utilizando nucleases (por exemplo, enzimas de restrição), por junção de ácidos nucléicos mais curtos ou nucleotídeos (por exemplo, utilizando ligases ou polimerases), de bibliotecas de cDNA ou genômicas, etc.

5 Ácido nucléico da invenção pode ser fixado em um suporte sólido (por exemplo, uma conta, placa, filtro, filme, lâmina, suporte de microconjuntos, resina, etc.). Ácido nucléico da invenção pode ser rotulado, por exemplo, com um rótulo radioativo ou fluorescente, ou um rótulo de
10 biotina. Isso é particularmente útil onde o ácido nucléico deve ser utilizado em técnicas de detecção, por exemplo, onde o ácido nucléico é um iniciador ou como uma sonda.

 O termo "ácido nucléico" em geral significa uma forma polimérica de qualquer comprimento, que contém
15 desoxirribonucleotídeos, ribonucleotídeos, e/ou seus análogos. Incluem-se DNA, RNA, híbridos de DNA/RNA. Também inclui análogos de DNA ou RNA, como aqueles contendo colunas modificadas (por exemplo, ácidos nucléicos de peptídeo (PNAs) ou fosforotioatos) ou bases modificadas.
20 Desse modo, a invenção inclui mRNA, tRNA, rRNA, ribozimas, DNA, cDNA, ácidos nucléicos recombinantes, ácidos nucléicos ramificados, plasmídeos, vetores, sondas, iniciadores, etc. Onde ácido nucléico da invenção tem a forma de RNA, pode ou não ter um cap 5'.

25 Ácidos nucléicos da invenção compreendem sequências meningocócicas, como definido acima, porém podem compreender também sequências não meningocócicas (por exemplo, em ácidos nucléicos da fórmula 5'-X-Y-Z-3', como definido acima). Isso é particularmente útil para
30 iniciadores, que podem desse modo compreender uma primeira

sequência complementar a um ácido nucléico PCAV alvo e uma segunda sequência que não é complementar ao ácido nucléico alvo. Quaisquer tais sequências não complementares no iniciador são preferivelmente 5' para as sequências complementares. Sequências não complementares típicas compreendem sítios de restrição ou sequências promotoras.

Ácidos nucléicos da invenção podem ser preparados de muitas maneiras, por exemplo, por síntese química (pelo menos em parte), por digestão de ácidos nucléicos mais longos utilizando nucleases (por exemplo, enzimas de restrição), por junção de ácidos nucléicos mais curtos (por exemplo, utilizando ligases ou polimerases), de bibliotecas de cDNA ou genômicas, etc.

Ácidos nucléicos da invenção podem fazer parte de um vetor, isto é, parte de uma construção de ácido nucléico projetada para transdução/transfecção de um ou mais tipos de células. Vetores podem ser, por exemplo, "vetores de clonagem" que são projetados para isolamento, propagação e replicação de nucleotídeos inseridos, "vetores de expressão" que são projetados para expressão de uma sequência de nucleotídeos em uma célula hospedeira, "vetores virais" que são projetados para resultar na produção de um vírus recombinante ou partícula semelhante a vírus, ou "vetores shuttle", que compreendem os atributos de mais de um tipo de vetor. Vetores preferidos são plasmídeos. Uma "célula hospedeira" inclui uma célula individual ou cultura de células que pode ser ou foi um receptor de ácido nucléico exógeno. Células hospedeiras incluem progênie de uma célula hospedeira única, e a progênie pode não necessariamente ser completamente

idêntica (em morfologia ou em complemento de DNA total) à célula pai de origem devido à mutação e/ou alteração natural, acidental ou deliberada. Células hospedeiras incluem células transfectadas ou infectadas in vivo ou in vitro com ácido nucléico da invenção.

Onde um ácido nucléico é DNA, será reconhecido que "U" em uma sequência de RNA será substituído por "T" no DNA. Similarmente, onde um ácido nucléico é RNA, será reconhecido que "T" em uma sequência de DNA será substituída por "U" no RNA.

O termo "complemento" ou "complementar", quando utilizado em relação a ácidos nucléicos se refere a emparelhamento de base Watson-Crick. Desse modo, o complemento de C é G, o complemento de G é C, o complemento de A é T (ou U), e o complemento de T (ou U) é A. Também é possível utilizar bases como I (inosina purina), por exemplo, para complementar as pirimidinas (C ou T). Os termos também indicam uma direção - o complemento de 5'-ACAGT-3' é 5'-ACTGT-3' em vez de 5'-TGTCA-3'.

Ácidos nucléicos da invenção podem ser utilizados, por exemplo: para produzir polipeptídeos; como sondas de hibridização para a detecção de ácido nucléico em amostras biológicas; para gerar cópias adicionais dos ácidos nucléicos; para gerar ribozimas ou oligonucleotídeos de anti-sentido; como iniciadores ou sondas de DNA de fita única; ou como oligonucleotídeos de formação de fita tripla.

A invenção provê um processo para produzir ácido nucléico da invenção, onde o ácido nucléico é sintetizado em parte ou totalmente utilizando meio químico.

A invenção provê vetores compreendendo sequências de nucleotídeo da invenção (por exemplo, vetores de expressão ou clonagem) e células hospedeiras transformadas com tais vetores.

5 A invenção também provê um kit que compreende iniciadores (por exemplo, iniciadores de PCR) para amplificar uma sequência padrão, contida em uma sequência de ácido nucléico de meningococo, o kit compreendendo um primeiro iniciador e um segundo iniciador, onde o primeiro
10 iniciador é substancialmente complementar à sequência padrão e o segundo iniciador é substancialmente complementar a um complemento da sequência gabarito, onde as partes dos iniciadores que têm complementaridade substancial definem as extremidades da sequência gabarito
15 como sendo amplificadas. O primeiro iniciador e/ou segundo iniciador pode incluir um rótulo detectável (por exemplo, um rótulo fluorescente).

A invenção também provê um kit que compreende primeiro e segundo oligonucleotídeos de fita única que
20 permitem amplificação de uma sequência de ácido nucléico meningocócico padrão, contida em um ácido nucléico de fita única ou dupla (ou mistura dos mesmos), onde: (a) o primeiro oligonucleotídeo compreende uma sequência iniciadora que é substancialmente complementar à sequência
25 de ácido nucléico gabarito; (b) o segundo oligonucleotídeo compreende uma sequência iniciadora que é substancialmente complementar ao complemento da sequência de ácido nucléico gabarito; (c) o primeiro oligonucleotídeo e/ou o segundo oligonucleotídeo compreende(m) sequência que não é
30 complementar ao ácido nucléico gabarito; e (d) as

sequências iniciadoras definem as extremidades da sequência gabarito como sendo amplificadas. A(s) sequência(s) não complementar(es) de característica (c) são preferivelmente a montante de (isto é, 5' para) sequências iniciadoras.

5 Uma ou ambas essas sequências (c) podem compreender um sítio de restrição [por exemplo, ref.33] ou uma sequência promotora [por exemplo, 34]. O primeiro oligonucleotídeo e/ou o segundo oligonucleotídeo pode incluir um rótulo detectável (por exemplo, um rótulo fluorescente).

10 A sequência gabarito pode ser qualquer parte de uma sequência de genoma.

A invenção provê um processo para detectar ácido nucléico da invenção, compreendendo as etapas de: (a) contatar uma sonda nucléica, de acordo com a invenção, com
15 uma amostra biológica sob condições de hibridização para formar duplex; e (b) detectar o duplex.

A invenção provê um processo para detectar meningococo em uma amostra biológica (por exemplo, sangue), compreendendo a etapa de contatar ácido nucléico, de acordo
20 com a invenção, com a amostra biológica sob condições de hibridização. O processo pode envolver amplificação de ácido nucléico (por exemplo, PCR, SDA, SSSR, LCR, TMA, NASBA, etc.) ou hibridização (por exemplo, microconjuntos, blots, hibridização com uma sonda em solução, etc.).

25 A invenção provê um processo para preparar um fragmento de uma sequência alvo, onde o fragmento é preparado por extensão de um iniciador de ácido nucléico. A sequência alvo e/ou o iniciador são ácidos nucléicos da invenção. A reação de extensão de iniciador pode envolver
30 amplificação de ácido nucléico (por exemplo, PCR, SDA,

SSSR, LCR, TMA, NASBA, etc.)

Amplificação de ácido nucléico, de acordo com a invenção, pode ser quantitativa e/ou tempo real.

Para certas modalidades da invenção, ácidos nucléicos são preferivelmente pelo menos 7 nucleotídeos de comprimento (por exemplo, 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, 45, 50, 55, 60, 65, 70, 75, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 224, 250, 275, 300 nucleotídeos ou mais longa).

Para certas modalidades da invenção, ácidos nucléicos são preferivelmente no máximo 500 nucleotídeos em comprimento (por exemplo, 450, 400, 350, 300, 250, 200, 150, 140, 130, 120, 110, 100, 90, 80, 75, 70, 65, 60, 55, 50, 45, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15 nucleotídeos ou mais curto).

Iniciadores e sondas da invenção e outros ácidos nucléicos utilizados para hibridização são preferivelmente entre 10 e 30 nucleotídeos em comprimento (por exemplo, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 ou 30 nucleotídeos).

Composições farmacêuticas

A invenção provê composições compreendendo: (a) polipeptídeo, anticorpo, e/ou ácido nucléico da invenção; e (b) um veículo farmacêuticamente aceitável. Essas composições podem ser apropriadas como composições imunogênicas, por exemplo, ou como reagentes de diagnóstico, ou como vacinas. Vacinas de acordo com a invenção podem ser profiláticas (isto é, para evitar

infecção) ou terapêuticas (isto é, para tratar infecção), porém serão tipicamente profiláticas.

Um "veículo farmacêuticamente aceitável" inclui qualquer veículo que não induz ele próprio a produção de anticorpos prejudiciais ao indivíduo que recebe a composição. Veículos apropriados são tipicamente macromoléculas grandes, lentamente metabolizadas como proteínas, polissacarídeos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácido, sacarose, trealose, lactose e agregados de lipídeo (como lipossomas ou gotículas de óleo). Tais veículos são bem conhecidos por aqueles com conhecimentos comuns na técnica. As vacinas também podem conter diluentes, como água, solução salina, glicerol, etc.

Adicionalmente, substâncias auxiliares, como agentes umectantes ou de emulsificação, substâncias de tamponamento de pH, e similares, podem estar presentes. Solução salina fisiológica tamponada como fosfato, isenta de pirogênio, estéril, é um veículo típico. Uma discussão completa de excipientes farmacêuticamente aceitáveis é disponível na ref. 141.

Composições da invenção podem incluir um antimicrobiano, particularmente se acondicionado em um formato de dose múltipla.

Composições da invenção podem compreender detergente, por exemplo, um Tween (polisorbato) como Tween 80. Detergentes são genericamente presentes em baixos níveis, por exemplo, <0,01%.

Composições da invenção podem incluir sais de sódio (por exemplo, cloreto de sódio) para fornecer

tonicidade. Uma concentração de 10 ± 2 mg/ml NaCl é típica.

Composições da invenção incluirão, genericamente, um tampão. Um tampão de fosfato é típico.

Composições da invenção podem compreender um
5 álcool de açúcar (por exemplo, manitol) ou um dissacarídeo
(por exemplo, sacarose ou trealose), por exemplo, em torno
de 15-30 mg/mol (por exemplo, 25 mg/ml), particularmente se
devem ser liofilizados ou se incluem material que foi
reconstituído a partir de material liofilizado. O pH de uma
10 composição para liofilização pode ser ajustado em torno de
6,1 antes da liofilização.

Polipeptídeos da invenção podem ser administrados
em combinação com outros agentes imunorreguladores. Em
particular, composições incluirão normalmente um adjuvante
15 de vacina. Adjuvantes que podem ser utilizados em
composição da invenção incluem, porém não são limitados a:

A. composições contendo mineral

Composições contendo mineral, apropriadas, para
uso como adjuvantes na invenção incluem sais minerais, como
20 sais de alumínio e sais de cálcio. A invenção inclui sais
minerais como hidróxidos (por exemplo, oxiidróxidos),
fosfatos (por exemplo, hidróxi fosfatos, ortofosfatos),
sulfatos, etc. [por exemplo, vide os capítulos 8 & 9 da
ref. 35] ou misturas de compostos minerais diferentes, com
25 os compostos assumindo qualquer forma apropriada (por
exemplo, gel, cristalina, amorfa, etc.), e com adsorção
sendo preferida. As composições contendo mineral também
podem ser formuladas como uma partícula de sal de metal
[36].

30 Fosfatos de alumínio são particularmente

preferidos, particularmente em composições que incluem um antígeno de sacarídeo H. influenzae, e um adjuvante típico é hidróxi fosfato de alumínio amorfo com razão molar de PO_4/Al entre 0,84 e 0,91, incluído em 0,6m Al^{3+}/ml . A adsorção com uma dose baixa de fosfato de alumínio pode ser utilizada, por exemplo, entre 50 e 100 $\mu g Al^{3+}$ por conjugado por dose. Onde há mais de um conjugado em uma composição, nem todos os conjugados necessitam ser adsorvidos.

10 **B. Emulsões de óleo**

Composições de emulsão de óleo apropriadas para uso como adjuvantes na invenção incluem emulsões de esqualeno-água, como MF59 [capítulo 10 da ref. 35; vide, também ref. 37] (5% Esqualeno, 0,5% Tween 80, e 0,5% Span 85, formulado em partículas de submícron utilizando um microfluidificador). Adjuvante de Freund completo (CFA) e o adjuvante de Freund incompleto (IFA) também podem ser utilizados.

20 **C. Formulações de saponina [capítulo 22 da ref. 35]**

Formulações de saponina também podem ser utilizadas como adjuvantes na invenção. Saponinas fazem parte de um grupo heterólogo de glicosídeos de esterol e glicosídeos de triterpenóide que são encontrados na casca, folhas, troncos, raízes e mesmo flores de uma ampla gama de espécies de plantas. Saponina a partir da casca da árvore Molina Quillaia saponaria tem sido amplamente estudada como adjuvantes. Saponina também pode ser comercialmente obtida de Smilax ornata (sarsapilla), Gypsophilia paniculata (véu de noiva), e Saponaria officianalis (soap root).

Formulações adjuvantes de saponina incluem formulações purificadas, como QS21, bem como formulações de lipídeo, como ISCOMs. QS21 é comercializado como Stimulon™.

Composições de saponina foram purificadas utilizando HPLC e RP-HPLC. Frações purificadas específicas utilizando essas técnicas foram identificadas, incluindo QS7, QS17, QS18, QS21, QH-A, QH-B e QH-C. Preferivelmente, a saponina é QS21. Um método de produção de QS21 é revelado na ref. 38. Formulações de saponina também podem compreender um esterol, como colesterol [39].

Combinações de saponinas e colesterolis podem ser utilizadas para formar partículas únicas, denominadas complexos de imunoestimulação (ISCOMs) [capítulo 23 da ref. 35]. ISCOMs também incluem, tipicamente, um fosfolipídeo como fosfatidil etanol amina ou fosfatidil colina. Qualquer saponina conhecida pode ser utilizada em ISCOMs. Preferivelmente, o ISCOM inclui uma ou mais de Quila, QHA & QHC. ISCOMs são adicionalmente descritos nas refs. 39-41. Opcionalmente, os ISCOMs podem ser isentos de detergente adicional [42].

Um exame do desenvolvimento de adjuvantes baseados em saponina pode ser encontrado nas refs. 43 & 44.

D. Virosomas e partículas semelhantes a vírus

Virosomas e partículas semelhantes a vírus (VLPs) também podem ser utilizados como adjuvantes na invenção. Essas estruturas contêm, genericamente uma ou mais proteínas a partir de um vírus opcionalmente combinado ou formulado com um fosfolipídeo. São genericamente não patogênicos, não replicantes, e genericamente não contêm nenhum do genoma viral nativo. As proteínas virais podem

ser recombinantemente produzidas ou isoladas a partir de vírus inteiros. Essas proteínas virais apropriadas para uso em Virossomas ou VLPs incluem proteínas derivadas de vírus de influenza (como HA ou NA), vírus de hepatite B (como proteínas de cápside ou núcleo), vírus de Hepatite E, vírus de sarampo, vírus Sindbis, Rotavírus, vírus de doença de febre aftosa, Retrovírus, vírus Norwalk, vírus de Papiloma humano, HIV, RNA-fagos, Q β -fago (como proteínas de revestimento), GA-fago, fr-fago, AP205 fago, e Ty (como proteína p1 Ty retrotransposon). VLPs são discutidos adicionalmente nas refs. 45-50. Virossomas são discutidos adicionalmente, por exemplo, na ref. 51.

E. derivados bacterianos ou microbianos

Adjuvantes apropriados para uso na invenção incluem derivados bacterianos ou microbianos como derivados não tóxicos de lipopolissacarídeo enterobacteriano (LPS), derivados de Lipídeo A, oligonucleotídeos imunostimuladores e toxinas de ADP-ribosilação e derivados detoxificados dos mesmos.

Derivados não tóxicos de LPS incluem lipídeo monofosforila A (MPL) e MPL 3-O-desacilado (3DMPL). 3DMPL é uma mistura de lipídeo monofosforila 3 de-O-acilado A com 4, 5 ou 6 cadeias aciladas. Uma forma de "partícula pequena" preferida de lipídeo monofosforila 3 De-O-acilado A, é revelada na ref. 52. Tais "partículas pequenas" de 3DMPL são pequenas o bastante para serem filtradas estéreis através de uma membrana de 0,22 μ m [52]. Outros derivados de LPS não tóxicos incluem simulações de lipídeo monofosforila A, como derivados de fosfato de glucosaminida aminoalquila, por exemplo, RC-529 [53, 54].

Derivados de lipídeo A incluem derivados de lipídeo A de *Escherichia coli* como OM-174. OM-174 é descrito, pro exemplo, nas refs. 55 & 56.

5 Oligonucleotídeos imunoestimuladores apropriados para uso como adjuvantes na invenção incluem sequências de nucleotídeos contendo um motivo CpG (uma sequência de dinucleotídeo contendo uma citosina não metilada ligada por uma ligação de fosfato a uma guanosina). RNAs de fita dupla e oligonucleotídeos contendo sequências palindrômicas ou
10 poli(dG) também foram mostradas como sendo imunoestimuladoras.

Os CpG's podem incluir modificações/análogos de nucleotídeo como modificações de fosforotioato e podem ser de fita dupla ou de fita única. As referências 57, 58 e 59
15 revelam possíveis substituições de análogo, por exemplo, substituição de guanosina com 2'-desóxi-7-deazaguanosina. O efeito adjuvante de oligonucleotídeos de CpG é adicionalmente discutido nas refs. 60-65.

A sequência de CpG pode ser dirigida a TLR9, como
20 o motivo GTCGTT ou TTCGTT [66]. A seqüência CpG pode ser específica para induzir uma resposta imune Th1, como um CpG-A ODN, ou pode ser mais específica para induzir uma resposta de célula B, como CpG-B ODN. CpG-A e CpG-B ODNs são discutidos nas refs. 67-69. Preferivelmente, o CpG é um
25 CpG-A ODN.

Preferivelmente, o oligonucleotídeo CpG é construído de modo que a extremidade 5' seja acessível para reconhecimento do receptor. Opcionalmente, duas sequências de oligonucleotídeo CgP podem ser fixadas em suas
30 extremidades 3' para formar "imunômeros". Vide, por

exemplo, refs. 66 & 70-72.

Toxinas de ADP-ribosilação bacterianas e derivados detoxificados das mesmas podem ser utilizados como adjuvantes na invenção. Preferivelmente, a proteína é derivada de E.coli (enterotoxina instável em calor E.coli "LT"), cólera ("CT"), ou coqueluche ("PT"). O uso de toxinas de ADP-ribosilação detoxificadas como adjuvantes de mucosa é descrito na ref. 73 e como adjuvantes parenterais na ref. 74. A toxina ou toxóide está preferivelmente na forma de uma holotoxina, compreendendo subunidades tanto A como B. Preferivelmente, a subunidade A contém uma mutação de detoxificação; preferivelmente a subunidade B não é mudada. Preferivelmente, o adjuvante é um mutante LT detoxificado como LT-K63, LT-R72 e LT-G192. O uso de toxinas ADP-ribosilação e derivados detoxificados dos mesmos, particularmente LT-K63 e LT-R72, como adjuvantes pode ser encontrado nas refs. 75-82. Referência numérica para substituições de aminoácido é preferivelmente baseada nos alinhamentos das subunidades A e B de toxinas ADP-ribosilação expostas na ref. 83, especificamente incorporada aqui a título de referência na íntegra.

F. imunomoduladores humanos

Imunomoduladores humanos apropriados para uso como adjuvantes na invenção incluem citocinas, como interleucinas (por exemplo, IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12 [84], etc.), interferons (por exemplo, interferon-

G. Bioadesivos e mucoadesivos

Bioadesivos e mucoadesivos também podem ser utilizados como adjuvantes na invenção. Bioadesivos

apropriados incluem microesferas de ácido hialurônico esterificado [86] ou mucoadesivos como derivados reticulados de ácido poliacrílico, álcool de polivinila, polivinil pirrolidona, polissacarídeos e carbóxi metil
5 celulose. Quitosana e derivados da mesma também podem ser utilizados como adjuvantes na invenção [87].

H. Micropartículas

Micropartículas também podem ser utilizadas como adjuvantes na invenção. Micropartículas (isto é, uma
10 partícula de

-tóxica (por exemplo, um ácido hidróxi), um ácido poliídrido, um poliortoéster, um polianidrido, um policaprolactona, etc., com poli(lactide-co-glicolide) são preferidos, opcionalmente tratados para ter uma superfície
15 negativamente carregada (por exemplo, com SDS) ou uma superfície positivamente carregada (por exemplo, com um detergente catiônico, como CTAB).

I. Lipossomas (capítulos 13 & 14 da ref. 35)

Os exemplos de formulações de lipossoma
20 apropriadas para uso como adjuvantes são descritos nas refs. 88-90).

J. Formulações de éster de polioxietileno e éter de polioxietileno

Adjuvantes apropriados para uso na invenção
25 incluem éteres de polioxietileno e ésteres de polioxietileno [91]. Tais formulações incluem ainda tensoativos de éster de sorbitana de polioxietileno em combinação com um octoxinol [92] bem como éteres de alquila de polioxietileno ou tensoativos de éster em combinação com
30 pelo menos um tensoativo não iônico adicional como

octoxinol [93]. Éteres de polioxietileno preferidos são selecionados a partir do seguinte grupo: éter de laurila-9-polioxietileno (laureth 9), éter de esteoríla-9-polioxietileno, éter de esteoríla-8-polioxietileno, éter de laurila-4-polioxietileno, éter de laurila-35-polioxietileno e éter de laurila-23-polioxietileno.

K. Polifosfazeno (PCPP)

Formulações PCPP são descritas, por exemplo, nas refs. 94 e 95.

L. Peptídeos de muramila

Os exemplos de peptídeos de muramila apropriados para uso como adjuvantes na invenção incluem N-acetil-muramila-L-treonila-D-isoglutamina (thr-MDP), N-acetil-normuramila-L-alanila-D-isoglutamina (nor-MDP), e N-acetilmuramila-L-alanila-D-isoglutaminila-L-alanina-2-(1'-2'-dipalmitoíla-sn-glicero-3-hidróxi fosforilóxi)-etil amina MTP-PE.

M. Compostos de imidazoquinolona

Os exemplos de imidazoquinolona apropriados para uso como adjuvantes na invenção incluem Imiquamod e seus homólogos (por exemplo, "Resiquimod 3M"), descritos adicionalmente nas refs. 96 e 97.

A invenção pode compreender também combinações e aspectos de um ou mais dos adjuvantes identificados acima. Por exemplo, as seguintes composições adjuvantes podem ser utilizadas na invenção: (1) uma emulsão de saponina e uma de óleo em água [98]; (2) uma saponina (por exemplo, QS21) + derivado LPS não tóxico (por exemplo, 3DMPL) [99]; (3) uma saponina (por exemplo, QS21) + um derivado LPS não tóxico (por exemplo, 3dMPL) + um colesterol; (4) uma

saponina (por exemplo, QS21) + 3dMPL + IL-12 (opcionalmente + um esteroide) [100]; (5) combinações de 3dMPL com, por exemplo, QS21 e/ou emulsões de óleo em água [101]; (6) SAF, contendo 10% de esqualeno, 0,4% de Tween 80TM, 5% de polímero de bloco-plurônico L121, e thr-MDP, microfluidificado em uma emulsão de submicron ou submetido a vórtice para gerar uma emulsão de tamanho de partícula maior. (7) sistema adjuvante Ribi TM (RAS), (Ribi Immunochem) contendo 2% esqualeno, 0,2% Tween 80 e um ou mais componentes de parede de célula bacteriana a partir do grupo que consiste em monofosforilipídeo A (MPL), dimicolato de trealose (TDM), e esqueleto de parede de célula (CWS), preferivelmente MPL + CWS (DetoxTM); e (8) um ou mais sais minerais (como sal de alumínio) + um derivado não tóxico de LPS (como 3dMPL).

Outras substâncias que atuam como agentes imunoestimuladores são reveladas no capítulo 7 da ref. 35.

O uso de um adjuvante de fosfato de alumínio ou hidróxido de alumínio é particularmente preferido, e antígenos são genericamente adsorvidos nesses sais. Fosfato de cálcio é outro adjuvante preferido.

O pH das composições da invenção está preferivelmente entre 6 e 8, preferivelmente aproximadamente 7. pH estável pode ser mantido pelo uso de um tampão. Onde uma composição compreende um sal de hidróxido de alumínio, prefere-se utilizar um tampão de histidina [102]. A composição pode ser estéril e/ou isenta de pirogênio. As composições da invenção podem ser isotônicas com relação a seres humanos.

As composições podem ser apresentadas em frascos,

ou podem ser apresentadas em seringas cheias-pontas. As seringas podem ser fornecidas com ou sem agulhas. Uma seringa incluirá uma única dose da composição, ao passo que um frasco pode incluir uma única dose ou múltiplas doses.

5 Composições injetáveis serão normalmente soluções ou suspensões líquidas. Alternativamente, podem ser apresentadas em forma sólida (por exemplo, liofilizada) para solução ou suspensão em veículos líquidos antes da injeção.

10 As composições da invenção podem ser embaladas em forma de dose unitária ou em forma de múltiplas doses. Para formas de múltiplas doses, frascos são preferidos em relação a seringas pré-cheias. Volumes de dosagem eficaz podem ser rotineiramente estabelecidos, porém uma dose
15 humana típica da composição para injeção tem um volume de 0,5 ml.

Onde uma composição da invenção deve ser preparada extemporaneamente antes do uso (por exemplo, onde um componente é apresentado em forma liofilizada) e é
20 apresentado como um kit, o kit pode compreender dois frascos, ou pode compreender uma seringa cheia-pronta e um frasco, com o conteúdo da seringa sendo utilizado para reativar o conteúdo do frasco antes da injeção.

Composições imunogênicas utilizadas como vacinas
25 compreendem uma quantidade imunologicamente eficaz de antígeno(s), bem como quaisquer outros componentes, conforme necessário. Por "quantidade imunologicamente eficaz", se quer dizer que a administração dessa quantidade em um indivíduo, em uma dose única ou como parte de uma
30 série, é eficaz para tratamento ou prevenção. Essa

quantidade varia dependendo da saúde e condição física do indivíduo a ser tratado; idade, grupo taxonômico de indivíduo a ser tratado (por exemplo, primata não humano, primata, etc.), a capacidade do sistema imune do indivíduo
5 sintetizar anticorpos, o grau de proteção desejado, a formulação da vacina, a avaliação do médico atendente sobre a situação médica, e outros fatores relevantes. Espera-se que a quantidade esteja compreendida em uma faixa relativamente ampla que pode ser determinada através de
10 experimentos de rotina, e uma quantidade típica de cada antígeno de sacarídeo meningocócico por dose está entre 1 µg e 10 mg por antígeno.

Usos farmacêuticos

A invenção também provê um método de tratar um
15 paciente, compreendendo administrar ao paciente uma quantidade terapêuticamente eficaz de uma composição da invenção. O paciente pode estar em risco da doença ele próprio ou pode ser uma mulher grávida ("imunização de material").

20 A invenção provê ácido nucléico, polipeptídeo, ou anticorpo da invenção para uso como medicamentos (por exemplo, como composições imunogênicas ou como vacinas, incluindo para uso no tratamento ou prevenção de doença e/infecção causada por meningococo); ou como reagentes de
25 diagnóstico. Também provê o uso de ácido nucléico, polipeptídeo, ou anticorpo da invenção na fabricação de:
(i) um medicamento para tratar ou prevenir doença e/ou infecção causada por meningococo; (ii) um reagente de diagnóstico para detectar a presença de meningococo ou de
30 anticorpos criados contra meningococo; e/ou (iii) um

reagente que pode criar anticorpos contra meningococo. O meningococo pode ser de qualquer sorogrupo ou cepa, porém está preferivelmente no sorogrupo B. A doença pode ser, por exemplo, meningite bacteriana (e particularmente meningite 5 meningocócica) ou septicemia.

O paciente é preferivelmente um ser humano. Onde a vacina é para uso profilático, o ser humano é preferivelmente uma criança (por exemplo, um bebê ou criança pequena) ou adolescente, por exemplo, idades de 0- 10 18 anos; onde a vacina é para uso terapêutico, o humano é preferivelmente um adulto, por exemplo, idade de 18-55 anos. Uma vacina destinada a crianças pode ser também administrada em adultos, por exemplo, para avaliar a segurança, dosagem, imunogenicidade, etc.

15 Um modo de verificar a eficácia de tratamento terapêutico envolve monitorar infecção meningocócica após administração da composição da invenção. Um modo de checar a eficácia de tratamento profilático envolve monitorar respostas imunes contra um polipeptídeo administrado após 20 administração. A imunogenicidade de composições da invenção pode ser determinada por administrar as mesmas em sujeitos em teste (por exemplo, crianças de 12-16 meses de idade, ou modelos de animal) e então determinar parâmetros padrão incluindo títulos ELISA (GMT) de IgG. Essas respostas 25 imunes serão genericamente determinadas em torno de 4 semanas após administração da composição, e comparadas com valores determinados antes da administração da composição. Onde mais de uma dose da composição é administrada, mais de uma determinação pós-administração pode ser feita. Um 30 método padrão para avaliar eficácia profilática para

meningococos é o ensaio bactericida de soro (SBA). A administração resulta, preferivelmente, em um aumento em título de SBA para o sorogrupo relevante de pelo menos 3 vezes, e preferivelmente pelo menos 8 vezes, medido com 5 complemento humano [103]. Se o complemento de coelho for utilizado para medir títulos de SBA então o aumento de título é preferivelmente pelo menos 128 vezes.

A administração de antígenos de polipeptídeo é um método de tratamento preferido para induzir imunidade. A 10 administração de anticorpos da invenção é outro método de tratamento preferido. Esse método de imunização passiva é particularmente útil para crianças recém-nascidas ou para mulheres grávidas. Esse método utilizará, tipicamente, anticorpos monoclonais, que serão humanizados ou totalmente 15 humanos.

As composições da invenção serão genericamente administradas diretamente em um paciente. A distribuição direta pode ser realizada por injeção parenteral (por exemplo, por via subcutânea, intraperitoneal, intravenosa, 20 intramuscular ou ao espaço intersticial de um tecido), ou por via retal, oral, vaginal, tópica, transdérmica, intranasal, ocular, aural, pulmonar ou outra administração em mucosa. A administração intramuscular na coxa ou no braço superior é preferida. A injeção pode ser via uma 25 agulha (por exemplo, uma agulha hipodérmica), porém injeção sem agulha pode ser alternativamente utilizada. Uma dose intramuscular típica é 0,5 ml.

A invenção pode ser utilizada para eliciar imunidade sistêmica e/ou de mucosa.

30 O tratamento de dosagem pode ser um programa de

dose única ou um programa de dose múltipla. Múltiplas doses podem ser utilizadas em um programa de imunização primária e/ou em um programa de imunização de reforço. Um programa de dose primária pode ser seguido por um programa de dose de reforço. O timing apropriado entre doses de iniciação (por exemplo, entre 4 - 16 semanas) e entre iniciação e reforço, pode ser rotineiramente determinada.

Infecções bacterianas afetam várias áreas do corpo e assim as composições podem ser preparadas de várias formas. Por exemplo, as composições podem ser preparadas como injetáveis, como soluções líquidas ou suspensões. Formas sólidas apropriadas para solução ou suspensão em veículos líquidos antes da injeção também podem ser preparadas (por exemplo, uma composição liofilizada). A composição pode ser preparada para administração tópica, por exemplo, como unguento, creme ou pó. A composição pode ser preparada para administração oral, por exemplo, como comprimido ou cápsula, ou como xarope (opcionalmente aromatizado). A composição pode ser preparada para administração pulmonar, por exemplo, como inalador, utilizando um pó fino ou uma pulverização. A composição pode ser preparada como supositório ou pessário. A composição pode ser preparada para administração nasal, aural ou ocular, por exemplo, como pulverização, gotas, gel ou pó [por exemplo, refs. 104 & 105].

Componentes antigênicos adicionais de composições da invenção

A invenção também provê uma composição que compreende um polipeptídeo ou a invenção e um ou mais dos seguintes antígenos adicionais:

- Um antígeno de sacarídeo a partir do sorogrupo N. meningitidis A, C, W135 e/ou Y (preferivelmente todos os quadro), como o oligossacarídeo revelado na ref. 106 a partir do sorogrupo C [vide, também, a ref. 107] ou os
5 oligossacarídeos da ref. 108.
- Um antígeno de sacarídeo da Streptococcus pneumoniae [por exemplo, 109, 110, 111].
- um antígeno do vírus de hepatite A, como vírus inativado [por exemplo, 112, 113].
- 10 - um antígeno do vírus de hepatite B, como os antígenos de superfície e/ou núcleo [por exemplo, 113, 114].
- um antígeno de difteria, como difteria toxóide [por exemplo, capítulo 3 de ref. 115], por exemplo, o
15 mutante CRM₁₉₇ [por exemplo, 116].
- um antígeno de tétano, como tétano toxóide [por exemplo, capítulo 4 da ref. 115].
- um antígeno de Bordetella pertussis, como holotoxina de coqueluche (PT) e hemaglutinina filamentosa
20 (FHA) de B. pertussis, também opcionalmente em combinação com pertactina e/ou aglutinogens 2 e 3 [por exemplo, refs. 117 & 118].
- um antígeno de sacarídeo de Haemophilus influenzae B [por exemplo, 107].
- 25 - Antígeno(s) de pólio [por exemplo, 119, 120] como IPV.
- antígenos de sarampo, caxumba e/ou rubéola [por exemplo, capítulos 9, 10 & 11 da ref. 115].
- antígeno(s) de influenza, [por exemplo,
30 capítulo 19 da ref. 115], como as proteínas de superfície

de hemaglutinina e/ou neuraminidase.

- um antígeno de *Moraxella catarrhalis* [por exemplo, 121].

5 - Um antígeno de proteína de *Streptococcus agalactiae* (estreptococo do grupo B) [por exemplo, 122, 123].

- um antígeno de sacarídeo de *Streptococcus agalactiae* (estreptococo do grupo B).

10 - um antígeno de *Streptococcus pyogenes* (estreptococo grupo A) [por exemplo, 123, 124, 125]

- um antígeno de *Staphylococcus aureus* [por exemplo, 126].

A composição pode compreender um ou mais desses antígenos adicionais.

15 Antígenos de proteína tóxica podem ser detoxificados onde necessário (por exemplo, detoxificação de toxina de coqueluche por meio químico e/ou genético [118]).

20 Onde um antígeno de difteria é incluído na composição, prefere-se também incluir antígeno de tétano e antígeno de pertussis. Similarmente, onde um antígeno de tétano é incluído, prefere-se também incluir antígenos de difteria e pertussis. Similarmente, onde um antígeno de pertussis é incluído prefere-se incluir também antígenos de
25 difteria e tétano. Combinações de DTP são assim preferidas.

Antígenos de sacarídeo, preferivelmente estão na forma de conjugados. Proteínas portadoras para os conjugados incluem toxina de difteria, toxina de tétano, a proteína de membrana externa de *N. meningitidis* [127],
30 peptídeos sintéticos [128, 129], proteínas de choque

térmico [130, 131], proteínas de pertussis [132, 133], proteína D a partir de H. influenzae [134], citocinas [135], linfocinas [135], proteínas estreptococais, hormônios [135], fatores de crescimento [135], toxina A ou B de C. difficile [136], proteínas de absorção de ferro [137], etc. Uma proteína portadora preferida é a difteria toxóide CRM197 [138].

Antígenos na composição estarão, tipicamente, presentes em uma concentração de pelo menos 1 µg/ml cada. Em geral, a concentração de qualquer antígeno dado será suficiente para eliciar uma resposta imune contra aquele antígeno.

Como alternativa ao uso de antígenos de proteína nas composições imunogênicas da invenção, pode ser utilizado ácido nucléico (preferivelmente DNA, por exemplo, na forma de um plasmídeo) codificando o antígeno.

Antígenos são preferivelmente adsorvidos em um sal de alumínio.

Métodos de classificação

A invenção provê um processo para determinar se um composto de teste se liga a um polipeptídeo da invenção. Se um composto de teste se liga a um polipeptídeo da invenção e essa ligação inibe o ciclo de vida de meningococo, então o composto de teste pode ser utilizado como um antibiótico ou como um composto principal para a elaboração de antibióticos. O processo compreenderá, tipicamente, as etapas de contatar um composto de teste com um polipeptídeo da invenção, e determinar se o composto de teste se liga ao polipeptídeo. Polipeptídeos preferidos da invenção para uso nesses processos são enzimas (por

exemplo, sintetase tRNA), transportadores de membrana e polipeptídeos ribossomais. Compostos de teste apropriados incluem polipeptídeos, polipeptídeos, carboidratos, lipídeos, ácidos nucleicos (por exemplo, DNA, RNA, e formas modificadas dos mesmos), bem como compostos orgânicos pequenos (por exemplo, MW entre 200 e 2000 Da). Os compostos de teste podem ser fornecidos individualmente, porém farão, tipicamente parte de uma biblioteca (por exemplo, uma biblioteca de combinação). Os métodos para detectar uma interação de ligação incluem NMR, ensaios de ligação de filtro, ensaios de retardação de gel, ensaios de deslocamento, ressonância de plasmônio superficial, híbrido-dois inverso, etc. Um composto que se liga a um polipeptídeo da invenção pode ser testado em relação à atividade de antibiótico pelo contato do composto com bactérias de meningococo e, então, monitorando-se em termos de inibição de crescimento. A invenção também provê um composto identificado utilizando esses métodos.

Preferivelmente, o processo compreende as etapas de: (a) contatar um polipeptídeo da invenção com um ou mais compostos candidatos para fornecer uma mistura; (b) incubar a mistura para permitir interação do polipeptídeo e o(s) composto(s) candidato(s); e (c) avaliar se o composto candidato se liga ao polipeptídeo ou modula sua atividade.

Após um composto candidato ter sido identificado in vitro como composto que se liga a um polipeptídeo da invenção então pode ser desejável executar experimentos adicionais para confirmar a função in vivo do composto em inibir crescimento e/ou sobrevivência bacteriana. Desse modo, o método compreende a etapa adicional de contatar o

composto com um meningococo e avaliar seu efeito.

O polipeptídeo utilizado no processo de classificação pode estar livre em solução, afixado a um suporte sólido, localizado em uma superfície de célula ou localizado intracelularmente. Preferivelmente, a ligação de um composto candidato ao polipeptídeo é detectada por intermédio de um rótulo direta ou indiretamente associado ao composto candidato. O rótulo pode ser um fluoróforo, radioisótopo ou outro rótulo detectável.

10

Geral

A invenção provê um meio legível por computador (por exemplo, um disco flexível, um disco rígido, um CD-ROM, um DVD, etc.) e/ou uma memória de computador e/ou um banco de dados de computador contendo uma ou mais das sequências na listagem de sequência.

15

O termo "compreendendo" abrange "incluindo" bem como "consistindo", por exemplo, uma composição "compreendendo" X pode consistir exclusivamente em X ou pode incluir algo adicional, por exemplo, X + Y.

20

O termo "aproximadamente" em relação a um valor numérico x significa, por exemplo, $x \pm 10\%$.

25

A palavra "substancialmente" não exclui "totalmente", por exemplo, uma composição que é "substancialmente isenta" de Y pode ser totalmente isenta de Y. Onde necessário, a palavra "substancialmente" pode ser omitida a partir da definição da invenção.

30

A identidade entre polipeptídeos é preferivelmente determinada pelo algoritmo de busca de homologia Smith-Waterman como implementado no programa MPSRCH (Oxford Molecular), utilizando uma busca de lacuna

de afinidade com parâmetros penalidade de abertura de lacuna = 12 e penalidade de extensão de lacuna = 1. A identidade entre sequências é preferivelmente determinada, também pelo algoritmo de busca de homologia Smith-Waterman.

5 Os resíduos de extremidade-N nas sequências de aminoácidos na listagem de sequência são dados como o aminoácido codificado pelo primeiro códon na sequência de nucleotídeos correspondente. Onde o primeiro códon não é ATG, será entendido que será traduzido como metionina
10 quando o códon é um códon de partida, porém será traduzido como o aminoácido não-Met indicado quando a sequência está na extremidade-C de um associado de fusão. A invenção revela especificamente e abrange cada uma das sequências de aminoácidos da listagem de sequência tendo um resíduo de
15 metionina de extremidade-N (por exemplo, um resíduo de metionina-formila) no lugar de qualquer resíduo não-Met indicado.

Códons de partida alternativos podem ser utilizados em biologia. As sequências de aminoácido na
20 listagem de sequência são baseadas em códons de partida específicos, porém códons de partida a jusante podem ser alternativamente utilizados. Desse modo, a invenção revela e abrange, especificamente, cada uma das sequências de aminoácido da listagem de sequência, iniciando em qualquer
25 resíduo de metionina a partir da sequência que está a jusante do resíduo de terminal-N mostrado na listagem de sequência (por exemplo, SEQ ID NOS: 5 & 10).

Como indicado no texto acima, os ácidos nucléicos e os polipeptídeos da invenção podem incluir sequências
30 que:

(a) são idênticas (isto é, 100% idênticas) às sequências reveladas na listagem de sequência;

(b) compartilham identidade de sequência com as sequências reveladas na listagem de sequência;

5 (c) têm 1, 2, 3, 4, 5, 6, 7, 8, 9 ou 10 alterações de aminoácido ou nucleotídeo único (deleções, inserções, substituições), que podem estar em locais separados ou podem ser contíguos, como comparado com as sequências de (a) ou (b); e

10 (d) quando alinhadas com uma sequência específica a partir da listagem de sequência utilizando um algoritmo de alinhamento de pares, uma janela móvel de x monômeros (aminoácidos ou nucleotídeos) que se movem do início (extremidade-N ou 5') para o final (extremidade-C de 3') de
15 tal modo que para um alinhamento que se estende até p monômeros (onde $p > x$), há $p-x+1$ de tais janelas, cada janela tem pelo menos $x.y$ monômeros alinhados idênticos, onde: x é selecionado de 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200; y é selecionado de 0,50, 0,60, 0,70, 0,75,
20 0,80, 0,85, 0,90, 0,91, 0,92, 0,93, 0,94, 0,95, 0,96, 0,97, 0,98, 0,99; e se $x.y$ não for um número inteiro então é arredondado até o número inteiro mais próximo. O algoritmo de alinhamento em pares preferido é o algoritmo de alinhamento global Needleman-Wunsch [139], utilizando
25 parâmetros default (por exemplo, com penalidade de abertura de Lacuna = 10,0, e com penalidade de extensão de Lacuna = 0,5, utilizando a matriz de marcação EBLOSUM62). Esse algoritmo é convenientemente implementado na ferramenta needle no pacote EMBOSS [140].

30 Os ácidos nucleicos e polipeptídeos da invenção

podem ter, adicionalmente, sequências adicionais à extremidade-N/5' e/ou extremidade-C/3' dessas sequências (a) a (d).

A prática da presente invenção empregará, a menos que de outro modo indicado, métodos convencionais de química, bioquímica, biologia molecular, imunologia e farmacologia, nos conhecimentos da técnica. Tais técnicas são explicadas completamente na literatura. Vide, por exemplo, as referências 141-148, etc.

10 **MODOS PARA REALIZAR A INVENÇÃO**

Várias sequências de aminoácidos codificados foram identificadas no genoma da cepa M04-240196 de sorogrupo B. N. meningitidis. 39 delas foram selecionadas como antígenos úteis, com base em vários critérios, e seu gene e sequências de aminoácido são fornecidos na listagem de sequência.

Funções biológicas previstas são fornecidas na tabela 1, porém os papéis precisos dos antígenos em biologia de meningococo não são tão importantes como sua capacidade de funcionar como imunógenos. A Tabela 1 também observa o casamento mais estreito nos genomas de sorogrupo A e B publicados nas referências 6 e 8, bem como no genoma não publicado de cepa FAM18 do sorogrupo C. Onde uma sequência tem mais de 95% de identidade com uma sequência conhecida (e particularmente onde tem 100% de identidade) então a invenção está mais preocupada com a identificação das propriedades antigênicas úteis da proteína do que identificação da proteína por si.

Além das anotações e comparações, características adicionais de interesse incluem: B269_17 contém um domínio

de intein; B269_34 tem uma sequência de junção em sua extremidade-C; um domínio de transmembrana está presente em B269_05, B269_10, B269_18, B269_24; e cinco domínios de transmembrana estão presentes em B269_15 e B269_29.

5 Utilizando a informação de sequência da presente invenção, as proteínas podem ser prontamente expressas em hospedeiros recombinantes e utilizadas para gerar respostas imunes utilizando técnicas conhecidas na arte.

Por exemplo, as sequências que codificam
10 proteínas B269_11, 13, 14, 15, 17, 24, 25, 26, 29, 31, 32, 34, 36, 37, 51, 52, 53, 54, 55, 56, 57, 58 e 59 foram inseridas nos vetores de expressão com um marcador polihistidina terminal-C. Proteínas B269_14, 29, 31, 34, 37 foram também expressas em uma forma de domínio-truncado.
15 Fusões de GST de 37, 54, 55 e 57 foram também preparadas. Proteínas expressas foram purificadas a partir de E. coli. Sem nenhuma otimização de expressão, vários graus de pureza foram vistos, por exemplo, de 20% de pureza com domínio B269_14 e B269_32, até 95% de pureza com B269_51. Expressão
20 solúvel foi vista com proteínas B269_13, 24, 25, 31_{domain}, 32, 51, 53 e 56.

Anticorpos foram criados contra proteínas expressas por injetar os mesmos em camundongos com adjuvante completo de Freund ou hidróxido de alumínio.
25 anti-soros foram então utilizados em western blots ou ensaios de ligação FACS contra meningococos. As seguintes proteínas B269_ poderiam ser detectadas por western blot: 13; 25; 29_{domain}; 51; 52 e 53. Além disso, as seguintes proteínas poderiam ser detectadas nos blots em MWS
30 específicos: 11 (40 kDa); 24 (20 kDa); e 26 (28 kDa). FACS

revelaram as seguintes proteínas: 17; 24; 26; 29_{domain}; 34_{domain}; e53.

5 Será entendido que a invenção foi descrita somente como exemplo e modificações podem ser feitas enquanto permanecem no escopo e espírito da invenção.

Tabela 1 - Anotações

aa = comprimento de polinucleotídeo

10

15

20

25

30

| B269 | SEQID | aa | Localização * | Anotação |
|------|-------|------|---------------|---|
| 01 | 2 | 588 | 0 | Proteína PilC |
| -11 | 4 | 320 | 0 | Adesina semelhante a MafA putativo adhesin |
| 13 | 6 | 387 | c | Proteína de superfamília Cupin |
| 14 | 8 | 420 | i | Proteína de fusão-membrana |
| 15 | 10 | 670 | i | peptidase, família C39 |
| 16 | 12 | 331 | p | cis-trans isomerase peptidil-profil putativo |
| 17 | 14 | 209 | s | Proteína hipotética conservada |
| 18 | 16 | 265 | 0 | Proteína de opacidade |
| 19 | 18 | 680 | i | Proteína de ligação-transferrina 2 |
| 20 | 20 | 1370 | s | Proteína relacionada a hemaglutinina-hemolisina |
| 21 | 22 | 734 | s | Proteína relacionada a hemaglutinina-hemolisina |
| 22 | 24 | 887 | s | Proteína relacionada a hemaglutinina-hemolisina |
| 23 | 26 | 794 | s | Proteína relacionada a hemaglutinina-hemolisina |
| 24 | 28 | 206 | c | Proteína hipotética conservada |
| 26 | 30 | 1502 | s | Proteína semelhante a hemolisina HlyJ protein |
| 26 | 32 | 257 | G | Proteína hipotética conservada TIGR00294 |
| 27 | 34 | 237 | i | transferase de Enxofre tiosulfato putativo |
| 28 | 36 | 402 | C | provável: GlcNAc transferase 19 Putativo |
| 29 | 38 | 297 | 0 | Proteína hipotética conservada |
| 30 | 40 | 226 | c | Proteína de opacidade Opal 15 |
| 31 | 42 | 588 | 0 | Família da região terminal-C semelhante a YadA |
| 32 | 44 | 201 | c | Proteína hipotética conservada |
| 33 | 46 | 337 | i | Glicosil transferase putativo |
| 34 | 48 | 529 | c | Proteína regulada por ferro frpA |
| 35 | 50 | 676 | c | Proteína de ligação de transferrina/lactoferrina B |
| 36 | 52 | 203 | c | mucina |
| 37 | 54 | 340 | c | Proteína hipotética conservada |
| 38 | 56 | 376 | i | Sensor de dois componentes putativo cinase 196 |
| 39 | 58 | 346 | p | Região de ligação-ATP, semelhante a ATPase-Histidina cinase A, N-terminal |
| 41 | 60 | 1028 | p | proteína PilC |
| 42 | 62 | 333 | 0 | Proteína hipotética conservada |
| 43 | 64 | 229 | c | glicosil transferase, proteína de família grupo 2 |
| 44 | 66 | 208 | c | Proteína hipotética conservada |
| 45 | 68 | 478 | c | Proteína mafB |
| 46 | 70 | 229 | c | adesina MafB |
| 47 | 72 | 432 | 0 | Proteína de ligação de transferrina B subunidade 19 |
| 48 | 74 | 809 | s | precursor HecA Proteína semelhante a hemolisina-hemaglutinina HecA putativo |
| 49 | 76 | 783 | s | Hemaglutinina possível(DUF637) família 1 |
| 50 | 78 | 300 | s | Proteína relacionada a hemaglutinina-hemolisina hemagglutinin-hemolysin-related protein |

*Chave de localização: O = membrana externa; C = citoplasma; I = membrana interna; P = espaço periplásmico; S = secretada

Tabela 2 - Relação para outras sequências
5 meningocócicas [6, 8]

'MC58' = correspondente mais fechado a partir da referência 6.

Colunas (a) a (c) são % de correspondentes para outras sequências: (B) = ref6; (A) = ref8; (C) = cepa FAM18

| B269 | SEQID | MC58 | (B) | (A) | (C) |
|------|-------|---------|------|------|------|
| 01 | 2 | NMB0049 | 77.6 | 78.2 | 73.1 |
| 11 | 4 | NMB0652 | 62,7 | 100 | 98,4 |
| 13 | 6 | NMB1786 | 51 | 48.4 | 80.3 |
| 14 | 8 | NMB0097 | 51.1 | 59.9 | 100 |
| 15 | 10 | NMB0098 | 64 | 28.9 | 94.9 |
| 16 | 12 | NMB0281 | 85.4 | 92.1 | 93.5 |
| 17 | 14 | NMB0369 | 83.2 | 61.7 | 91.9 |
| 18 | 16 | NMB0442 | 86.4 | 81.5 | 87.1 |
| 19 | 18 | NMB0460 | 71.5 | 70.9 | 41.8 |
| 20 | 20 | NMB1779 | 67.4 | 70 | 97.4 |
| 21 | 22 | NMB1775 | 82.6 | 87.9 | 96 |
| 22 | 24 | NMB1779 | 81.9 | 80.5 | 95.2 |
| 23 ; | 26 | NMB1775 | 79.2 | 89.2 | 88.9 |
| 24 | 28 | NMB0515 | 66.1 | 99.5 | 93.7 |
| 25 | 30 | NMB0585 | 85,5 | 91 | 90,9 |
| 26 | 32 | NMB0803 | 86.8 | 86.8 | 86.8 |
| 27 | 34 | NMB0841 | 84 | 95.4 | 88.1 |
| 28 | 36 | NMB0846 | 89 | 56.1 | 85.1 |
| 29 | 38 | NMB0888 | 87.9 | 79.8 | 85.2 |
| 30 | 40 | NMB1636 | 80.4 | 81.6 | 92 |
| 31 | 42 | NMB0992 | 87.4 | 90.1 | 94.6 |
| 32 | 44 | nenhum | 41.2 | 41.2 | 70.9 |
| 33 | 46 | NMB1255 | 73 | 95.4 | 97.5 |
| 34 | 48 | NMB1415 | 89.5 | 80.5 | 98.7 |
| 35 | 50 | NMB1541 | 68.3 | 78.2 | 56.4 |
| 36 | 52 | NMB0891 | 49 | 53.4 | 78.6 |
| 37 | 54 | nenhum | 30.7 | 33.3 | 100 |
| 38 | 56 | NMB1606 | 82.2 | 80.1 | 87 |
| 39 | 58 | NMB1606 | 83 | 85.9 | 90.9 |
| 41 | 60 | NMB1847 | 80.4 | 81.6 | 78.2 |

| | | | | | |
|----|----|---------|------|------|------|
| 42 | 62 | NMB1870 | 87.4 | 87.4 | 74.2 |
| 43 | 64 | NMB1929 | 48.2 | 48.2 | 48.2 |
| 44 | 66 | NMB1992 | 89.9 | 81.5 | 92.8 |
| 45 | 68 | NMB2105 | 89.7 | 88.5 | 95.9 |
| 46 | 70 | NMB2105 | 88 | 96.7 | 89.9 |
| 47 | 72 | NMB2132 | 76,5 | 68,6 | 68,7 |
| 48 | 74 | NMB0493 | 85,4 | 55 | 68,2 |
| 49 | 76 | NMB1775 | 67,2 | 82,9 | 99,4 |
| 50 | 78 | NMB1779 | 51,2 | 84 | 83,3 |

REFERÊNCIAS (cujos conteúdos são incorporados pelo presente a título de referência)

- [1] WO99/24578
- [2] WO99/36544
- 5 [3] WO99/57280.
- [4] WO00/22430.
- [5] WO00/66791.
- [6] Tettelin e outros (2000) *Science* 287: 1809-16.
- 10 [7] Pizza e outros (2000) *Science* 288:1-1816-20.
- [8] Parkhill e outros (2000) *Nature* 404: 502-8.
- [9] Geysen e outros (1984) *PNAS USA* 81:3998-4002.
- [10] Carter (1994) *Methods Mol Bio.* 36:207-23.
- [11] Jameson, BA e outros 1988, *CABIOS*4(1):181-15 186.
- [12] Raddrizzani & Hammer (2000) *Brief Bioinform* 1(2):179-89.
- [13] De Lalla e outros (1999) *J. Immunol.* 163:1725-29.
- 20 [14] Brusica e outros (1998) *Bioinformatics* 14(2): 121-30
- [15] Meister e outros (1995) *Vaccine* 13(6):581-91.
- [16] Roberts e outros (1996) *AIDS Res Hum Retroviruses* 12(7):593-610.

- [17] Maksyutov & Zagrebelnaya (1993) *Comput Appl Biosci* 9(3):291-7.
- [18] Feller & de la Cruz (1991) *Nature* 349(6311):720-1.
- 5 [19] Hopp (1993) *Peptide Research* 6:183-190.
- [20] Welling e outros (1985) *FEBS Lett.* 188:215-218.
- [21] Davenport e outros (1995) *Immunogenetics* 42:392-297.
- 10 [22] Bodanszky (1993) *Principles of Peptide Synthesis* (ISBN: 0387564314).
- [23] Fields e outros (1997) *Meth Enzymol* 289: *Solid-Phase Peptide Synthesis*. ISBN: 0121821900.
- [24] Chan & White (2000) *Fmoc Solid Phase Peptide*
15 *Synthesis*. ISBN: 0199637245.
- [25] Kullmann (1987) *Enzymatic Peptide Synthesis*. ISBN: 0849368413.
- [26] Ibba (1996) *Biotechnol Genet Eng Rev* 13:197-216.
- 20 [27] Breedveld (2000) *Lancet* 355(9205):735-740.
- [28] Gorman & Clark (1990) *Semin. Immunol.* 2:457-466.
- [29] Sambrook e outros (1989) *Molecular Cloning: A Laboratory Manual*.
- 25 [30] *Short protocols in molecular biology* (4{a ed., 1999) Ausubel e outros. eds. ISBN 0-471-32938-X.
- [31] patente US 5.707.829
- [32] *Current Protocols in Molecular Biology* (F.M. Ausubei e outros eds., 1987) Supplement 30.
- 30 [33] EP-B-0509612.

- [34] EP-B-0505012.
- [35] *Vaccine Design...* (1995) eds. Powell & Newman. ISBN: 030644867X. Plenum.
- [36] WO00/23105.
- 5 [37] WO90/14837.
- [38] patente US 5.057.540.
- [39] WO96/33739.
- [40] EP-A-0109942.
- [41] WO96/11711.
- 10 [42] WO00/07621.
- [43] Barr e outros. (1998) *Advanced Drug Delivery Reviews* 32:247-271.
- [44] Sjolanderet e outros (1998) *Advanced Drug Delivery Reviews* 32:321-338.
- 15 [45] Nikura e outros (2002) *Virology* 293:273-280.
- [46] Lenz e outros. (2001) *J Immunol* 166:5346-5355.
- [47] Pinto e outros. (2003) *J Infect Dis* 188:327-338.
- 20 [48] Gerber e outros. (2001) *Virology* 75:4752-4760.
- [49] WO03/024480
- [50] WO03/024481
- [51] Gluck e outros. (2002) *Vaccine* 20:B10-B16.
- [52] EP-A-0689454.
- 25 [53] Johnson e outros. (1999) *Bioorg Med Chem Lett* 9:2273-2278.
- [54] Evans e outros. (2003) *Expert Rev Vaccines* 2:219-229.
- [55] Meraldi e outros. (2003) *Vaccine* 21:2485-30 2491.

- [56] Pajak e outros (2003) *Vaccine* 21:836-842.
- [57] Kandimalla e outros. (2003) *Nucleic Acids Research* 31 -.2393-2400.
- [58] WO02/26757.
- 5 [59] W099/62923.
- [60] Krieg (2003) *Nature Medicine* 9:831-835.
- [61] McCluskie e outros. (2002) *FEMS Immunology and Medical Microbiology* 32:179-185.
- [62]WO98/40100.
- 10 [63] patente US 6.207.646.
- [64] patente US 6.239.116.
- [65] patente US 6.429.199.
- [66] Kandimalla e outros. (2003) *Biochemical Society Transactions* 31 (part 3):654-658.
- 15 [67] Blackwell e outros. (2003) *J Immunol* 170:4061-4068.
- [68] Krieg (2002) *Trends Immunol* 23:64-65.
- [69] WO01/95935.
- [70] Kandimalla e outros. (2003) *BBRC* 306:948-953.
- 20 [71] Bhagat e outros. (2003) *BBRC* 300:853-861.
- [72] WO03/035836.
- [73] W095/17211.
- [74] W098/42375.
- [75] Beignon e outros. (2002) *Infect Immun* 25 70:3012-3019.
- [76] Pizza e outros. (2001) *Vaccine* 19:2534-2541.
- [77] Pizza e outros. (2000) *Int J Med Microbiol* 290:455-461.
- [78] Scharton-Kersten e outros. (2000) *Infect* 30 *Immun* 68:5306-5313.

- [79] Ryan e outros. (1999) *Infect Immun* 67:6270-6280.
- [80] Partidos e outros. (1999) *Immunol Lett* 67:209-216.
- 5 [81] Peppoloni e outros. (2003) *Expert Rev Vaccines* 2:285-293.
- [82] Pine e outros (2002) *J Control Release* 85:263-270.
- [83] Domenighini e outros. (1995) *Mol Microbiol*
10 15:1165-1167.
- [84] WO99/40936.
- [85] WO99/44636.
- [86] Singh e outros (2001) *J Cont Release* 70:267-276.
- 15 [87] WO99/27960.
- [88] patente US 6.090.406
- [89] patente US 5.916.588
- [90] EP-A-0626169.
- [91] WO99/52549.
- 20 [92] WO01/21207.
- [93] WO01/21152.
- [94] Andrianov e outros. (1998) *Biomaterials* 19:109-115.
- [95] Payne e outros. (1998) *Adv Drug Delivery*
25 *Review* 31:185-196.
- [96] Stanley (2002) *Clin Exp Dermatol* 27:571-577.
- [97] Jones (2003) *Curr Opin Investig Drugs* 4:214-218.
- [98] WO99/11241.
- 30 [99] WO94/00153.

- [100] WO98/57659.
- [101] pedidos de patente europeia 0835318, 0735898 e 0761231.
- [102] WO03/009869.
- 5 [103] *W.H.O. Tech. Rep. Ser.* 594:51, 1976.
- [104] Almeida & Alpar (1996) *J. Drug Targeting* 3:455-467.
- [105] Agarwal & Mishra (1999) *Indian J Exp Biol* 37:6-16.
- 10 [106] Costantino e outros. (1992) *Vaccine* 10:691-698.
- [107] Costantino e outros. (1999) *Vaccine* 17:1251-1263.
- [108] WO03/007985.
- 15 [109] Watson (2000) *Pediatr Infect Dis J* 19:331-332.
- [110] Rubin (2000) *Pediatr Clin North Am* 47:269-285, v.
- [111] Jedrzejas (2001) *Microbiol Mol Biol Rev* 20 65:187-207.
- [112] Bell (2000) *Pediatr Infect Dis J* 19:1187-1188.
- [113] Iwarson (1995) *APMIS* 103:321-326.
- [114] Gerlich e outros. (1990) *Vaccine* 8 25 SuppI:S63-68 & 79-80.
- [115] *Vaccines* (1988) eds. Plotkin & Mortimer. ISBN 0-7216-1946-0.
- [116] Del Guidice e outros. (1998) *Molecular Aspects of Medicine* 19:1-70.

- [117] Gustafsson e outros. (1996) *TV. Engl. J. Med.* 334:349-355.
- [118] Rappuoli e outros. (1991) *TIBTECH* 9:232-238.
- [119] Sutter e outros. (2000) *Pediatr Clin North Am* 47:287-308.
- [120] Zimmerman & Spann (1999) ,4/77 *Fam Physician* 59:113-118, 125-126.
- [121] McMichael (2000) *Vaccine* 19 Suppl 1:S101-107.
- [122] Schuchat (1999) *Lancet* 353(9146):51 -6.
- [123] WO02/34771.
- [124] Dale (1999) *Infect Dis Clin North Am* 13:227-43, viii.
- [125] Ferretti e outros. (2001) *PNAS USA* 98: 4658-4663.
- [126] Kuroda e outros. (2001) *Lancet* 357(9264): 1225-1240; vide também as páginas 1218-1219.
- [127] EP-A-0372501
- [128] EP-A-0378881
- [129] EP-A-0427347
- [130] WO93/17712
- [131] WO94/03208
- [132] W098/58668
- [133] EP-A-0471177
- [134] WO00/56360
- [135] WO91/01146
- [136] WO00/61761
- [137] WO 01/72337
- [138] *Research Disclosure*, 453077 (Jan 2002)

[139] Needleman & Wunsch (1970) *J. Mol. Biol.* 48, 443-453.

[140] Rice e outros. (2000) *Trends Genet* 16:276-277.

5 [141] Gennaro (2000) *Remington: The Science and Practice of Pharmacy*. 20ª edição, ISBN: 0683306472.

[142] *Methods In Enzymology* (S. Colowick and N. Kaplan, eds., Academic Press, Inc.)

[143] *Handbook of Experimental Immunology*, Vols. I-IV (D.M. Weir e C.C. Blackwell, eds., 1986, Blackwell Scientific Publications)

[144] Sambrook, e outros, *Molecular Cloning: A Laboratory Manual* (2ª edição, 1989).

15 [145] *Handbook of Surface and Colloidal Chemistry* (Birdi, K.S. ed., CRC Press, 1997)

[146] *Short Protocols in Molecular Biology*, 4ª ed. (Ausubel e outros eds., 1999, John Wiley & Sons)

[147] *Molecular Biology Techniques: An Intensive Laboratory Course*, (Ream e outros, eds., 1998, Academic Press)

20 [148] *PCR (Introduction to Biotechniques Series)*, 2ª ed. (Newton & Graham eds., 1997, Springer Verlag).

LISTAGEM DE SEQUÊNCIAS

SEQ ID NO: 1

ATGAACCGCAAAATCTGCCTGAGGTAAAGTGGGGCAGGACTATAGAAAATTTGGCGCAAAAAGCAATGAACGCCAATTTACCC
 ATACGACTAATTTTTACACCAAAAAAAGCTAAGTTTATCATTCAATAATACCGATGAAGTTGTTGCTAAAAAAGCGTACTGT
 CGTTTTCGGCGCGGCGACCTACCTGCCGCCCTACGGCAAGTTTCCGGTTTTGATGACAAAAGGCTGAAAGAGCGCGGCAATGCC
 GTTAATTGGATTATACGACCGGACCCAGGGTTGGTAGGCTACAGCTACGGAAGATGTTACATGCAACAGCGGGCAACTGTCTGAAG
 TTAGCTATAAAACCCAAATTTATCTTGCATAATCACCAGTTGGCAAAAAGAAAACAGACAGCAAGCTGGATATATACGAAGACAA
 AAGCCGCGCAAAATTCGCCCATTTACAAATTCAGGATTTATCCTTGGTTGGCGGTGTCTTTCAATTTGGCGGAGAGAGCTCTTC
 AAACCAAAGAGCAAGGTTCTTTGGTATCTTTTAGCGAGGATGTGACGCGAGCAAAATGGTACACAAGACCAATACAAAAGGCA
 AAAACCTCGTTATACGACAGAAATTCAATAATCAGGGTAATCGTAACCATCAGGACAAAACCCACGCCATCCCTTTTATCT
 GAACGCCAAACTGCACCTGCTGGACAAAAGGGATTAAAGATATCACCACAAAACAGTGCAGTTGGGTGTCTTGAACCGCGC
 ATCGATTTGACGGAAAGCGTGGAAAACAGGCGATGGGAGCTTTTTGGTAATGGTAATGGACGTTTTGAAGATAAAGGAGCAGTCA
 GCGTCAAATTTATCTTGGCGAAGTCAAAGCAGGGCGTGCATCAATAAACCGAACCCCAATCCCAAAGCCCAAGCCCTTTCCCC
 CGCACTGACTGCCCGCGCTGTGGTTCCGACCTGTGCAAAATGGTAAGATGGAGATGTAATCCGCTTCCGTTTTCTACCTACCCC
 GACAGTTGAGCAGCGCCATCTTCTTCAAATCTGAAAAGAAAACGACCCCAACAAACCCGGCCGCTATTCCCTCGCAGACT
 TGAGCGGTGCGAGATTTAAAGTAAGAGCCGACTTTCACAGGGCGGCAACCGTCAATCCGATTGGATAAAGGCGTACATCAGAT
 CAAACTAAAGGCAATGAGGTGCAAGGTTTTAAGGAAACAGCGCAACGACACTTTCCGCTTGTAGTGAAGGAGCTTATG
 CCTGATGACGCGAAGTGGAAAAGTGTCTGCTTGGACGTTTGGTCTGCTCAATGATGACGGTCAATTTAACACAGTCAACA
 AAGAAGAAAACAGCGCAAGCAAAATACAGTCAAATAACCGCAGCCCAACCAACGGCAAGCAGGCGCAATTTGGGCGACAT
 CGTCAACAGCCCATCGTGGCGGTGGCGAGTATTTGGCTACTCCGCAACAGCAGGATGGTGCATATCTTCAAAGGAGCGCG
 GGGGATGACCGCAACTATAGTCTAAGCTCAGTTATATCCCGGCGCAGTGGCGCGCAAGGATTTCAAAGCCAAAGCTCCACCC
 TTGCCAAAGAGCTGGCGCTTTGCCGAAAAGCTATGTGGCGACCGCTACGGCGTGGACGGCGGCTTTGTCTTGGCCCAAGT
 CGAATGAAAAGGGCAAAACCGCGTGTATGTTCCGCGGATGNNNNNACACACTAATTAATTA

SEQ ID NO: 2

MNAQNLPEVKWGDYRKLQKSNRQFHTHTNFYTKKNVLSFNNTDEVVAKKNGTVVFGAATYLPYKGVSGFDDKRLKERGNA
 VNWIHTTDPGLVGYSEYEDVTCNSGNCPEVSYKQFIFDNHQLAKKTDKSLDIYEDKSRDNPISYKLDQYVPLGVSNLGGESSF
 KPKRQSSLYSFFSEVDVTKQNGTQDQYKGNLYTTEDYNNQGNRNHQDKHIIAFYLNALHLLDKKGKIDITDKTQVLGVLKPR
 IDLTEAWKNRHGSFFGNGNWTFFEDKGAVSVKLIPEVKAGRCINKPNPNPKAQLSPAL TAPALWFGPVQNGKMEMYSASVSTYP
 DSSSRIFLQNLKRKNDPNKPGRYSLADLSASEIKSKEPTFTGRQTVIRLDKGVHQIKLKGNEVEFGKNGNNDTFGIVSEGSFM
 PDDSEWKVLLPWTVRAFNDQDFNTVKEENNGPKYKQYRSRNNKHERNLGDIVNSPIVAVGEYLATASANDGMWHIIFKKS
 GDDRNYSKLSYIPGTPMRKDIQSQDSTLAKELRAFAEKSVVGDYVGDGFFVLRQVEWKQNRVFMFGAMXXTHLIN

SEQ ID NO: 3

ATGCAAGCACGGCTGCTGATACCTATTCTTTTTTTCAGTTTTTATTTTATCCGCTGCCGGACACTGACAGGTATCCATCGCATG
 CGCGAGGTAACGCTTTGCGGTGCAACAAGAACTTGTGGCGCTTCTCGCAGAGCTGCCGTTAAAGACATGGATTTACAGGCATT
 ACACGGACGAAAAGTTGCATTGTACATTGCAACTATGGGCGACCAAGGTTGAGGCAAGTTTGAAGGGGGTCCGTAATCCATGAT
 GCACGATTTGTTGGCGAATACATAAACAGCCCTGCCGTCCGTACCGATTACACCTATCCACGTTACGAAACCAACCGCTGAAACAA
 CTAAGCGGTTGAGCAGGTTAACCATTCTTTATCTACACTTAATGCCCTGACTCTCGCGCACCAATCAGAGGTGAGCGG
 AAGTAAAAGCAGTCTGGCTTAAATATTGGCGGATGGGGATTTATCGAAATGAAACCTTGCAGCTAACCCCGCGCACACTGCC
 TTTCTTTCCCACTTGTACAGACCGTATTTTCTGCGCGGCATAGACGTTGTTTCTCCTGCCAATGCCGATACCGATGTGTTTA
 TTAACAICGAGCTATTGGAACGATACGCAACAGAACGAAATGCACTATAACAATGCCGAAACACTGAAAGCCCAACAAACT
 GGAATATTTGCGAGTAGACAGAACCAATAAAAAATTTGCTCATCAAAACAAAACCAATGCGTTTGAAGCTGCCTATAAAGAAAA
 TACGATTTGGATGGGACCGTATAAAGTAAGCAAGGAAATTAACCCGACAGAAAGGATTAATGTTGATTTCTCCGATATCAAC
 CATACGGCAATCATATGGTAACTCTGCCCATCCGTAGAGGCTGATAACAGTCATGAGGGGTATGGATACAGCGATGAAGCAGT
 GCGACGATAGACAAGGGCAACCTTGA

SEQ ID NO: 4

MQARLLIPILFSVFIILSACGLTGIPISSHGGKRFAVEQELVAASARAAYKMDMLQALHGRKVALYIATMGDQGSGLTGGRYSID
 ALIRGEYINSPAVRDTYTPRYETTAETSSGLTGLTSLSTLNAPALSRTOQSDGSGSKSLNLNIGMGDYRNETLTTNPRDTA
 FLSHLVQTVFLLRGIDVVSPANADTDVFINIDVFTIRNRTEMLHLYNAEFLKQATKLEYFAVDRNKKLLIKPKTNAFEAAAYKEN
 YALWMPYKVKSGIKPTEGLMVDSDIIPYGNHMGNSAPSVYADNSHEGYGYSDAIVRRHRQQP

SEQ ID NO: 5

GTGCATTAATTAACITTTTCAATGGAATATAAAGAAATTAATGAAATTTATATATAAAAAGCCGTTTATTTTTAAAAAGCTTTAG
 ACGTAAGITCTATCTCATGGAAGAAATTAATGAGTTGATCAACTGCACGCCCTACTGATTGGCAATTTAAATTTGCAAAAGG
 GGAATAATCCCAAAGAACCTATGTTGAATCATTCAATGATGTGGGTAGAATACGTCACTCGGTTTATAAAAACAGCCGATAT
 CAGTATCTGCAAGATGGTCAACAATGGTTATAACCGTATTGATAATGAGCCATTTGTTGACAGTATTGCTAAAACAAATTTGCC
 AGTTTGTCAAGCAACAACTGTTGTTAGTGGTTATTTGGCGTTGGTATTTCCATCATATCGTAATCATTGGGCAACTCGTGA
 TGTGTTGGCGTTCAATGATTGGCAAAAACACTGGACAGTTTCTGCTCCAAATTTGATATGCCAATTTGATATGCAAGCAAGCA
 AAAGATATGCCGACATTAACCTCGGCTACAAACAGTAGATATGGAAGTCATTTAGAAGCAGGAGATATTTATATATATCCCGGTG
 GTTGGTGGCATAATCCTATGCCTATGAATGTGAAACATTCATTTAGCTATTGGAATTTCCGCCAAATGGTTATTAATATAT
 GGAATGGTTGATGAAAAAATACCTGATATCCAGAGTATTCGTAATTTTATAGACTGGGAACACGACCAAAAAAATATAGAT
 AATGCCGCTCAGGCAGTTACTGAAATGATGAAGAACAAGAAAAATTAACAAGCATTACACAAGATTTTTGGGAAATCAACGCG
 TAAATACTGCATTTAATATGCAAAATTTTTGGAAATCTGGAATAATGACAGATTACCTGAAATAGTACAATTAACCTCAATTCGTT
 GGATAACCGTACAATTAACCAAGGATATATCATAGCAATGGCATCAAAACCAACTGGACAATGATAGCCAAACGATTTTGCAA
 TGGATTGCTGACAAACATAGTGTAAAGTTAACTCAACTGTATGAATTTTCCCAAAATCAAAACATTAATTTGGAAAAGTGGAAA
 AATTAGTGTGATTTGACAATGATCGATGATTTGGAATGTTAACTGATGAAAGATA

SEQ ID NO: 6

MHINFSEYKEFNENLYKPPFIKALDVSSISWKEINELYQRADPTDQWFKFRKGEIIPKEAYVESFNDVGRIRHRFNKTAVY
 QYLQDGATMVYNRIDNEPFDVSIKQIAQFAQAQTVVSGYLAFGSSPSYRNHWDRDVFVQLIGKHHWTVSAPNFDMPLYMQQA
 KDMPHITPPTTVDMVEVLEAGDILYIPRGWHPNMPMNCFTFLAIGTFPPNGYNYMEWLMKKIPDIQSIQRFIDWEHDQKNID
 NAAQAVTEMMKNQENYQAFIQDFLGNQRVNFAFNMQIFGNLDNRLPENSTIKLNSLDNRITIKQGYIIANGIKTNLNDSDQITLQ
 WIADKHSVLTQLYEFQCNQINLEKVEKLVDFLTMIDVLECLTDER

SEQ ID NO: 7

ATGACATATTTCCGCCCAAGTATTCCAAGCTAAAAAATCGCTGGACAGGGCAAAATGTTTTAGTGCCTCTTTTCTCTTGC
 AGTTTTAAACATTTTTCCGCGTCTATTAGCGGCTATTTTAGTGTCTTTTTAATTTTTGGGAGCTATACCAATAAAACACCGT
 AACAGGACAACTTTTGCCGCAACAGGGGTAGTACGCGTATATTCGAGGATATGGGGTTATTGCTCATCAACAGTTATGAAT

GCGGATTTTGTGAAAAAAGCGGATGTGTTATTTACATTATCTACTTCACGAAATGACAATAATGGCAGCATAACAAGCAGGATTAC
TCGCCAAGCTGAATTAATAAAATCACTATCAGAACAGAAATTTATTAAGAAAAACGTGTTTCATGCAGCAGAAAAAACCGCACA
AGAAAAACCGGTACACCGTTTTCAGAACCAATGCAACACGTTAGAAATCAAAATTCATGCAAGAAAAACGTTATGCGATTTCT
GAAAAATGTTAGAAAAACAGCGTTATTTAGCCAAAATGGACGCAATTTCTGAATAGAAAAAATAGTTATGAAATGCTTTAT
TGGAGCTGAAAGCAGGTTTACCCGCTTACCAACGAGAAGCAGATAATCTTGCTCGGAAATAACTGTACAGCAAGCAATCTAAA
AAATCTGCCTGAACAGCAAGCTACTGAAATCAGTCAATTAGAACCGCGAGTATCTGTCTATCAACAAGAAATATTGGATTATCAA
CAACGTAAACGAAACAACTATTCGCGCAACCATATCAGGCTATGTCAGTTCTATTAACACGGAATTTGGTCAGCAGGTAGATACTA
ACAACTATTGATGAGTATCGTCCGGAAGAAAGTGAATTAAGTAAATTTATATGTCCTAGTCGCGCAATCGGTTTTGTA
ACCAATGACAAGGTAATTTGCGTTATCAAGCATATCCTTATCAAAAATTCGGACACGAGAAAGGACACGTCATTTCTATTGCT
CAACCGGCACTAGGGCGCAAGAAATGGACAAATTTAGGCAATTTTTACTCAACAGCACAAAGTGAATGAGCCTGTTTATTAA
TCAAAGTAAAACGGATAATCAGCATATTCGATATATATGGCACGGAGAAGAAATGCAAATTTGGTATGATACTAGAAGCCGAT
TTGTCATGAGAATAAACGGTTATATGAATGGATACTTGACCCCTTGTATCAAGTAATGGGAAAAATATCATAG

SEQ ID NO: 8

MTLFRPEVFAQKKNRWTGQIVLVRPFLQFLTFFAVSLAAILVAFLIFGSYNTKTTVTGQLLPTTGVVVRVYSQDMGVIAHQHVMN
GDFVKKGDVLFLLSTRNDNNGSIQARLLAEELKKSLEQEIIMKRVHAAEKTAQENTVHRFQNMQHRVNIIMQEKRIAS
EKMLEKQRYLAKMDAISLEKNSYIEALLEKAGLAAYQREADNLAREITVQSNLKNLPEQQATEISQLERAVSVYQIILDYQ
QRNEQTRATISGYVSSNTEIGQQVDNKLMSIVPKESELLANLYVPSRAITGFVKPNDKVLIRYQAYPYQKFGHAEHGVISIA
QALGAQEWTLNGNIFQTQAVNEPVYLIKVKLDNQHIRIYGTAKLQIGMILEADILHENKRLYEWILDPLVQVMGKIS

SEQ ID NO: 9

TTGGCGTGTGGTATCGGTATGGGGTTTCATGGTTTTATACGGATTTGCGCCATTAAGGGCGCGTTTTCTTTGTCTCTTA
AAGCGCAACATTTGGCAGATTTGGTACGGTTTGTCTAATAGCATGAATTTAAGTGTCTGAGCTGTTCCGGTTGGAATAGATGAAT
GGTAAATTTGCGCTTACCCTGTATTTACATTTGGGATTTAAATCATTTTGGTGTGTCATGAAGTTCATCGTAATCACATTGTG
ATTATGAACCTGCTTGGGTTGGCAAAAAGTGCATGGAAGAAAGTCTCGGCTGTTTACAGGCATAGCCTGGAGCTATTCC
CTAACACGCAATTTGAAGGAAACGAAACCGGTAATAATTCGCATACTCCGATGTTGCGCGGTGGTTGGATTAAAAAGGTC
TTTATTTGAGCTTCTGCTGCTGGCGGAGTATTACAGGTGTTGCTTTGGCAAGCCCTTTTTTATGCAATGGGTAATAGCAT
GCCATTTGATCGGACCGTGAATTTGCTAACATTTGGCATTGGCTTTGGTGTGTTGTTGATGATTGTCAGCAGTTGGTGCAT
TGTGCAAACTTTGGCAGGTATGATTTGCTACTTCAATGAATTTCAATGGAAAGCCAACTGATTTCCGCGCTTAAATGGATTT
ACCTGTTTCTTTTCACTAAACGACACTTTGGGCGATGTTGTTTACGTTTTGGCTCGGTGGATAGTATTCAAAGTACGCTCACT
TCTACTTTCTTTGTGCTGACTCAACAGCATTATGGCAGTATTACATAGGGTTAATGATTTATACAGCCCTAGTTTAAACGG
CTGTTGATTTGGTGTGTTTACTGATTTATATCGGTATTCGTTGGTGGCATATTATCCGTTACGCCATGCAACCGAAGAAATAT
TGTTCTGCTGCCAAGCAAGCAGTATTTCATGGAACCAATACCGGTTATCAAACCGTTAAATTTGATAAAAATGCAACA
CGTCAATGCGGTTGGCTGAACTTATTTGTTGATACAAATCAACACAGTTTAAACGACACAAAAACTTTCCGCAATGTTCCGTTTTG
CGAACGCGCTTTTGTGGTATCGCTAATATTTAATTTGTTGTTTGGTGGGTTCTGTGTTGGACGGTGTGTTTACCGTTGG
GGCGTTAATGGCGTTTATGCTTACAAAAGTCAAGTTGAGAGTAAAGCAGTTTATTGATTGACAGTTTGTACAAATCAAAATG
CTAGGATTTGACGCGGAAACGCTTGGCGGATTTGCTAGAAAAACGGAACGAGCAGAACCAAGTTCAGGCTGCCTAAAA
ATTTAGAACACTTTGATATTAATAAGAAAAATATATCATTTCCGATGCTGAAAAATGAACCATACATTTTGAAGATTTCAAGTT
AATTATCAAAACAGGAACAGCCCTTTTCTTTGCTGGACATTTAGGCTGTTGGGAAATCTACGCTCATTCAAATTTGACAGGCA
CTGAAACCTGAAAGCGGTCAAGTGTACTAGGCGAGAATGATTTATGCTGTTTCCACCTGCATTTATCCGCAATGGAGCGCAA
GTGTAATGCAAGATGATGTTATTTGCTGGCTCAATCGCAGAAAAACATCAGTTTTTTGATGATACGCCCAATCAAAAT
CGCATTTTGGCTCAAATGGCGAATTTATCATGAAATTTGCTGATGCTTATGGCGTATGAAACCTTAAATGGTACATGGGA
AGCGCATTTGAGCGGTGGGCAAAAGCAGCGTATTTGCTAGCTGCTGTTTATATCGTGAACCCAAATTTTGTGTTTATAGACGA
GCAAGTACGCTTTGCAATCAACCAATGAACGATGATTAACGATAATCTACGCCAACTCAACATTACCAAAATCATGGTGC
TCGTCAGGAACTTTAAATACTGCTGATAGCGTATTTATTTAGACAAGGTAGCCTGA

SEQ ID NO: 10

MACLVSVLGFHGFYDLRHLRFRFSLSLKGLADLVRFANSMLNLTARAVRLDLDELVNLRLPCILHWDLNFVVLHEVHRNHIV
IMNPASGWQKVRMEEVSRCTGIALELFPNTQFEKHETRKIRILPMLRGVGLKRSFLQLLLAAVLQVVALASPFMQWVIDH
AIVSADRDILLTLALGFLMLIVQQLVSLQWAGMYLSTSLNIQWKANVFRRLMDLPVSYFTKRHLGDVVSFRGSDVSIQSTLT
STFFVLVLSMAYFTLNGMYLSPSLTAVVLLVLLIYIGIRWVAYPLRHATEENIVHAAKQSSYFMETIRGIQTVKLFDKNAQ
RHAAWLNLFDVTINTGLTQKLSAMFGFANSLFGIANILIVFYGAVSVLDGVFTVGALMAFMSYKQFESKAGSLIDQFVQIKM
LGLHAERLADIVLEKTEENQNDLRLPKNLEHFDIKIENISFRYAENEPYIQLDFSLIHKQGTIAIFAGHSKGGKSTLIQILTGS
LKPESGHVLLGENDIHAFFPAFIRKWSASVMQDDVLFAGSIAENISFFDDTPNMEKIAFCAQMANIHHEIVAMPAYETLIGDMG
SALSGGQKQRIVLARALYREPKILFLDEASSHLIDINNEVINDNLRQLNITKIMVAHRQETLNTADSVVYLDKVA

SEQ ID NO: 11

ATGAAAAACAAAGCCCTGATGATTGCCGCCGATGCTGGCAGCAGCCGATGTCCACGCCGACCCGAAAAGGCAAAAACCGCAT
CCGCCAAAGCTGCCAAAGCTGCCAAAGCTGCCAAAGCTGCCAAAGCTGCCAAAGCTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCC
AAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAAGTTGCCAAA
GTTGCCAAAGTTGCCGCCACGGCGCAAAAGAACGCCACCCGACAAACAGCAGGGCGGATCCGCTTTTCAGACGGCATTTGCCG
CCGTTGCCGACAAAGCAAGTCAACGCGCCGCGGCTTCCGGAAGCCGTTGCCGAAGCCAAAGCCAACTGCCAAAGACCGCGCA
GATAAGCGAATCCGAGCTGTCCGACAGGTGCTGATGACGCTTGTCAACCAATCCCTGATTGTACAGGCGGGCAAAACCGCGCAAC
ATTCAAGCAAGCGAAGCGGAAATCGATGCCGTGCTGCCAAAAAATCCGCGCTCAAACCTCAGCCCGCCCAACCGCGCGGATT
TTGCCGCAACATCATTTGCCGAAAAGTCCGCGCAGCAGCGAGTGTGATGACAGAACAGCCGAGTGAGCGAAGCTGAAATCGAT
CCTCGAGCAGGCGCAAAAACAAGGCATCACCTGCCGAAAGGCGCACCGTTGCGCCAAATACCAGCGCCCAACACATCCTGATTA
GCCCAGCAGCAAAAACCGCGCGTCCGCGCGGAAAGCACCATCCGCAAAATCTACGGAGAGGCCCGCAGCGGACAGACTTTTCCA
GCCTAGCGCGCAATATTCCGAAGACGCGAGCGCGGCAACCGCGGAGATTGGGCTGGTTGCCCGCGCGGTGATGGTTCCCGC
CTTTGAAGAAGCGCTCACCGCTCAAACCCGGACAGGTCCGCGCANNNNACACACTTAA

SEQ ID NO: 12

MKIKALMIAAALLAAADVHAAPQKAKTASAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAKAAK
VAKVAATAQKEAAPAQQGGIRFSDGIAAVADNEVITRRRLAEAVAEAKANLPKDAQISELSRQVLMQLVNQSLIVQAGKRRN
IQASEAEIDAVVAKNPKLNLSPAQRDFADNIIAEKVRQQAQVMSRVSEAEIDAFLEQAQKQGITLPEGAPLRQYRAQHILIK
ADSENAAVGAESTIRKIYGEARSGTDFSSLARQYSQDASAGNGDGLWFDAGVMVPAFEEAVHALKPGQVGAAXHT

SEQ ID NO: 13

TTGGTCAAAAACGGCAGACGGCTACAAAGCTATTGCCCGTATCCGAGCCGGCAGAGCGTCTCTCAAAGGACGAGGCAAGCGGAA
AAATGGGATACAAAACCGTTACCGCCGATACCGCAATCCGATGAAGAAACCGTTTACATTAAGTTTTCAGACGGCATCGGCA
CAGCCAAACCTGATTTCCAACCGCATCCACTCGTTTTATTCGGCGGCAAAATGGATTAAAGCAGGAAATTTGAAAGCGGGAATC
AGGCTGTTATCCGAAAAGCGCAAAAACCAAACCGTCCGCAACATCGTTGTCAAACCAAAACCGCTCAAAGCTCAAACTCGACCG
TTGCTGATGCATACCTACTTCCGCAAGGCGATCAGCGGAAACCGAAGGGTTTGGGTTTCAATAATGCGTCTCCGCTAAAAG

TAGCTTGAGTGTGCGTCAAAGTAATCTGCAATCCAGTGGCGGTAATGTGCAGATGAGCGGACTAAAGGCAACATCAGCTTGAAT
CAGAGTTGGATAAATGCCAGTCAAGTAATTTGATACGGCAGCACTTCAAGGCAATATTTTTCAGACGGTCTGACTGCTGTTGCCG
AAGTTGGACCGGTATCTCTTCTTCCAAACGGCAATGTTGATTTCAACGGTCTAAATACCTTGATTGCGGAAGGAGACATTAATGC
TGGTTTCAGTGGTAAAGGCGCTGTAAGGAAAGGCAATACCGATATTTATGCTTCTGACGGGATGTAAACTGGTTCGCCGGAGT
CAATTAGATTTAGGCAACGGCACCGTTAACGGCGGTATGTCAGCTTAGACAGCAATAAAGGCAAGTGGTGGTGCAGAAATGATAC
ACCTGAATGCCCGGCATCACTGAAAGTGGATGCGGATCAAACTTTAACCAATTAATAACAGCAAGCTCAATTCGGATCACAATAAC
CCAGATTAATACAAATCATGGTATATGACGCTTAATCAGCTTGTGATGCTCATTACAGTGTGATGATGATCAGTGCACAGGGT
AAAGGCAAGGAAAGACAGCGGTCAAATTTTACAAAACGACAGCAAAACAGTAAAGTACTTTGGCGGACAGATGGTGTATTGT
CATTGAACAGCAGCGCATACAGTGTGGCAATACTACCTGCGCGGTGGCGGATAAACATCAAAGCTGGCGGAGGCATCAT
CAAACGGGGCATATCGATTTGGGAAACCAAGTACGGCAACTGCGTTACAGCGAACTGAAACCACTGTCCGGTATGATGTCG
ATAGAATCGGGCGGTGATAACCCCTTTAACCGTTGAACCGGTAACCGTATCGTTTCTGCAGGTGATTTGGCCGTGAACCAACAG
GCACATTCAAAATCAGTGGCCAGAGCAGGAAATAACGGCAATCCGAGTCCGAGTGGCGCAACAGCCAGCGTTTTCAGCTAAAGCAATCCG
GATTGTGGCGGAGAGGTGGATATTGATGACGCAATATTGCCGAGGAAAGATTTGGCTTATGTTGCAACTAAAGGGAATATT
TCACTAAACAGTATTAGGAATACGTTTAGTAATCAATGAAAACAGATAAACACAATATCACACAGCAACTACTGATGTAG
AGCAGGAGCTTAGTAAAGTTAACGAGCGATCCTAAATATCGTAAGCAAGGATCTTTCAAAATGTTGAGGCGTAAATAAAGAG
GCGTGACAAAGTATTCGGTGTAGTGTAGGCGGATACGTTGTTTACGAGCAAGATAAATGCTGCTGATGAGGCTTGGGCGAGAA
GCACATTCAAAATCAGTGGCCAGAGCAAGTAAACCAATTTTCAACAAGCCCTTTTAAACAGTTTCCAAACCGGTTAGTGTCC
ACGAAAAACAAGGTAGCACCTTAGCGGTCAAATATCAAATGCTCGTGCAGGTGGTATCCGTATCCAGGGCTCAAAGTTGC
TGCCACCCAAACAGGCAAAATCCAAGCTGCAGGATTTTACCTGCACAGCCGACAGGAACTGCAAGAAAGTTCAGCTTCAATCA
GCCATTTGACATCAGCGGTGACTTGTATCTTTGAATGATGTCAGCAAGGACGCGATAAATACGGCTATGCCATTTTCAGCAGGC
CATCTGAAATTTTCAGGAAAAACAGGTGTTACCCTCTCTGCACAAATGCCAATGAGAACAGCCGCGATCAGTCTGAGTGGCGGAAA
CATCGAAGTGTAAACCGGCAAAATCAAATCAATCCTACCGCGCACTTCTATTACCGCGGACAGGCGCAACTCTATACCTTT
GAACGCGGACGCTACAAAACCGGCAAAATGGTACAACCGCAACACATTACCGAAGTCAAAGAACCAAAAAACGCAAGCCGACG
CAGTAAACCTCAGCGCATCCAAAGGCATCGACATCAATCTGGTGGCAGCATCGACGCTACGCCACCGCATTCGATGCCCCCAA
AGGCAGCATTAAACATCGAAGCCGGCGGAAATGACACTCTATGCGGTAGAAGAGCTCAACTACGACAAATAGACAGCCAAAAA
AGGCGCAGATTTCTCGGCATCAGCTACAGCAAGCACACGACACCCACCAAGTATGAAACCGCGCTGCCCTCAAGGTTAG
TTGCAAGTTCAGCCAACTGCAATCAGTTGGGATACCAAATGCAAGGCAACAGTTTGAACCACTACTGGTGGCGGCAACT
ACGCGCAGGCGTAGGCGAGCAGGCACGGGCGAGTGCAGATTATCCTCGAAGGATCAAAGCAGCATCCACACAGAAACCGTG
AGCAGCAGCAATCTACTCTATGGCAAAACAGGCGAGGCGGGCAGTAAACATCGAAACCTTGAATTCGCGAGTTTACCGGCTC
CCGTTGCGCCGCTACTGTCCGACCCGCGGTTACATTTGCGATATTTCCCAAGGCAATCGAAAAACCAATCGAAACCTCAC
CAAACAGCCGAGTATGCTTATTTGAACAACCTCAAGTTGCGAAAAACATCAACTGGAATCAGTGCAGCTTCTACGATAAAT
GGACTACAAACAGGANNNNNCACACTTAA

SEQ ID NO: 20

MXXXPNGRGLSHNRYTQFDVNDKGAFLNDRNNPFLVKGSQALILNEVRGTASKLNGIVTVGGQKADVIANPNGITVNGGGFK
NVRGILTTGTPQIGKDGALTFDVRQGLTVGTSWVNDKGGADYEVLEVARAVALQGLQKLNVAVSTIAQKVDYASGEISAGTA
AGTKPTVALDAAALGGMYADSIILIANEKGVGVKNAGTLEAGKLVITVSSGRIENSGRIATTDADTEASPTLSIETTEKGAAGT
FISNGGRIEESKLLVIFETGEDISLRNAGVAVQNNRSRPTTVLNAGHNLVIESKTNVNNAKGSANLSAGGRTTINDATIQAQSVY
SSTKGDDELGENRIIAENVTVLNSGSIASAIVI EAKDTHIEAGKPLSLETSNVASNIRLNNNGSIKGGKQVVLMADDDIQAQAS
HLNASGNLYIHAGKDLDLNADKDLSTQISLRLADNTALISSNGNTLTAEKNLDIQAQSLSVRQSNLQSSGGNVQMSATKGNISLN
QSWINASQNDITAAALQGNISDGLTAVAEVGRVSLLANGVDFNGLNLTIAEGDINAGSVGKRLKMDNTDIYASAGDVKLVAGG
QLDLNGTVNGGHSLSLNSKGSVQVQVHNLNARASLKVADQTLTINNSKLSNDHNTQINTNHGHMTLNQIDAHSRRHMSISAQG
KRGKDSGQILQNDQNSKSTL AADGVL SLNSALQVLDNTTLRGGAINIKAGGGIKRGRHIDWETQDATTMRSAELKPLSGMMS
IESGGDNLPTVEPNRIVSAGDLAVNHNGTFQISARAGNNGNPSAQTASVSAKGNIGIVAGEVDIDAANIAGKDLALVATKGNIS
RLNSIRNFTFSNYQLKTDKHNTIQQLTDVEQELSKLSDPKYRKAQDLQSQLRKYRKRDRKVFQDSEARLRLRAKINAADAEAWAE
RQSPVKALLQQLLTVSQPQSGHENQSTLSGQNIKLTLAAGGIRIQGSKVAATQQANIQAAGFLPAPAAEELQEGRLQS
AIDISGLVDTFEYQQQSDKYGYAIFSRPSEISGGTGVTL SAPNANENSRI SL SAANIEAENGKIKIQSYDQSYAGGTYLTF
ERRSYKTKWYNRKHITTEVKEHKNAPDAVNL SASQGDIDKSGSSIDAYATAFDAPKGSINIEAGRKLTL YAVEELNYDKLDSQK
RRRFLGISYSKAHDTTQVMKTLALPSRVVAESANLQSGWDTKLQGTQFETLLGGATIRAGVGEQARADAKIILEGKSSIHETV
SSSKSTLWQKQAGRGSNIETLQLPSFTGPVAPVL SAPGGYIVDIPQGNLKTQIETLTKQPEYAYLNNFKLRKTSTGITSLLTIN
GTTNRXXHT

SEQ ID NO: 21

GTGNNNNNGCAAATCAAACCTGCAATCCTACGGCGACCAATCCTATTACGCCGGACAGGGCGAACTCTATACCTTTGAACGCCGCA
GCTACAAAACCGGCAAAATGGTACAACCGCAAAACACATTACCGAAGTCAAAGAACCAAAAAACGCCAAGCCGACGCGAGTAAACCT
CAGCGCATCCCAAGGCATCGACATCAAATCTGGTGGCAGCATCGACGCTACGCCACCGCATTTCGATGCCCCAAAGGCAGCATT
AACATCGAAGCCGGGCGGAAATGACACTCTATGCGGTAGAAGAGCTCAACTACGACAAACTAGACAGCCAAAAAGGCGCAGAT
TTCTCGGCATCAGCTACGCAAAAGCACAGCACACCCACCAAGTCAATGAAAACCGCGCTGCCCTCAAGGGTATGTCAGAAATC
AGCCAATCTGCAATCAGGTTGGGATACCAAATGCAAGGCACACAGTTTGAACCACTGGGTGGCGCAACCATACGCGCAGGC
GTAGGCGAGCAGGCACGGGCGAGTGCAGATTAATCCTCGAAGGGATCAAAGCAGCATCCACACAGAAACCGTGAGCAGCAGCA
AATCTACTCTATGCGAAAAACAGGCAAGCAGGCGGCGAGTAAACATCGAAACCTTGAATTTGCCGAGTTTACCGGTTCCGCTTGC
CGTACTGTCCGACCCCGCGGTTACATTTGTCGATATTTCCCAAGGCAATCTGAAAACCAAAATCGAAACCTTCAACAAACAGCC
GAGTATGCTTATTTGAAACAACCTCAAGTTGCGAAAAACATCAACTGGAATCAGGTGCAGCTTGTTCAGATAAATGGGACTACA
AACAGGAGGGCTTAACCGAAGCAGGTGCGGCGATTATCGCACTGGCCGTTACCGTGGTCACTCAGGCGCAGGAACCGGAGCCGT
ATTTGGGATTAACCGGTGCGGCCCGCCGCAACCGATGCAGCATTTGCCCTTTTGGCCAGCCAGGCTTCCGTATCGTTTATCAAC
AACAAAAGGCAATATCGGTAACACCTGAAAGAGCTGGGCAAGAGCAGCAGCGTGAATAAATCTGGTGGTTGCCGCGCTACCGCAG
GCGTATCCAAACAAATCGGTGCTTCTCCCTTGCACCTTGGAGCGAAACCCCTGGGTAACAACCTCAAGTCAACTTGGCCAA
TGCGGCGAGTGGCGCATGATTAATACCGCTGTCAACGGCGGACGCTGAAAGCAATCTGGAAGCGAATCTTCTGCGGCTTTG
GTCAATACCGCGCATGGAGAAGCAGCCAGTAAAAATCAAACAGTTGGATCAGCACTACATAGTCCACAAGATTGCCATGCCATAG
CGGGCTGTGCGGCGAGCGCGGCAATGAGGGCAAGTGTGAGGATGGTGCATAGGCGCTGCAGTGGTGGATTTGTTGGTGGAGC
TTTTGTTAAGAAATACCGATTTTACGGGTATGACTGCTTCTGAAATGAAAAAGCTAAAGCGAATATTACTGCGTATGCAAAATG
GTAGCCGGAGCGACTGTAGGTGTACAGGAGGCAATGTTGATGTGGCGGCAAAATGCTTCCGAAACAGCTGTTAAAAATAATGCAT
TAGATAATTTTGGGATATCGGCAACCTCGTATGGGACGGCGGTAATGATTTACGCCAAATCTATTGGCGATAAGCAGATGGC
TGAGAAGCGGGGATTTGTTTGGTGGGATGCGCGCCGACGCTGCCGTTCCCTTTGTTCCGGCAGTGCAGTAAATACGCCGA
GGCGGGCTTATGTTCTGAAGGCGGGAGACGAAGCAGTGTGATACCGCTAAAGCCATACAGGAAATTCAGAAAGCAGACCGGAAATCA
AGCTTACTTATGATAAGGTTAAAGGTTTGGACAACACCGCGGGGTAGATTTAGATGCTAAGCATGTAAGGATTAAGGAT
TAAACATGTTTTAGCCATACAATCAAATCAAACAAACCTGTTCACTTCTGTTTTAATGTGTCCCGTAAAGAAATTTGCCCT
TTGGTTGATGAAGCTTGGAGAAATGAAAGGAAATCCTTTGCCAAATGATTCATCCGTATATCTTGTAGATGAAAGAAACCTATTG
GAACAAAAGGAGAAACAAAAGTGGCGATTGTTGCAAAAAGGAAACAAATAAAATCAATTTCTGCATATCTCAGAAATAA

SEQ ID NO: 22

MXXQIKLQSYGDQSYYYAGQGLYTFERRSYKTKGWYRKHITVEVKEHKNKPAVNLSASQGDIDKSGGSDAYATAFDAPKGSINIEAGRKLTLAYEELNYDKLDSQKRRRFLGISYSKAHDTTQVMKTLPSRVVAESANLQSGWDTKLQGTQFETTLGGATIRAGVGEQARADAKIILEGKSSIHTEVSSSKSTLWQKQAGRGSNIETLQLPSFTGPVAPVLSAPGGYIVDIPQGNLKTQIETLTKQPEYAYLKLQVAKNINWQVQLAYDKWDYKQEGLEAGAAITLAVTVVTSAGTGAVALGLNGAAAAATDAAFASLASQASVSFINNKGNIENLTKELGRSSTVKNLVVAAAATAGVSNKIGASSLATWSETPWVNNLNVNLANAGSAAALINTAVNGGSLKDNLEANILAAALVNTAHGAAASKIKQLDQHYIVHKIAHAIAGCAAAAANKKQDGAIGAAGVEIVGEALVKNTDFSGMTASETEKAKANITAYAKLVAGATVGVGTGGNVDAANAASETAVKNNALDIWIDIGNLVWDGGKWIYAKSIGDKQMAREAAIDFGVDAAAAAVFPVPAGATKISRGGAYVLKAGDEAVDTAKAIQEIQKQTGIKLYTDKVNKVVTTAGLDYGLDPAKHGRIKHVLAHTIPNPNKPVHSVFNVSRRKEVLPVDEAWRMKGNPLPNDSSVYLVDMKKPIGKGETKVRIVVQKGTNKIISAYPQK

SEQ ID NO: 23

GTGAGCATTAGCGCACCGTATGCCAATGAGAACAGTCGCATCCTGCTCAGCACCACGGATATCAGTTCGGAACCGGCAAAATCAAAATCAATCTTACGGTGACCAATATTAATACTGCGAGACAGAGCGAACTCTATACCTTTGAACGCCGAGCTACAAAACCGGCAAATGGTACAACCGCAACACATTAACCGAAGTCAAAGAACACAAAACCGCAAGCCGACGCGAGTAACCTCAGCGCATCCCAAGGCATCGACATCAAAATCGGTGGCAGCATCGACGGCTACGCCACCGCATTCGATGCCCCAAAGGCGAGCATTAAACATCGAAGCCGGCGGAAATGACACTCTATCCCGTAGAAGAGCTCAACTACGCAAACTAGACAGCCAAAAGGCGCAGATTCTCGGCATCAGCTACAGCAAAAGCACACGACCCCAAGTCAAAAAACCGCGCTGCCCTCAAGGGTAGTTGCAGAATCAGCAACCTCAAACTCGGGTGGGATACCAACTGCAAGGCACACAGTTTGAACCACTGGTGGCGCAACCATACGCGCAGGCGTAGGTGAGCAGGCACGGGCAGATGCCAAGATTCTCGAAGGCATCAAAACCGCATCCACAACGAAACCGTGGCAGCAGCAAACTGCTCTATGGCAAAAACAGGACGGACGGGGCAGTAACATCGAAACCTTGCAATTCGCGAGTTTACCGGTCCTGTTGGCCGCTACTGTCGCAACCCGGCGTTACTGTCGACATCCCAAGGCAATCTGAAAACCGAAATCGAAAAGCTGGCCAAACAGCCCGAATACCGCTACCTGAACAGCTTCAGACGCAAGAACGCTCGATTGGAAACAGGTGCGAGTGGCTACGACAATGGGACTATAAAACAGGAAGCTTACCGGAGCCGGTGCAGCGATTGTGGTATTATTGTAACCGCTTAACCTTATGGATACGGAGCGCTGCAGCGGTAGCGTAACCTGCCGCAGGAAGTAGTACAGCCGCGCTGCAACACGACAGCGGCAAGTACCACTGTTTCTACTGCAGCTGCCATGCAACCCGACGCTTACTAGCTTGTATAGCCAAAGCAGCTGTAGCATCATCAATAAAGGCGGATGAGGCAAGCATTAAAAGATCTCGGCACAGTGATACGGTCAAGCAGATTGTTACTTCGCGACTACGCGGGTGCATTAACCCAGATGGCGCAGATATTGCCAAATGAAACAGCAAGGTAAGAACCAGAACTGTTAGCAGTACGGGCAATCAAACTTCCAACTTGGAGGAGATTGGCCAAACCGCAATCGCAATGCAATGCTCAGCTGGTATCAATACCGCGTTAACCGCGGAGCTTGAAGACAACTAGGCAATGCCGATTGGGAGCATTGGTTAATAGTCTCAAGGGGAAGCCGCGCAAAAATCAAAAACCTTCAAGCAGCATTATGTTGCCAAACAGTTCGCGCACTTGGCGGTTGTGTTAGCGAAGTGGTACAGGGAAATGTAAGATGGGCAATGGCGCAGCAGTTGGGAAATCGTAGCTGAATCCATATTTGGCGGAGAAACCTGCTACACTCAGCGATGCAAGAAAGCATAAAGTTATCAGTTACTCGAAGATTATTGGCGGAGCGTGGCGGCACTCAACCGCGGCGATGTAATACTGCGGCGAATCGGCTGAGGTAGCTGTGGTGAATAATGAAATTTGACAGTACCCCTACCAATGCAAAAAGCATCAACCGCAGAAAGCCGCAAAAACCGCACTGAAAATAATCATCAAAAGATTATGCTGTCACATGCGGCGAGTGGCAGTAACTCCGAGGATAAGGATGCTGCCATTTGGATAAGCAATATCCGTAATGGCATCAGCGCCGATTGTGATTACAGCTATGGAGTTTATGCAGCAGGTTGGACAGCTCCGCTGATCGGTACAGCTGGTAAAGCAGCTATCAGCACCCTGTATGGTAACTCTTGTGTTACTGTCTGTTACGCGAGGCTGCTGAAGCGGGCGCGGGAATCGTACGCGGTGCGGTAAACGGTAGGCAACCGCTTGGGAAGCTCCAGTAGGAGCGTTGCGAAAGCGAAGGGCGGTAAGCAAGCTGCTCTAAAGAAACATAAACAATTTGGCAAAATTTAGCAAAAGCAGAACAGCAGATTTTATTCCGTAATGGCAACCGGATACGCAACTGGATGATGGAAGACGGGATTTAAACAATAGAGTAAGGAAAGGAGCAGGCTTGTCTGATGCAAGTAATATTCCGATAACCAATTAACGGAAACCAATCAAACTTACAAAGCCATAAGCTTAAAGGGAGCACCCGTTTACAGCGCGTAAAGCGAACAGGAGATTTTGGCCTTTATCGGCAGATGACTGGCCGAATCCGAATTTAGAGTTTTGCTGACGGAAGATTAGCAAAATGGCATTATCAGTACTGGAGAAATGGGCGAGAAACAAAATTCATTAAGAAATTTTTCAAAAACAGAGAAATCAACTCAAGCAGCATGGACATTAGATTGCAGAAATCTCCATCTTATTAAAGGTACTAAATTTGGAGCTTAAATCCAATAA

SEQ ID NO: 24

MSISAPYANENSRLLSTTDIISSENGKIKIQSYGDQYRYARQSELYTFERRSYKTKGWYRKHITVEVKEHKNKPAVTLASQGDIDIKSGGSDAYATAFDAPKGSINIEAGRKLTLAYEELNYDKLDSQKRRRFLGISYSKAHDTTQVMKTLPSRVVAESANLQSGWDTKLQGTQFETTLGGATIRAGVGEQARADAKIILEGKSSIHTEVSSSKSALWQKQAGRGSNIETLQLPSFTGPVAPVLSAPGGYIVDIPQGNLKTIEIKLAKQPEYAYLKLQVAKNINWQVQLAYDKWDYKQEGLEAGAAIVVIVITALTYGGAAGASVTAAGSSTAAAATTTAAATTTVSTAAAMQTAALASLYSQAAVAIIINNKGDVKGALKDLGTSVTVKQIVTSALTAGALNQMAGDIAQLNSKVRTELFSSTGNQITIANLGGRLATNLSNAGTSAGINTAVNNGSLKDNLNGAALGALVNSFQGEAAASKIKTTFSDDYAKQFAHALAGCVSGLVQGGCKDGAIGAAVGEIVAESILGRRNPATLSDAEKHVYISYSKIIAGSVAALNGGDVNTAANAEEVAVVNNALNFDSTPTNAKHKHPQKPDKTALEKIIQSIIMPAHAAGAMTNPQDKDAATWISNIRNGITGPVITISYGVYAGVTAAPLIGTAGVNNALNFDGMANPSGCTVMVTOAAEAGAGIVTGAVTVGNAWEAPVGLSKAKAAKQAPKETINNLANLAKAEQIILFRIAQDRDQLDANKTGFNNRVRKAGLLDASNIPITINGKTIKPVQATISLKGAPVYSGVSEQEIFALYRQMTGQNPFRVLPDGRLANGIISTGEWAGTKIALRNFskTENSTQARWTLDLQNPSPFikgTKLELKFQ

SEQ ID NO: 25

TTGACACTCTATGCCGTAGAGAGCTCAACTACGCAAACTAGACAGCCAAAAGGCGCAGATTTCTCGGCATCAGCTACAGCAAAGCACACGACACCACCCCAAGTCAAGTCAAAAAACCGCGCTGCCCTCAAGGGTAGTTGCAGAAATCAGCCAAATCTGCAATCAGGTTGGATACCAAACTGCAAGGCACACAGTTTGAACCACTGGTGGCGCAACCATACGCGCAGGCGTAGGCGAGCAGGCACGGGCGATGCCAAGATTCTCGAAGGGATCAAAAAGCAGCATCCACACAGAAACCGTGGCAGCAGCAAACTACTCTATGGCAAAAACAGGACAGGACGGGGCAGTAACATCGAAACCTTGCAATTCGCGAGTTTACCGCGTCCCGTTGGCCCGCTACTGTCGCGCACCCGCGGTTACATTGTCGATTTCCCAAGGCAATCTGAAAACCCAACTGAAACCCCTCACCAAACAGCCGAGTAGTCTTATTGAAACAACTTCAAGTTGCGAAAAACATCAACTGGAATCAGGTGCGACTTGTCTACGATAAATGGGACTACAAAACAGGAGGGCTTAAACCGAAGCAGGTGCGGCGATTATCGCACTGGCCGTTACCGTGGTCACTCAGGCGCAGGAACCGGAGCCGTAATGGGATTAACCGGTGCGGCGCCGCGCAACCAGTGCAGCATTTGCCCTTTTGGCCAGCCAGGCTCCGTAATCGTTTATCAACAACAAAGGCAATATCGTTAACACCTGAAAAGAGCTGGGCAAGCAGCAGCGTGAAGAAATCTGGTGGTGGCCGCGTACCAGCAGGCGTAGCCGACAAAATCGGCGCTTCCGCACTGAACTGTACGCGATAAGCAGTGGATCAACAACCTGACCGTCAACTAGCCAAATGGGCGAGTCCGCACTGATTAATACCGCTGTCAACGGCGGCGCTGAAAGACAATCTGGAAGCAATATCTTGGCGCTTGGTCAATACCGCGCATGGAGAACGACGCAAGTAAAATCAAAACAGTTGGATCAGCAGTACATAGTCCACAAGATTGCCATGCCATAGCGGGCTGTGCGGCGAGCGGCGCGAATAAGGGTAAAGTGTAGGATGGTGGCAGTCCGTTGGCGGCTGTGGGCGAGATAGTGGGGAGGCTTTGACAAAACCGCAAAAATCTTGCCACTTTGACAGCTAAAGAACCGCAAGCAGATTTGGCATAACAGCAAACTGGTGGCGGTACGGTAAAGCGGTGTGGTGGCGGCGGATGTGAATACAGCGGCAATCGGCTAAAGTCCGATGAAATTAACCTTATCTCAAGAAGAGTATGCTCTAGAGAAAATTTGATCAAAAAGCCAAAAGGAAAGGCTAATTTCTTATAGATTGGGCGAGCCTGACCGAAACAGAGGCAAGGCGAGTTTATCTATTTGATTGAGAAAAGTTCGATTTCTAATCAATGCTTACCGGATATCAAAAATCCAAAGTAGTTTAAATAATCAAGAAAAAAATTTCTTGGATAATTTTAAACCAACCTCGAGGTAACACAGCTGGGCGACTCGATACTGAAAACCGCCAGTCAATGGGTATCTACTATTCCTTCAAAAGATTAATAACACCTTATCGAAAGCCTTCAAAACATTTAGTCTGTTATGATTCTTTGATTACAAATCAGCTGTGGCGCAACCTGCATTTACTTATAAACGGACCGCTTGGCTTCAAGTCAAGCAGCTACTGTGGCAGCAGGAGGA

TATAACATTGGACAGGGAGCGAAAGCAATCTCTAATGGAGAATCTGCATGGTACAGTTGAGTTGTTAATGGCACATTGATGG
TTGCAGGATCTGTATCTGCACAGGCTGCAATATCGGCCAAGCTGCACCTGTACCCGTTATCTGAGCAATGACAGTGCTCCTGC
TTTAAGACAAGCTTTAACTGCTGAAAGCCAGAGAATCCGCATGAAACTGCCGGAAGGATCGACAAATAGGGAATCTTGCAGATA
GCAAAAATTGATGTTAAAGGATTACCGCAAAGGATGGAAGCATTAGTTCTTCCAAAAAGGGGAACATGGATTTATTTGTTTAC
CTGAAACAAAAATTTTAAACCTATATCTGTTGATAAATATCATAATTTGCTCTCCTCTAGAGGAACATTAAGAAATATAGA
TGGAGATATAAATTACTTGAACCTATAGCACAGCAACTCGGAAATAATCGTAATGTATCAGGTAGAATTGATCTATTTACAGAA
TTAAAGCCCTGTCAATCTTGACGAATGTTATTTAGAGTTTGAAGATCGCTATCCAAATATTCAATTAATATTTTACAGGAA
AATAG

SEQ ID NO: 26

MTLYAVEELNYDKLDSQKRRRFLGISYSKAHDTTTTQVMKALPSRVVAESANLQSGWDTKLQGTQFETLLGGATIRAGVGEQARA
DAKIIILEGIKSSIHTEVSSSKSTLWQKQAGRGSNIETLQLPSTFGPVAPVLSAPGGYIVDIPQGNLKTQIETLTKQPEYAYLKQ
LQVAKNINWVQLAYDKWYKQEGLEAGAAIILAVTVVTSAGTGAVLGLNGAAAAATDAAFASLASQASVSFINNKGNIIGN
TLKELGRSSTVKNLVVAAATAGVADKIGASALNNVSDKQWINNLTVNLANAGSAALINTAVNGGSLKDNLEANILAAALVNTAHGE
AASKIKQLDQHYIVHKIAHAIGACAAAAANKGKQDQGAIGAAVGEIVGEALTNGKNPATLAKEREQILAYSKLVAQTVSGVVG
DVNTAANAAKVAIENLLSQEYALREKLKAKKGLSLDWGSLTEQEARQFIYLEKDRYSNQLLDRYQKNPSSLNNQEKNI
LAFYINQTSVGNATAWAASILKTPQSMGNLTIPSKDINNTLSKAYQTLTRYDSFDYKSAVAAPALYLLNGLPFGSVKAAATVAAGG
YNIQGGAKAISNGEYHLGTVQVVGNTLMVAGSVSAQAASAKPAPVTRYLSNDSAPALRQALTAESQIRIMKLPPEYRQIGNLAI
AKIDVKGLPQRMEAFSFKGEHGFISLPETKIFKPIVSDKYHNIASPPRGLRNLIDGEYKLETTIAQQLGNNRNVSGRIDLFT
LKACQCSNVILEFRNRYPNQLNIFTKG

SEQ ID NO: 27

TTGCTGGGTGTGGTCCGGGTATCGGTGAATCGATACAGGCTATAAAGTAGCGAAAGCGGCAAAAAATTTACAAGGCATGAAAA
AAGCCTTGGACAAGGCAGCAACCGTTGCCACTGCACAGGCTATGTGAGTAAAACCAAAATCAAAATCGGTCAAACCTGAATTAAG
GGTACTGCAGCAACTGACAAAACAAATGCTGAAAGCTATTTGGCGAAGGAAGGGACACGACAGGTAATAACGACGAGTATTT
GACTCTTTAGCTAAACAAAATGGCTTCAGAGTGTCTTTCGGGCGGCAAAATACGGCGGAAATACCGTTTTGATCATGTATGGCAGG
CTGCCATGGTGTGCTGTTTTGATTGTAGAAAAGTAAGCAGATTAGGAACGGTACGGTACAGCTGAATCCGAATGGTGTGGCGGTGG
ATATACGCAAAATGAGTGGAGGATTGGATTAGACAAGTTTATAGCTAAATACCCGATGGTAGTCCCGCTAAAGCTGCTCTTCAAA
GCAATAAAGAACGGCACATTAACAAACAGCAATAGCAGGCGTTGATCGTCAAAACAGGTAAGGCCGTTATTTCTCTGTCAAAGTTC
CTTCTAAAACCAATAAAGGAGATAA

SEQ ID NO: 28

MLGVVPGIGESIQAQYKAKAANKLQGMKKALDKAATVATAQGYVSKTKIKIGQTEL RVTAATDKQLLKAIGEDRDTTGKMTQLF
DSLAKQNGFRVLSGGKYGGNNGFDHVWQAADGVSVLIVESKQIRNGTVQLNPNAGAGGYTQMSSEWIRQVLDQLPDGSPAKAAVFK
ANKNGTLKTAIAGVDRQTGKAVILPVKVPSTNIRR

SEQ ID NO: 29

ATGAATGAGGGTGAAGTTGTTTTAACACCAGAAACAAATCCAAACCTTGGCTGGTTATGCTTCCGTGGCGATACCTATGGCGGT
GGCGTATTTGGCTAAATTTGGGTGACCGTTATGCGGATGATGCTGCTGCAATTTGCGGTAAGGATGCAAACTTAAATGGTTTGAA
TTTATGGATGAAAAAAGGTGTGGAAACCTATGGGATGATACGGTGGTAAAAAGACCCGTTTAGAGAAATTTGATCGGGTTGCA
TTGCAACATTTACGCCAATATGTAGATCTAATTAATGAAAATAATGGTAGATTACCTAACACTAGTGAATTTAGAGAAAGTTACT
ATAAAGCCGTTACCGAAAAATGGTGTCTTCTAGTGCAGCTATTTGATTTAGTTAATCGCTCACTCCGGATATGGCAGATGG
TTATTTGGGCAATAGGTTTTGGGGATAGAAGCCGAACGTATCCACAATGCGCAAGCAGTAATAAATCCGAACGGTACGCAATGGGAT
AATAGAAAGCAGTTAATATGCTTTAGATAAAGGATTTGATGGATCTTTTAAAGAGAAGCATTACTTTTTCAACTTGTGA
TAATGGATGTAACAAAGTTAGTGTGAATATACAATAGATGGTTGGCAAAAAATTTGGAGGTTGGGGTAAATGGGATAATCAATGA
TTTATATAAAGTGTGCAAAAAAGAGAGTGGACTGAAATTTTGAAGTCTTAATAAATCAACCAAGCAATTTAGAGATCTGTTTC
CCAAATCCGGAAGGCTGGATCGATGATGTTCCCAATGTTTCGCTGGTTAAAGAAAATAAAAAACGCAATGGCAAAATATC
ATGCTACGACCCCTCGCCCTAGATTTGGACGGAGCAGGTATAGAAAACCGTTCGCCACAAAGGCTTTGCAGGACGCTATTTGA
TCACACCAACCAAGGTTATCGCACCGCCACCGGTTCTGCGGATGACGGTTACTCGTCCGCGATTTGACGGCAACCGC
ATCATCGACAACGGTGGCAACTCTTCCGCGCAACACCAAACTGGCAGACGGTTCTTTGCCAAACACGGCTATGCAGCTTTGG
CCGAATTTGGATTTCAAACGGCGACAACATCATCAACGGCGCAGACGCGCATTCCAAACCCCTCGGTGATTTGGCAGGATCTCAATCA
GGACGGCATTCCCAAACCAACGAACTCCGCAACCTTGAAGATTTGGGTTAATCAATCTTTGGATCTCGCCTATAAAGATGTAAT
AAAAATCTCGGTAAACGGTAACACTTTGGCTCAGCAAGTAACTACACCAAAACCAACGGTACAACCGCAAAAAATGGGGGTTTAC
TTTTAGCAGCCGCAACTCTGCAACGCGCTTCAAAGCAAAAGTGAACACTACTGCCAAACAGGCAAAAGCCGCAATCTTGGCGG
CATTGGTCTGTCGCGGATTTGCGCGAAGCTGCCGATTTGTCGCGGATTTGGCCAAATGCTGAAAGCTTATTTCTGCCGCGAA
ACTAAAGAAGCACAGTTGGCATTGTTAGATAAATTTGATTCACAAATGGCGGAAACCGATTGCAACTGGGGCAAAAAATCGCCAA
TGGCAGCTTTCAACCGATTTGGACGCAACCGCTAATGAAGGATTTGCACTGACACCATCCAAAGTAGCACAACTAAAAAAGAACGC
TTTAGTTTCCCTTTCTGATAAAGCTAAAGCAGCTATTGACGCGCCCGGACCGCAATTTGCCGTGCTTGTGCTACACGGGGCAG
GATTCAGCACACTCTATTACATGAGCGAAGAAGACGCGCTTAATATCGTCAAAGTAACCAACGATACATACGACCATCTCGCCA
AAAAATCTACCAAAACCTGTTGTTCCAAACCGTTTGCAGCCATATTTGAATCAAATCAGTTTCAAAATGGAAAATGATACGTT
CACITTTGGATTTTAGTGGTCTTGTCAAGCATTAAACCATGTCAAAGAACTAATCCGCAAAAAAGCTTTTGTGGATTTGGCCGAG
ATGTTGCAATGGCGAATCTGTTCTTGGTATGAAGGCCGAAGCAACTAATGGCCGATTATGTGGAGGAGGCAAAAAAAGCAGGTA
AATTTGAAGATTACAGAAAAGTGTGGGTGAGGACCGTTGCATTATTAGCTAAAACATCGGGTACGCAAGCAGATGATCTCT
GCAAAATGTAGGCTTTGGTCATAAATAAATGTTCTTTATATGGTAATGACGGCAACGACACTCTGATCGGGCGGTGCCGGTAAT
GATTAATTTGGGGGGCGGCGAGCGGTTCCGATACCTATGTTCTTCCGCAAAAGGCTTCCGTGAGGATACGGTCTATAATACGACTACG
CTACCGGACGCAAAAGACATCATCCGCTTACCAGCGGATTTACAGCCGATATGCTGACTTTTACCAGAGGGGCAACCATCTTCT
TATCAAGGCAAAAGACGCGCAGTGGACAAGTGAAGTACTTCTAGTACTTTTCCAGAACGATGGCTCAGGAGCTTACCGTATCGACGAG
ATTCATTTGATAACGGCAAAGTACTGGATGTTGCCACTGTCAAAGAAGTGGTACAGCAATCCACCGACGGCTCGGACAGATTGT
ATGCTACCAATCCGGAATACCTTAAATGGCGGATTTGGCGATGACTATCTGTACGGTCCGACGGGGATGACTGCTGAATGG
TGATGCAAGGCAACGACAGTATCTACAGTGGCAATGGCAATGATACGCTCGATGGAGGAGAAGGCAACGACCGCCCTGTACCGGTAT
AATGGTAACGATGACTGAATGGTGGCGAAGGCAATGATCATTTGAACGGCGAAGACGGTAAACGACACTCTGATCGGGCGGCGG
GTAATGTAATCTTGGAGGGCGGCGAGCGGTTCCGATACCTATGTTCTTCCGCAAAAGGCTTCCGTGAGGATACGGTCTATAATACGA
CTACGCTACCGGACGCAAAAGACATCATCCGCTTACCAGCGGATTTACAGCCGATATGCTGACTTTTACCAGAGGGGCAACCAT
CTTCTTAAAGGCAAAAGACGCGCAGTGGACAAGTGAAGTGTTCAGTCTTATTTCCAGAACGATGGCTCAGGTGCTTACCGTATCG
ATGAGATCTATTCCGAAACCGGCAAGTACTGGATGTTGCCACTGTCAAAGAAGTGGTACAGCAATCCACCGACGGCTCGGACAG
ATTGTATGCTACCAATCCGGAATACCTTAAATGGCGGATTTGGCGATGACTATCTGTACGGTCCGACGGGGATGACTGCTGT
AATGGTGTGACGGCAACGACAGTATCTACAGTGGCAATGGCAATGATACGCTCGATGGAGGAGAAGGCAACGACCGCTGTACG
GCTATAATGGTAAACGATGACTGAATGGTGGCGAAGGCAATGATCATTTGAACGGCGAAGACGGTAAACGACACTCTGATCGGGCGG
TGCAGGCAATGATTACTTTGGAGGGCGGACGCGGTTCCGATACCTTACCGAGTATTACAGCCGATATGCTGACTTTTACCAGAGGGCA
TACGACTACGCTACCGGACGCAAAAGACATCATCCGCTTACCAGCGTATTACAGCCGATATGCTGACTTTTACCAGAGGGCA
ACCATCTTCTTAAAGGCAAAAGACGCGCAGTGGACAAGTGAAGTGTTCAGTCTTATTTCCAGAACGATGGCTCAGGTGCTTACCG

MSEPLHVLVIPSWPQSEQVDVGIFFQNLALQKRGIKTAVLAPMFRYLKREKASILTGPYGFAYRQKGLDIYAWHGMYPFR
FPLIDIRIRWVRAGLKAFFERYIGENGIPIIDIIHAHGMNYAGILAFKISQKYGIPYVVEHSSTITRGLVVRPHQWPMKNAHAAS
ALLAVSCRFAQVQLQHKYGTWQYLPNLLGNIFTRAFNPPQINRPDKYFTFCTISHLRRLKGDHVLVLSAFARALAKHPNLRLNIGG
SGQEETSLKRQARQLGTAHAVTFLGALQPEAVLDLMRNSDAFVLSRTEFTGVVYIEALSQGLPVIATRCGGAESIVSDGNGLV
PVDDDDALADALIKMYEHHSDFEPDRLLRENCLNEFGEDTVIGKLGIFRQAIAYEGKKIPVK

SEQ ID NO: 37

ATGAAGGGTTTTACCATTATTGAATTTTTGGTTGCGGGCATGCTCAGTATGATTGCTGATGGCGGTGGATCGAGTTACTTTA
CGTCCCAGAAATTAATGATGCGGCAACGAGCGTCTTCCGCGCAACAGGATTTGCGGAATGCGGCAACATTGATTGTCGCGCA
TGCGAATGGCGGGCGGCTCGGTTGTTCAATATGTCGAGCATTCTGCAACTGATGTTATCCCGATACGACGCAACAAAAT
TCTCCTTTTTCTTAAAAAGGAACGGTATAGATAAACTTATCCATACGGGAATCTTCAAATATCGGATACCGGATTTACCC
AGCGCCTTAACGCATTGATTTTCAAATACGGAAATCGATGATGTTAATGCAAGCGCGGATACTACCGTCTGAGCAGCTGTGCCAA
AATAGCAAAACCGGGTAAGAAAATATCTACCTTGCAGAAAGCAAGAGTGCAATTACAGATTAATAATGATGATAAAACAAAATGGA
AATATACCCCGTCAAAGGCATGTGGTCAATGCCTATGCGGTGCGGAGGATTGCCGGTGAGGAAGGTTTGTCCGCTTCAAATTA
ATGATAATGGGCAGTGGGGTAACTCCTCAGTTGTTGGTAAAAAGATTAATAAAATGGATATACGGTATATTTATGTTTCAAAGT
TCCTGAAGATGACGATGCGGCAAGAGGAAACATTAATAATACGGATAAAATCAACAGCGCCAAAATGCTGTTACGCCCGCC
GGGGTGGAGTTTTATTGAGTAGCGGTACTGATACCAAGATTGCGGCTTCTTCCAGACAATCATATTTATGCTTACCGTATCGATG
CGACAATACCGGGGAAATGTATGCGCAACAGAACACTTTGA

SEQ ID NO: 38

MKGFTTIEFLVAGMLSMIVLMAVGVSSYFTSRKLNDAANERLAAQQDLRNAATLIVRDARMAGGFGCFNMSEHSATDVIPDPTTQQN
SPFSLKRNIGDKLIPIAESSNIGYDPTQRLNALIFQYGI DDVNASADTTVVSSCAKIAKPGKKISTLQEAKSALQITNDDKQNG
NITRQRHVNYAVVRIAGEEGLFRFQLNDNGWGNPQLLVKKINKMDIRIYVSNCPEDDAGKEETFKYTDKFNQAQNAVTPA
GVEVLLSSGTDTKIAASSDNHIIYAYRIDATIRGGNVCANRTL

SEQ ID NO: 39

GTGCAGGCGGATTTAGCTTATGCGCGCAACGCTATTACCCACGATTATCCGAAAGCAACCGGTGCAGACAAAGACAAAATAGCA
CAGTAAGCGATTATTTAGAAAACATCCGTGCGCATTTCCATCCACCCTCGTGTGCGGTGCGCTACGACTTCCGCGGCTGGAGAAT
AGCGGCGAGATTTAGCCAGTTACAGAAAATGGAACAACAATAAATATTCCGTCAACACAAAAGGTGCAAGAAAACCAAGGCACC
AGAAAAGACCTGAAGACCGGAAAATCAGGAAAACCGGACGTTCCACCGCCGTTCTTCTCTCGGCTTACCGCCATTTACGATTTCA
AACTCAACGATAAATCAAACCTATATCGGTGCGCGCGTGCCTACGGACACGTGAGACACAGCATCGATTGACCAAAAAAAC
AACAGAGGTTGTTACTCCACCCATGGTGGTGTGACACAAAACCTACGATTTAATGGGAAAGTACCGAAAACGCTATCAC
GAAAGCCACAGCTCCGCGCTGGGTCTTGGTGTGCTGCGCGGTGCGGTTTCCGACTACGCGCAAGCTGACTTTAGACACCG
GATACCGCTACCACAACCTGGGACGCTTGGAAAAACCCGCTTAAAACCCACGAAGTCTCATTGGGCATGCGCTACCGCTTCTG
A

SEQ ID NO: 40

MQADLAYAERITHDYPKATGADKDKISTVSDYFRNIRAHSIHPRVSVGYDFGGWRIAADYASYRKWNNNKYSVNTKKVQENQGT
RKDLKTENQENGSFHAASLGLSAIYDFKLNDFKPYIGARVAYGHVRHSIDSTKTEVVTSTHGGADTKPTIYNGESTQNAHY
ESHDIRLLGLVAVGVGFDITPKLTLDTGYRYHNWGRLENTRFKTHEVSLGMRYRF

SEQ ID NO: 41

ATGAACAAAATATACCGCATATTTGGAATAGTGCCTCAATGCCTGGGTGCGGCTATCCGAGCTCACACGCAACACACCAAAAC
GCGCCTCCGCAACCGTGGCGACCGCGTATTGGCGACACTGTTGTTTGCACCGTTGAGCGGAGTACTACCGATGAAGATGAAGA
AGAAGAGTTAGAACCCGTAGTACGCTCTGCTCTGGTGTGCAATTCATGATCGATAAAGAAAGGCAATGGAGAAAACGAATCTACA
GGAAAATATAGGTTGGAGTATATATTACGACAATCACAACACTCTACACCGCGCAACCGTTACCCCTCAAAGCCGGGACAACTGA
AAATCAACAAAACCGGCAAAAATCACTACTCGTGAAAAAAGCCCTCACAGGTCTGACCAAGTGTGGAATGAAGAATTTATC
GTTTGGCGCAACCGGCAAGAAAGTCAACATCAAGCGCACACCAAGGCTTGAATTTTGGAAAAGAAACCGCTGGGACGAACCGG
GACACCACCGTTTCTGACCGGTAATGGTCTACTTTAACCGATACGCTTGGGGTCTTCTGCTTCTCAGCTGATGCGGGTA
ACCAAAGTACCAATCACTCTGTCAGCAAGTATTAAGGATGTTGAAATGCGGGTTGGAATAATTAAGGGTGTAAAACTGGCTC
AACAACCTGGTCACTAGAAAATGTCGATTTGTCGCGACTTACGACACAGTTCGAGTTCTTGAGCGCAGTACGAAAACCAACGCTC
GTTAATGTTGGAAGCAAGGCAACCGGCAAGAGAAGCAAGTAAATCGGTGCGAAGACTTCTGTTATCAAGAAAAGAGCGGTA
AGTTGTTTACTGGTAAAGGCAAGGCGAAGTGGTCTTCTACAGACGAAAGCGAAGCTTAGTGACTGCAAAAAGAGTGAATGA
TGCAAGTAAACAGGCTGGTTGGAGAAATGAAAACAACACCGCTAATGGTCAAACAGGTCAAGCTGACAAGTTTGAACCGTTACA
TCAGGCACAAATGTAACCTTTGCTAGTGGTAAAGGTACAACCTGCGACTGTAAGTAAAGATGCTCAAGGCAACATCACTGTTAAGT
ATGATGTAAGTGTGCGCGATGCCCTAAACGCTCAATCAGCTGCAAAAACAGCGGTTGGAATTTGGATTCCAAAGCGGTTGCGGTTT
TTCGGGCAAAAGTCACTCAGCGGCAATGTTTCCGCGAGCAAGGGAAGATGGATGAAACCGTCAACATTAATGCCGGCAACCAACATC
GAGATTACCGCAACCGTAAAAATATCGCATCGCCACTTCGATGACCCCGCAGTTTCCAGCGTTTCCGTCGCGCGGGGGCGG
ATGCGCCACTTTGAGCGTGGATGGGACGCAATGAATGTCGCGAGCAAGAAAGGACAAACAAACCGTCCGCAATTAACAAATGTCG
CCCGGGCGTTAAAGAGGGGGATGTTACAAACGTCGCAACATTAAGGCGTGGCGCAAACTGAAACAACCGCATCGACAATGTG
GACGGCAACCGCGCTGCGGGCATGCCCAAGCGATTGCAACCGCAGGTCTGGTTCAAGCGTATTTGCCCGGCAAGAGTATGATGG
CGATCGCGCGGCGCACTTATCGCGCGAAGCGGTTACGCCATCGGCTACTCCAGTATTTCCGACGGCGGAAATGGATTATCAA
AGGACCGGCTCCGGCAATTCGCGCGGCAATTTCCGTGCTTCCGCATCTGTCCGTTATCAGTGGTAA

SEQ ID NO: 42

MNKIYRIIWNLSALNAWVAVSELTRNHTKRASATVAVLATLFLFATVQASTTDEDEEELEPVVRSALVLQFMIDKEGNGENEST
GNIGWSIYYDNHNTLHGATVTLKAGDNLIKQNGTNFTYSLKKALTLGLTSVGTTEL SFGANGKKNVITSDTKLNFAKETAGTNG
DTTVHLNIGSTLTDLAGSSASHVDAGNQSTHYTRAASIKDVLNAGWNKGVKGTSTTGGQSENVDVFRYTDVFLSADTKTTT
VNVESKDNKRTEVKIGAKTSVIKEKDGKLVTKGKGENGSSTDEGEGLVTAKEVIDAVNKAGWRMKTTFANGQTGOADKFEVTV
SGTNVTFASGKGTATYKDAQGNITVKYDVNVGDALNVNQLQNSGWNLDSKAVAGSSGKVISGNVSPSKGMDET VNIAGNINI
EITRNGKINDIATSMTPQFSSVSLGAGADAPTLSDVDGALNVGSKKDNKPVRIITNVAPGVKEGDVTVNVAQLKGVAQNLNRRIDNV
DGNARAGIAQAIATAGLVQAYLPGKSMMAIGGGTYRGEAGYAIYSSISDGGNWIKGTASGNSRGHFGASASVGYQW

SEQ ID NO: 43

ATGATTGAATCCATAGAAAAACACAGGAATCTATCAGAAAAAACAGAAACCCCAACCGGTCATTCCCGCAAAAGCGGGAAATCC
AGAAAACGCAACGCAACAGGAATTTATCGGAAAAAACAGAAACCTCACCGCGTCATTCCCGCAAAAGCTGGAATCCAAAAACGCA
ACCGCGCAGGACTTTATCGGAAAAAACAGAAACCCCAACCGGTCATTCCCGCAAAAGCGGGAAATCCAAAAACGCAACGCAACA
GGAATTTATCGGAAAAAACAGAAACCCCAACCGGTCATTCCCGCAAAAGCGGGAAATCCAGCAACCGGAAACCAACAGGAATCT
ATCAGCAAAACAGAAACCCCAACCGGTCATTCCCGCGCAGCGGGAAATCCAGAAACCAACCGCGCAGGACTTTATCGGAA
AAAAACAGAAACTCCACCGGTCATTCCCGCAAAAGCTGGAATCCAAAAACGCAACCGCAACAGGAATTTATCGGAAAAAACAG
AACCCCAACCGGTCATTCCCGCAAAAGCGGGAAATCCAGACCGGTCGCGCACGGAATCTTACCAGGATAAAACAGTTTCTTAGATT
CCACGTCCTAG

SEQ ID NO: 44

MIESIEKPKQESIRKRNHRVPAKAGIQKRNATGIYRKKQKPHRRHSRKSWNPKTQRGRTLSEKTETPTVIPAAGIQKRNATGIYRKKQKPHRPSFPQKRESSNRKTTGIYQQKQKPPPTVIPAAGIQKHNAAGLYRKKQKLRPSFPQKLESKNATQQEFIGKRNPTAVIPAAGIQTRRHGILPDKTVSLDSTS

SEQ ID NO: 45

ATGATTTTCTCCATCATCGTCCCTATTACAATGTGGAAAAATACCTCCGCTGCTGCGTGGATTCCGTGCTTCCGAAAAATTTTGCCGATFATGAAATGATTTTGGTCGATGACGGTTCCGCGGATGGCTGCGGGAAGATTGCGACGAATATGCAGGCAAAATACCGCATATAAAAGTGATTCATCAAGAAAACGGCGGGCTGCGGATGCCCGCAACGCCGATCCGAGCGGCAAAAGCGGATTACCTGATCTTTTGGACAGCGACGACTATTGGCCGATACCAACCGTTCAAAAAACCGCGGGGGGGGGGATTCTCTTTGATTTACAACAACCTGCAGACAAAAAGGTTGATTTGATCTGCTCCCTCGTCCCTCAATTACCACGTCATCCCAAAGGGGCGGACTTTCCGATAATGATTTTGTCCGCCATTTTGAACGTTGGTGGAGGGGGCGTACTATATCGCCAACCGCTGGACAAAAATCGTCAGGCGGAAAAATAATCATCAAAAAACATCTGTTTTCCCAAAGGATACATTACAGGATTTCCTGACAGTTTGCATTTGGCCGTTTTATCAAGACTTTTGCCTTTTACGATAACCTTTTACCAGTACCGCTTCCGCGGCTCTATCAGCCACAACATCAATACAAAAAATTCAGCGATGTGCTGACGCATCTCGACCGGGGTGGATTTTTGTAGTCGAAAAACAAAAATCCCCCATCTACGGCGGTTTGCAAAAATTTGCTTTCCGCAATATCCGGCTATCTGAGGCTATATTTGGTAAGGCTTTATTTTCCAAAAACATTCATCTACCTACCGGAAATTTTTTTCATTTAAAGAAAAATTCAGAAAAAGATATTCGGCGCAAGGCAATCCGTCGGTTTTTATCGGGAAGACCGCATTCATAGGATTCGCCGATTTGCGCCTGCTCGTACCGCCTATGCTGTACCCGGCAATCAAGCCGTTTTATCAAAAAATTTTTTCCGGAATA

SEQ ID NO: 46

MIFSIIVPIYNVEKYLRCVDSVLAENFADYEMILVDDGSPDGCCKICDEYAGKYPHIKVIHQENGLSDARNAGIRAAGDYLI FLDDDDYADTNRSKNAGGGILFDLQQLADKIKVDLILHPSLNYHVIKPKGAFSDNDVFRHFETLVEGRYYIANAWTKIVRREII IKNNLFFPKGYIHEDFPYSLQLARFIKTFAYDNPFYQYRVLGGSISHNIKYKNSDVLTHLDRGVDFLENKNSPIYGGKQFVFDNIGYLRSLVRLYFSKNIIVIRKYFSFKEKCRKIFGAKAIRPVIKGTAFIIGLPILRLLVPPMLYPAIKAVYQKFFSE

SEQ ID NO: 47

ATGACATCTGCAAAATTTAATAATTAACGGTTTTGGAGATGTGAAATTAACCCCTATTCACCACTCTTGGGATATAAGCTTGGGATTCATTATTTGGTTCAATTCCTTATCTGATTTAATCTATAATGTGGATAACAATAGAAAATAAAATGGAAATTAATGTTAA TAATGCTATCCAAGCTGCAGATAGCTTTTTAAGCAGTATTTGGAAGGATAACAAAAATAACAAATCTGCTTCTTTACTTGCATCCCTCGATAACATTTTTTAAATTAAGAAATGTATCTGAGATACAGAAAAACAGGAAAAATTTAAACCTAATGATATTCAACAAG CAATTTGGTGATATATTCTGCTGCTGGTATGATGATACAATATATAAAACAACAACAGAGGGCGATGGCTCAAAGCAAATCTT ACCAACTAAATTAACAACTGGTTTAAATGATGCTTAAATCTAGAATGCTAAAATCCTCTACTGTTTTACAGCATGAATGAAT TATTTGGGATTTAAATAAAGGATTTGAAACGAGAGGCTTGGCGAATCTATAATGAATATAGATGATTTTACACCAAGTAAGA TAGCAAACTTTTTTCCGGATCCTGATACATACAGCAATGATTTAGAAGAAATCTAGGTTTATATATTCTTACTGTTCTGATGA TGCAAACTTTGGAAGGGGGCGAAGATTATATTGGACGAGGATAAGTGAATGGGGAGAGTTACTGGAAAAATGGTATAAAACAA GATTTTCTCCCTTATCTTGAAGAAATGGGACCAATTTCCGAAATTTGAAGATTGGCTGCCTGAATTCCTGAATGGGCAAGAG AGTGGTTGAAATAGATCCCAACCGTTGAGGCAAAATATCATGCTACGACCCCTCGCCCTAGATCTAGACGGCGCAGGATAGA AACCGTTGCCACCAAGGCTTTGACGGCAGCTTATTTGATCACACAAACCGGATCCGACCCGCAACCGGTTGGGTTCTGCC GATGACCGGTTACTCGTCCGCGATTTGACCGCAACGGCATCTCGCAACCGGTCGGAACTCTTCCGGCACAACACCAAACTGG CAGACGGTTCTTTTGGCAAAACACCGCTATGCGCTTTGGCCGAATTTGATTCAAACGGCGACAACATCATCAACGGCGCAGACGC CGCATTCCAAACCTGCGTGTATGCGAGGATCTCAATCAGGACGGCATTTCCAAACCAACGAACTCCGCAACCTTGAAGAAATG GGTATTCATCTTTGGATCTCGCTATAAAGATGTAATAAAAAATCTCGGTAAACGTTAACACTTTGGCTCAGCAAGTACTACA CCAAAACAAACCGTGAACCGCAAAAAATGGGGGATTTACTTTTAGCAGCCGCAACTGACAGCCGCTTCAAAGCAAAAGTGA AACTACTGCCAAACAGGCAAAAGCCGCAATCTTGGCGGCAATTTGGTCTGCTGCGCGATTTGCGCGAAGCTGCCGATTTGCCGG GATTTGGCCAAATGTGTAAGGCTTATCTGCGCGCGAAANNNNNCACACACTTAATTA

SEQ ID NO: 48

MTSANFNINGFGDVKLTPYSPLLYGKAWDSFIGSIQSLSDLIYNVDNRRNKMEITVNNAIQAADSFLLSSIGRDNKITNTASLLAS LDNIFLNLRNVSRIRETGFKPNDIQQAIGDIFIAAGDGLYIKQTEAMAQSKFLPTKLTGLNDVLSNRMLKSTVLRQHELN YLGFKIKDYGNERLGESIMNIDDFTPSKIANFFADPDTYSNVLEEVSRFIYSLVPPDANPWKGGEDYIGRGISEWELLEKWKYKQ DFLPYLEKEWDQFPKFDWLPPEFPEWAREWLKLDPKRSKYHYVDPLALDLDDGGIETVATKGFAGSLFDHTNNGIRTATGWVSA DDGLLVDRDLNNGNJIIDNGAELFGDNTKLDGFSFAKHGYAALAE.LDSNGDNIINAADAAFTLRVWQDLNQGDISQINELRLEEL GIQSLDLAYKDVNKNLGNNGTLAQQGSYTKTNGTTAKMGDLLLLAADNLSRFKDKVELTAKQAKAANLAGIGRLRDLREAAALS GDLANMLKAYSAAEXHTLN

SEQ ID NO: 49

TTGAGCGGGAATATCGCAACTCTGATAAGGATGGCAATGATTTCCAATAGCAAAACAGCAGAAGAAAAAGCTGTCGTTTTAAAG AGGAAGATATCTGTTTTATACGGTTCCAAAAAAGATCAACGTCAGCAGCTTGAAGATAAAAATCATCAACGCAATCCTAATGT AAAAAATAGGACATCAGAAAAATGAAAAATAAAAATATGGTTATAACTTTGTAGATGCAGGTTATGATATGTAAGGGAGAAGAT GAAATTTAGAGGACTTCAAAATCAAGTATTTTACTCACCGGTTTGGTTATGACGGTTTTGTATATTCCGGAGAACGTCCTT CCAATCTTTACCGATTTCCGGAAACGGTGAATATTTCCGGCACTGGCAATATATGACCGATGCCATACGTCATCGGACAGGAAA AGCAGGAGATCTAGCGAAGATTTGGGTTATCTGTTTTATTACCGTCAAATGTCGGAGCAACTCTTATGCTGCGACTGCCGAC GACCGGGAGGGGGAGGAAAAACATCCTGCCGAATACGCGTTGATTTGACAAAAAACCTGAATGGCAAGCTGATTAATAATC AGTATGTGCAAAAGAAAGGACGATGAAAAAGAACCGCTGACATTTACAACATTAAGTCAAAAATGGACGGCAACCGCTTACC GG CAGTGCCAAGGTCAATCCTGATTTAGCGAAAAACCATGCCAAGAAGGAGCAATTTGTTTTCCATACCGATGCCGATCAGCGCTT GAGGCGGTTTTTTTCCGCGATAAGGGGGAAAGACTTCCGCGGAGGTTTTATCAGCAACGCAACAGCGTATTTCCGGATTTCCGAG GCAAAACAAAAACAGAGGCGAGCAACGCAAGCAGATACAAAAATTTGCCCTTGGCGTCTAGAAAAACACCAAAAAATCTTGATTTCTT CAAAAATTTCCGTTGACGAGGCAAGTGATAACAAATGCCCGCAGTTTCCATTTCTCTATGCCCAGTCTTGGTCATCCCACAAA CTTCTTGTGCAAGGGCGTGAAATTCCTTTGGTAAACCAAGAACAAACCATCGATCTTCCGACCGGACAGGAAATGACCGTCCGTG CTTGTTGCGATTTTCTGACCTATGTGAACTCGGACGGATAAAAAACCGACCGCCCGCAAGTAAATCAAAGCGGAAAGATAAAGG GGAGGATGAAGAGGGTGCAGGCTTGTAAACCGCGAAGAAAGCAAAATCAGCGATGAAGACGAAATTTCCGAAGTGAAGGCGAA GAAGACGAAGAAATCTGCAAGAAAGAAACCGAAGAACGAAACCGAAGAACGAAAGAGGGGGAAGCTGAAGAGGAAAGAACCGG AAGAAGCAGAAAGAACTGAAGAAAAATCGCCGCGAGAGGCAACGGCGATTCCAGCAGCATCTGCTGCCCGGAAAGCCCTAA AGGACGGACATCGACTTTTCTGAAAGGATTCGCAACCGCGGAAAGCCGACATTCGCAAAACCGGAAAGCACACTATACCGGC ACTTGGGAAGCGCGTATCAGCAAAACCCATCAATGGGATAATAAGCGGATAAAGCGGCAAAAGCAGAGTTAATGTTGATTTCCG GCGAAGATCGATTTCCGGAACGCTGACGAGAAAAATGGTGTAGAAGCTGTTTTCTATATTGAAAAAGGTTGATTTGAGGGCAA CGGTTTCCACGCGACAGCGCGCACTCGGAGAACGGCATCGACTTTCCGGGCGAGGTTCCAGCCAGATCGCAGAGCTTCAAAGCC GACAATCTTCTGTAAACGGGCGGCTTTTACGGCCGAAAGGCGGAAATTTGGCGGTAATTTCAATAATGATGGGAAATCTC TTAGATAAATGAAAAATTTGAAAAATGAAGTTGAAGTTGAAGCTGAAGCTGAAGTTGAAGTTGAAGCTGATTTGGCAACAGTT AGAACCTGATGAAGTTAAACCAAAATTCGGCGTGGTATTTGGTGCAGAAAGATAATAAAGAGGTTGAAAAATGA

SEQ ID NO: 50

SEQ ID NO: 63

GTGGCGAGATGGAAAAAGACGGGATATTGCCCGATGGGTTCTATTTGGAGATTTGGCGGAAGAAAAACAATAAAAGCGTGC
TTGCCGCCATTGCCGAAACCGGCGAAATTTGGAAAAACCCCTGACCCATCAGGAAATCACGTCTGCCTTCCCCTCCGCAACCC
CATAACAACAACAGCAGTATTATCGGCGCAGCGTCACTGACGGCGTTTGGGTTCCATCCCGCTATATCCACGCGGAAGAC
TATAAGTTTTGGTACGAAGCCGCGAAATTTGGCAGGCTGGCGAATTTACCCCGAAGCTTGGTCAAATACCGTTCCATCAAGACC
AGACTTCTTCAAACACAACCTGCAACAGCGTAAGACGGCGTGGAAAAATCAAAGAAAGAAATCAGGGCGGGGATTTGGAAGCGGCG
AGGCATAACCGTGGGTGGACTGCCTGAATACGGGCTTTTGAATCAACGGCATATGCGTTGCACGAAAAAGCCTTGTCCGGA
CAGGATATCCGATACCTCCGCTGTTCTGTACGAATATTTCTGTCTGTTGAAAAAGTATTCTTTGACCGATTACTGGATTTTC
TGACAGACCGGATGAGGAAAGCTGTTTCCGCGACCGCAATATAGGAAAAATCTGAAAAAAATGTTACGCCCTTGGAAATACCG
GGCTATTGA

SEQ ID NO: 64

MGEMEKDADIAAMGSYLEILAEENKSYLAAIARNGEIWKNP LTHQEITSAFPLRNP IHNNTMIMRRSVIDGGLRFDPAYIHAED
YKFWYEAGKLRLANYPEALVKYRFHQDQTSKHNLRQRKTAWKIKEEIRAGYKWAAGITVGSCLNYGLLSTAYALHEKALSG
QDIGYLRFLYEYFLSLEKYSLTDLDFLDRVMRKLFAAPQYRKLKMLRPWKYRGY

SEQ ID NO: 65

ATGACGCAAGAACGTTTACCCGAATTTTTCGACCGCGCCCGACGCTGACCGTACAAGACCCGCTTCCGCGATTCTCGGCGCGG
CCGAAAAACGGCATCTCACTTACCGCTACGCCGATGCCGTGCGCTGTGCGGACATTCTGCCCGACCGTCCGCGGGCGTACTT
GATGGTTATCAAAGTCTGAAAGCCCTTACGGCGAAGAGCTGCCGAAACCGGCGGCATCGAAGCCGCCATGCAGGAGCGGCGC
GACGAAGGCACGGTCCGCGTAACCGCATCCGCTGCTCAACTCTCACCGGCGCAGCCCGAAAAACGGCTTCCGAGGCGATCGGAA
TACAGGGACGCTTCCGCGCGCGCAACCTCTTATCTTTGGTGCAGGGCGAAATCAACGGCACACTCGCGCTCCGCCCGCGGACAC
CGCAAAAACCGTCCGCTGAGCCTCAACGCCCGCTGCAACCTTCCGACCGCAATTCGACCGCAATCATGCCAAAGCCGTCAGC
GGCAGCGCAAGTACCGACGAACCTCAAACACTTCGGACAACCTGCGCAGGCACGCGTTAAAGCATTTTTGACCGAATCGGCGGACG
ACCCGCGATTCTCATCTGTCGCGGAAGTGTGA

SEQ ID NO: 66

MTQERLPEFFDRAPTLTVQDPLAFLGAAENGILTYRYADAVRLCGHSCTPVAGAYLMVIKGLKALYGEELPERGGIEAAMQGAR
DEGTVGVTASVVQLLTGAAPETGFGGIGIQRFARRNLLSFGAGEINGTLALRRRDTGKTVAVSLNLAALQFPAPQMRDIMPRAVS
GSASTDELKHFQQLWQARVKAFLTESADDPFVIVRE

SEQ ID NO: 67

STGTGNNNNCGAACGGTTTGGATGCCCGTTTACCGGATGATATGCAGGCAAAAACACTACGAACCCGGTGGTAAATACCATCTGT
TTGGTAATGCTCGCGGAGTGTAAAAATCGGGTTACGCGCTCCAAACATTTGATGCAACTGCGGTCCGCCCCATACTGCCAT
TACACACGAACGGACAGGATTTGAAGGTGTTATCGGCTATGAAACCCATTTTTTCAGGGCACGGACATGAAGTACACAGCTCCGTT
GATAATCATGATTCAAGAAGCACTTCTGATTTACGCGCGGTGAGACGGTGGTTTTACTGTTTACCACTCATCGGACGGGT
CGGAAATCCATCCGGAGGATGGATATGACGGACCGCAAGCGAGGATTATCCGCCCCCGGAGGAGCAAGGGATATATACAGCTA
CTATGTCAAAGCACTTCTACAAAAACAAAGATAAACACTGTTCCGCAAGCCCATTTTTAGACCGTTGGCTAAAAGAAAAATGCC
GGTCCCGCTCTGGTTTTCTAGCCGTCGGGATGAAGCAGGAAAAACTGATATGGGAAAAACGACCCCAATCAAAAATTTGGTGGGTA
ACCGTATGGATGATATTCGCGGCATCATCAAGGTGACGCAATCCTTTCTAACGGTTTTTCAGGGATGGGAGTTGGGGCAAT
TACAGACAGTCCGTTAAACCCGGTAACTATGCGGCAGCACGGAAAACTTTACAGGGTATTCAAAATTTAGGAAATTTAAGTCCG
GAAGCACAACTTCCGCGCGCAACCGCATACAAGACAGTCTTTTCCGGTAAAAGACAGTATCAACTCTGCCAGACAATGGGCTG
ATGCCATCCGAATAAATGCAACAGCCAACTGCCCTTCCGTAGCAGAGGCGCAGGTACGGTTTGGGGCGGTAAAAAAGT
AGAATTAACCCGACTAAATGGGATTGGGTTAAAAATACCGATTATAAAAAACCTGCTGCCGACCTATGCAGACTTTAGATGGG
GAAATGGCCGGTGGGAAATAACCCGCTAAATCTATAACGCTGGAGGAAAAAGCCAATGCTGCAACTTATCTCAATTTAGTTAATC
AATTAACCTGGCAAAACTTAAAAAACATTGCGGCTCAAGATCCAAGATTGAGTCTAGCTCTTATAAAGAGTAAAAAAATTTCC
AATAGGAACTGCAACTTATGAAGAGGCAGATAGGCTAGTAAAAATTTGGGTTGGTGAAGGTTCAAGACAACTAGTGGAGGCGGA
TGTTAAGTATAGATTGGCACTCGACAAATATCGGCCACCAACAGAAAAAATTTACAAATTTGCAACTACAGGATTTCAAGCAAAAT
TTGAAACTTATACTATTGATCAATGAAAAAGGAATAAAAAATTTAAAAATGGACATTTAAATATAATAGTAA

SEQ ID NO: 68

MXXNGLDARLRDDMQAKHYEPGGKYHLFGNARGSVKNRNVYAVQTFDATAVGPILPIHTERTGFEGVIGYETHFSGHGHEVHSPF
DNHDSRSTSDFSGGVDGGFTVYQLHRTGSEIHPEDGDYDGPQGSDDYPPGGARDIYSYVVKGTSTKTKINTVPQAPFSDRWLKENA
GAASGFLSRADAEAGKLIWENDPNQNWGNRMDDIRGIIQGAANPFLTFGQFLGVGAIITDSAVNPVTYAARKTLQGIHNLGNLSP
EAQLAAATALQDSFAVAKDSINSARQWADAHPNITATAQTALVAEAAAGTVWRGKVNLPKWDWVKNTGYKTPAARPMQTLDG
EMAGGNPKPKSITSGGKANAATYPQLVNQLTGQNLKNIAAQDPRLLSLALHKSEKNFPIGTATYEAADRLGKIWVGEARQTSGGG
WLSIDGTRQYRPPTEKNSQFATTGIQANFETYTIDSNKRNKIKNGHLNIR

SEQ ID NO: 69

GTGGGCATCAATGCCAATCCTAAGTGTGCTGATGAAGCAGGAAAACTGATATGGGAAAAACGACCCCAATAAAAAATTTGGTGGGCTA
ACCGTATGGATGATATTCGCGCATCGTCCAAGGTGCGGTTAATCCTTTTTAATGGGTTTTCAAGGAGTAGGGATTGGGGCAAT
TACAGACAGTGCAGTAAGCCCGGTACAGATACAGCCGCGCAGCAGACTCTACAAGGATTAATCATTAGGAAATTTAAGTCCC
GAAGCACACTTGGCGCTGCAACCGCATTACAAGACAGTGTCTTTGCGGTAAAAGACGGTATCAATTTCCGCGACACAATGGGCTG
ATGCCCATCCGAATATAACTGCAACAGCCAAAACCTGCCCTTCCGTAGCAGAGGCGCAGGTACGGTTTGGCGCGTAAAAAAGT
AAACCTAACCCGACCAAGTGGGATTGGGTTAAAAATACCGGCTATAAAAAACCTGCTGCCGACCTATGCAGACTTTAGATGGG
GAGATGGCAGGGGGGAATAGACCGCTAAATCTATAACGCTCAACAGCAAAGCAGATGCTTCCACACAACCGCTTTTAAAGGCG
AACTAATTTGGAGAACAATAGTAGTGGGCTGCTATAACCAAGCATGTATAAGACAACAAGAAATTTACGGCTTTAAANNNNC
ACACACTAA

SEQ ID NO: 70

MGINANPNCADEAGKLIWENDPNKNWVANRMDDIRGIVQGANPFLMGFQGVGIGAITDSAVSPVTDAAQQLQGINHLGNLSP
EAQLAAATALQDSFAVAKDINSARQWADAHPNITATAQTALVAEAAAGTVWRGKVNLPKWDWVKNTGYKTPAARPMQTLDG
EMAGGNRPPKSITSNKADASTQPSLQQLIGEIQISSGHAYNKHVIRQEFIDLXXXHT

SEQ ID NO: 71

ATGAAGGGGATGATGATGTTTGAACGCAAGTGTGATTGCAATGGCTTGTATTTTTGCCCTTTCAGCCTGTGGGGCGGCGGTGGCG
GATCGCCGATGTTAAATCGGCGGACACGCTGTCAAAAACCGGCGCTCTGTTGTTGCTGAAAAAGAGACAGAGTAAAAAGAAGA
TGGCCACAGGCAAGTCTCAAGGACAGGGTGCCTATCCACACAAGGCAGCCAAGATATGGCGGCGAGTTTCCGCGAGAAAAACA
GGCAATGGCGGTGCGGCAACCAACCGGCAAAACCAAAAATGAAGACGAGGGACCGCAAAATGATGCTGCAAAATTTCCGCCGAAT
CCGCAAAATCAACAGGGGAAACAACCAACCGCGATTCTTCAAGTTCCGCGCCCGCTCAAAACCTGCACTCGCAATGGCGGTAG
CAATTTTGAAGGGTGTATTTGGCTAATGGCGTTTTGATTGATGGGCGCTGCAAAAATATAACGTTGACCCACTGTAAGGCGGAT
TCTTTGTAATGGTAAATTTTATGGATGAAGAAGCACCGTCAAAATTCAGAAATTTGAAAAATTTAAATGAGTCTGAACGAATTTGAGA
AATATAAGAAAGATGGGAAAAAGCATAAATTTACTAATTTGGTTCGACAGCAGTTCAGCTAATGGAAGTAAACAAATGTCAT

CATTTATAAAGACAAGTCCGCTTCATCTTCATTTGCGCGATTGAGCGGTTCTGCACGGTCGAGGAGGTGCTTCTGCCGAGATG
CCGCTAATCCCGTCAATCAGCGGATACGCTGATTGTCGATGGGGAAGCGGTCAGCCTGACGGGGCATTCCGGCAATATCTTCG
CGCCGAAGGGAATACCGGTATCTGACTACGGGGCGGAAAAATTCGCCGGCGGATCGTATGCCCTCCGTGTCAAGGCGAAC
GGCAAAAGCGAAATGCTTGGTGGCACGGCCGTGTACAACGGCAAGTGTGCAATTTTCATACGGAAAAACGGCCGTCCGTACCCG
ACTAGAGGCGAGTTTCCGCCAAAAGTCGATTTCCGGCAGCAATCTGTGGACGGCATTATCGACAGCGGCGGATGATTTGCATAGG
GTACGCAAAAATCAAAGCGCCATCGATGGAAACGGCTTAAAGGGAATTTGGACGGAAAAATGGCGCGGGGATGTTTCCGGAAG
GTTTTATGGCCCGCCGGCGAGGAGTGGCGGGGAAATACAGCTATCGCCCGACAGATGCGGAAAAAGGCGGATTCGGCGTGT
GCCGGCAAAAAGAGCAGGATTA

SEQ ID NO: 72

MKGMMFERSVIAMACIFALSACGGGGGSPDVKASDLSKPAAPVVAEKETEVEKEDAPQAGSQGGAPSTQGSQDMAVSAENT
GNGGAATTDKPKNEDEGPQNDMLQNSAESANQTGNNQPADSSDAPASNPAANGSNFGRVLDLANGVLIDGPSQNTLTHCKGD
SCNGDNLDEEAPSKSEFENLNERIEKYKDKGSDKFTNLVATAVQANGTNKYVITYKDKSASSFARFRRSARSRRSLPAEM
PLIPVNOADTLVDGEAVSLTGHSGNIFAPEGNRYLYTYGAELKPGGSYALRVQGEPAKGEMLAGTAVYNGEVLHFHTENGRPY
TRGRFAAKVDFGSKSVDGIISDGDLLHMGTKFKAAIDGNGFKGTWTENGGDVSGRFYGPAGEEVAGKYSYRPTDAEKGGFV
AGKKEQD

SEQ ID NO: 73

ATGCGGTTCTCAAGGTTTGGAGCCAAAGGTCGTCTGAAACAACAAATACGGTTTCAGACGACCTTTCTTTCAACAAGCCACCAG
GCAATCAGACAAAAGCAGCACATCGCCACATCCATGTCGGCAGTACGGCCGGCACAACACCATCCGACGGCGGGGATACCA
CCTCAAAGGTCGCGAGCTCATCGGCAAGGCAATACAGGCGATACGCGCAACCTGCATATGAAAGTGTACAAAGTACTGAAACC
TCTCAGAGCAAAACGCAAAAACGGCAATGCCAAGTTACTGTCTGGTTACGGATTTCAGTCAAGCGGCGATTACAGCCAAAGCAA
TCAAAGCAGACCATGCTCCGTAAACGAGCAAAAGCGGATTTATGCGGAGAAAGACGGCTATCAATCAAGGTCGGAACCATAC
AGACCTCAAGGCGGATCATCACCTCCGGCAAGAGTGGCGAAGCAAAAGGAAAAACCTTTTTCAGACGGCCACCCTTACTGCC
AGCGACATTCAMAAACACAGCCGCTCAAGAGGCAAGGATTCGGCATAGGGCGGAGTTTCGACCTGAACCGGCGGATACGAA
CGGTTACCGCAAAACAGGCGAGGCTACCGACAGGATAAGCTGGCAGCCGGTACGGCAGCGACGACGACGTAAGAGCAGCAT
CACA AAAAGCGGATCAACACCCGCAACATACACATCACCAGCAAGCGGCAACTTCCGCAACAGGCGGATGCAAAAAGAA
ACCGAAGCGGATCTACACCGGATCGACACCGAAACTGCGGATCAACACACAGGCGCTCTGAAAAACAGCTTCGAAAGAGC
CGGTCCGCAAGAGATCAACTGCAAGGGGAAGTAAACGAAAGGATTCGGCAGAAACCGCCGCAAGCCGTAGCGGGCGTTGCCGA
CAAACTCGGTAATACCAAAGTACGAACCGTATCAGAAAGCGCAACTCGGCTGGAGGCGGAACTGCAAAAACAGGACAGCAAA
CGCGAAAAAGCGCCATCCGCGCATCCCTCGGCCAAGTAAACGCTATCTTGC GAAAACCAAAGCGCTACGACACCTGGAAAG
AAGCGGCGATAGGCGGAGCATACTGACGGGGCGGCGGATGACGACCGGAAGCTCGGCGGCTACTGGCCGCGAGCGG
CACTTCCCTTGTGCACCATATTTGGACAAAAGCGGCGGAAACCTCGGTCGGCGGGCAAAGCGGCGGTCAACGCACTGGGCGG
GCGGTATCGGCTATGCGGCGGGCGGAAATGTCGGTACGGCGGCAAGTGGGGCGCAATGTCGATGGAAACAATAGGCACTGCATC
CGAAAGAAACACAAATCCTTAACAAACTGTCAAAGGCAAACTCGGCTGAAGAAGCAGTACCGCTAAAAGCCGCTGCATGTGCATT
AACCCGGTGC GCGGAAAGCGTACCTGACTTCGACCTCTTTATAAAGGACTAAAACCTCAAAGATGCCGGTAAACAGTTTGT
GCGGACAAAACGATTTGATGCGGACGGATGCAATTTAAATGGAACATGGAACAGCCTGAATGATATACGCGAGTACGACC
GTGCTGCCACAAAATTAAGGTCGCGGCAATATGGGATTTGGTGCACAGCCTTTGTCCGTTCCGGTGTATAGCGGAGGCT
GTGCAGTACCGGATTTGGCTGTGCGGCGGCGGAACTTATTTGCAACGGCAGGATGACCGGTTGGTATACACAGGCTCAGAAGGA
AGCCGGCAATTTGTTGGCACTTACAGTCCGATTTGGTAAAGGCTTGTCTATCTTTGGTACCAATAAGAACTACGAATCGC
CGTTAGTATCTGATGCGAAAACTAGCCGTATGGGATTTGAAACCGTATTACGCGCAAAATGGGAACTTGGCAACCGGTT
GAAAACCTCTTCTGACTCCGAAAACCTGACTGACTACAGGCAAACTCTGTCCCAATCCGAAGTCGGTATCAAGTGGGCAAGGG
ATTGAAGGACAGGGAATGCCTTGGGAGGATATGTCGGTAAAGGCTTGTCTGCCAATGCAAGGTTACTAAAATTTTAAAAACAT
TTGATTTATTTGATCGTGTACAGGACCGCAATCAGTGCCAAAACCTGGATACGCAAACTACGCGCGCTGTCCAAAACCGGA
ACAGCTTTACAGTACCATGAAAGGTAATCATCGATGAGACGGCAAAATTTCAAAGTTTATGAATTAATCAGAAGTACCTGTAAAGGCA
GACATGATCAAACAGCGCAAAATCCATCTGGCCATACCCGCAAACTAATAAGGAGCAAAGATTGCAAGTGCACGTTGGT
AGTATGGCAAAAGTCAAAACATTACAGTCAAATTAACGGAGATCGAATAA

SEQ ID NO: 74

MRFSRFPKGRLLKQIRFQTTFLSTSHHGNQTKAAHRHIVHGSTAGTTTIRSGDITLKGALIGKGIQADTRNLHIESVQDTE
SOSKQNGNVQVTGYYGFSASGSYSQSKVKADHASVTEQSGIYAGEDYQIKVGNHDLKGGIITSGKSAEDKGNLFTATLTA
SDIQNHSRYEGRSFGGDFDLNGWDTVTDKQGRPTDRISLAAGYSDSDSQQSITKSGINTRNHITDEAGLQARTGRTAKE
TEARIYTGIDTETADQHTGRLKNSFDKDAVAKEINLQREVTKEFRNAAQVAVAADKLGNTQSYERYEARTLLEAELQNTDSK
AEKAAIRASLQVNVNAYLAENQSRDYDWTKEGGIGRSILHGAAGLTTGSLGGILAGSGTSLAAPYLDKAEENLPGAKAAVNLGG
AVIGYAAGNVGTAAVGANVDWNRQLHPKETQILNKLKSKGSAEEQYRLKAAACALTRCAEGVPDFDPLYKGLKNLQDAGKQFV
AEQNVLMRTDAFKYGTWNSLDIRSSYDRAATKIKGAGNMGGLATTIVFGSGAIGGGLCSTGIGCAAGGLIATAGMTGGVYQASEG
SRQLFVYQSDFGKVVLSLGTPIEYESPLVSDAKNLAVWGLETLITRKLGNLATGVKTSLTPKTDVQRNLSQSEVGIKWGK
IEGQGMPEWEDYVGGKLSANARLPKNFKTFDYFORGTGTASAKTLDTQTARLSKPEQLYSTMKGYIDKTNFKSYELSEVPLRA
DMIKQREIHLAIPAQTNKEQRLQLQRVVEYKGSQNTITVKITEIE

SEQ ID NO: 75

GTGTGTGNNNNNGATAAACCGCAGCTATAAAACCGGTAAGTGGTACAAACTAAAACATGTTACTGAAATCAAAGAGCATAAAAACG
CCAAAGCCGACCCGGTGAACCTCAGTGCCTCACAAGGTTTGAATCAATCCGGCGGCAATATCGGTGCCACGCCACCTTGT
TGATGCACCCCGCGCTCCGTTAAAATCGAAGCGGACGTCGGGCTGTTCTCTATGCCGTGGAAGATCTCAACTACGCAAACT
GACACCCGTACCAAGCGCAAAATTTATCGGCATTAACCTACGACAAAGTGCACGACACCACCCACACCATGAAAACCGCCCTGC
CCTCAAGGGTAGTTGCGAATCGGCCAACCTGCAATCAGGCTGGGACGCCAACTGCAAGGCAACCGGCTTTGAAAACCGGCTGG
CGGCGCAGCCATCCGTGCAAGTGTAGGCGATCAGGACAGGACAGATGCCAAGATTAITCTTGAAGGCATCAAAGTGTGTGCGC
ACTGAAAACAGTAAAGCAGTAGCAAACTGCCCCCTGGCAGAAACAGCCGGACGCGGCGAGCAATATCGAAAACCTGCAACTGCCAA
GTTTACAGGCTCCGTTGCGCCGTACTCTGCCCCCGGCGGCTATCGTTGATATTCGAAAAGGCAATGAAAACCGGAAAT
CGAAAAGCTGGCCAAACAGCCGGAATACGCCCTACCTGAAAACAGCTTCAGACGGCAAGAACGTCGATGGAAAACAGGTGCAGCTT
GTCTACGACAAGTGGGACTATAAAGCCGAAGGCTGACCGGAGCCGGAGCCGCAATTAATCGCACTGGCCGTTACCGTGGTCACT
CAGGCGCAGGAACCGGAGCCGATTGGGATTAACCGTGGCGCCCGCCGCAACCCGATGCAGCATTGCGCTCTTTGGCCAGCCA
GGCTCCGATCGTTATCAACAACAAAGGCAATATCGTAAACACCTGAAAAGAGCTGGGCAAGGACGACCGGTGAAAATCTG
GTGGTTGCCCGCTACCGCAGGCGTAGCCGACAAAATCGCCGCTCGGCAATGAAACATGTGAGCGATAAGCGATGGATCAACA
ACCTGACCGTCAACCTGGCCAAATGCGGCGAGTCCGCCATGATTAATACCGCTGTCAACGGCGGCGAGCTGAAAAGCAACTGGA
AGCGAATATCTTGGCGCTTTGGTGAATATGCGCATGGAGAGGCAAGTAAAATCAAACATTAAGTACGACTACATAGTC
CACAAGATGCCCATGCCATAGCGGCTGTGCGGACGCGGCGGCAATAAGGCAAGTGTGAGGATGGCGGATAGGCGCTGCAG
TGGTGTGAGATTTGTTGTTGAGGCTTTGGTTAAGAACTGATTTTCACTGCTATGAGTGCAGCCGAAAATCGAAAATCTAAAAGCGAA
GATTACTGCTTAACTAAAACCTGGTGGCGGCACTGCTGCTGCGGCTGAGGCGGGGATGTAATACAGCGGCGAATCGGCGCAG
ATAGCGGTGGAGAAATAACTTTGATCCTAGATGCTTGGTGAAGTGTGATGAATTTCAAAGGAAACAAAATGGATAC
GTGAAAATCTGAAGAAATCGAGAAGTTTGGCTCTCAGACAGGATTTATCCAATTAATCGGTGATATACAGAGTTTGTACA

AGCACAGACCGCTGCCGATCACCTGTTTGCTTTGCTGGGTGTGGTTCGGGTATCGGTGAATCGATACAGGCCTATAAAGTAGCG
AAAGCGGCAAAAAATTTACAAGGCATGAAAAAGCCTTGGACAAGGCAGCAACCCTTCCCACTGCACAGGGCTATGTCAGTAAAA
CCAAAATCAAAATCGGTCAAACCTGAATTAAAGGGTACTGCAGCAACTGACAAACAATTGCTGAAAGCTATTGGCGAAGGAAGGGA
CACGACAGGTAAAATGACCGAGCAGTTATTTGACTCTTTAGCTAAACAAAATGGCTTCAGAGTCTTTTCGGGCGGCAAAATACGGC
GGAAATAACGGTTTTGATCATGTATGGCAGGCTGCCGATGGTAGTGTGTTGTTTATTGTTAGAAAAGTAAGCAGATTAGGAACGGTA
CGGTACAGCTGAATCCGAATGGTCCGGGTGGATACCGCAGATGAGTCTGAAATGGATTAAACAAGTTGTA AAAAGTTTACCTGA
TGGTAGTCTCTGCTAAGGCAGTTGTCTTAAAAGCAAATCAGAACGGCAAAATAAAAACGGCAATAGCAGGCGTTGATCGTCAACA
GGTAAGGCCGTTATTCTTTCTGTCAAAGTCTCTTCTAAAACCAATATAAGGAGATAA

SEQ ID NO: 76

MCXXDKRSYKTKGWYKLVKHTVEIKEHKNAKADPVLSASQGIEIKSGGNI GAHATLFDAPRGSVKIEAGRGLVLYAVEDLN YDKL
DTRTKRKF IGITYDKVHDTTTHMTALPSRVVAESANLQSGWDAKLQGTQFETTLGGAAIRAGVGDQARADAKIILEGIKSSVR
TETVSSKSALWQKQAGRGSNIETLQLPSFTGVSAPVLSAPGGYIVDIPKGNLKEIEKLAKQPEYAYLKQLQTAKNVDWKVQL
VYDKWYKAEGLTGAGAAIIALAVTVVTSGAGTGAVLGLNAAAAATDAAFASLASQASVSFINNKGNIGNTLKELRSTVKNL
VVAATAGVADKIGASALNNVSDKQWNNLTVNLANAGSAALINTAVNGGSLKDNLEANI LAAALVNTAHGEAASKIKQLDQHYIV
HKIAHA IAGCAAAAANKGKQDGAIGA AVGEIVGEALVKNDFSRMSATEIEKSKAKITAYSKLVAGTASAVVGGDVNTAANAQ
IAVENNTLYPRCVGAKCDEFQKEQQWIRENPEEYREVL LQTGFIP IIGDIQS FVQAQTAADHL FALLGVVPGIGESI QAYKVA
KAAKNLQGMKKALDKAATVATAQGYVSKTIKIQGTEL RVTAATDKQLLKAIGEGRDTTKMTEQLFDSLAKQNGFRVLSGGKYG
GNNGFHDHWQAADGSVVLIVESKQIRNGTVQLNPNAGGYTQMSREWIKQVVKSLPDGSPAKAVVLLKANQNGKLTAIAGVDRQT
GKAVILSVKVPSTNIRR

SEQ ID NO: 77

GTGAGCATTAGCGCGCGTATGCCAATGAAAACAGCCGCATCTGCTGAGCACCACGGATATCAGTTCGGAAAAACGGCAAAATCA
AACTGCAATCTACGGCGACAGTTCTACTACGGCGGACAGGGTGAGCTCTACACCTTCGATAAACGCAGCTATAAAAACCGGTAA
GTGGTACAAAATAAAACATGTTACTGAAATCAAAGAGCATAAAAACGCCAAAAGCCGACCCGGTGAGCCTCAGTGCCTCACAAGGT
ATTGAAATCAAATCCGGCGGCAATATCGGTGCCACGCCACCTTGTGATGCACCCCGGGCTCCGTTAAAAATCGAAGCCGGAC
GTGGGCTGGTTCTCTATGCCGTGGAAGATCTCAACTACGACAAAATTTGACACCCGTACCAAGCGCAAAATTTATCGGCATTACCTA
CGACAAGGTGCACGACACCACCCACCCATGAAAACCGCCCTGCCCTCAAGGGTAGTTGCAGAAATCGGCCAACCTGCAATCA
GGCTGGGACGCCAAACTGCAAGGCACCCAGTTTGA AACCA CGCTGGGCGGCGCAGCCATCCGTGCAGGTGTAGGGCATCAGGCAC
GAGCAGATGCAAGATTATTTCTGAAGGCATCAAAGTAGTGTGCGCACTGAAAACAGTAAGCAGTAGCAAAATCTGCCCTCTGGCA
GAAACAGGCCGGACGCGGACAGCAATATCGAAACCTTCAACTGCCAAGTTTACAGGCTCCGTTGCGCCCGTACTCTCTGCCCCC
GGCGGCTATATCGTTGATATCCGAAAGGCAATCGAAAACCGAAAATCGAAAAGCTGGCCAAACAGCCCGAATACGCCCTACCTGA
AACAGCTTCAGACGGCCAAGAACGTCGATNNNNNCACACTTAATTAATTA

SEQ ID NO: 78

MSISAPYANENSRI LLSTDISSENGKIKIQSYGDQFYAAGQGLYTFDKRSYKTKGWYKLVKHTVEIKEHKNAKADPVLSASQ
G I EIKSGGNI GAHATLFDAPRGSVKIEAGRGLVLYAVEDLN YDKLDTRTKRKF IGITYDKVHDTTTHMTALPSRVVAESANLQ
S GWDAKLQGTQFETTLGGAAIRAGVGDQARADAKIILEGIKSSVR TETVSSKSALWQKQAGRGSNIETLQLPSFTGVSAPVLSAP
GGYIVDIPKGNLKEIEKLAKQPEYAYLKQLQTAKNVDXXTHLIN

REIVINDICAÇÕES

1. Polipeptídeo caracterizado por compreender uma seqüência de aminoácidos que tem pelo menos 75% de identidade de seqüência com uma ou mais das SEQ ID N^{os}: 32, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78.
2. Polipeptídeo, de acordo com a reivindicação 1, caracterizado por compreender uma ou mais das seqüências de aminoácidos SEQ ID N^{os}: 32, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78.
3. Polipeptídeo caracterizado por compreender um fragmento de pelo menos 7 aminoácidos consecutivos a partir de uma ou mais das SEQ ID N^{os}: 32, 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78.
4. Polipeptídeo, de acordo com a reivindicação 3, caracterizado pelo fato de que o fragmento compreende um epítipo de célula-T ou célula-B a partir da SEQ ID N^o: seqüência de aminoácido.
5. Anticorpo caracterizado por se ligar ao polipeptídeo de qualquer uma das reivindicações 1, 2, 3 ou 4.
6. Anticorpo, de acordo com a reivindicação 5, caracterizado pelo fato de que o anticorpo é um anticorpo monoclonal.
7. Ácido nucléico caracterizado por compreender uma seqüência de nucleotídeo que tem pelo menos 75% de

identidade de seqüência com uma ou mais das SEQ ID N^{os}: 31, 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77.

5 8. Ácido nucléico, de acordo com a reivindicação 7, caracterizado por compreender uma seqüência de nucleotídeos selecionada a partir das SEQ ID N^{os}: 31, 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65,
10 67, 69, 71, 73, 75, 77.

9. Ácido nucléico caracterizado pelo fato de que pode hibridizar-se ao ácido nucléico, de acordo com a reivindicação 8, sob condições de severidade elevada.

15 10. Ácido nucléico caracterizado por compreender um fragmento de 10 ou mais nucleotídeos consecutivos a partir de uma ou mais das SEQ ID N^{os}: 31, 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77.

20 11. Ácido nucléico caracterizado pelo fato de que codifica o polipeptídeo de qualquer uma das reivindicações 1, 2, 3 ou 4.

12. Composição caracterizada por compreender: (a) polipeptídeo, anticorpo, e/ou ácido nucléico de qualquer
25 reivindicação anterior; e (b) um veículo farmacologicamente aceitável.

13. Composição, de acordo com a reivindicação 12, caracterizada por compreender ainda um adjuvante de vacina.

30 14. Ácido nucléico, polipeptídeo, ou anticorpo, de acordo com qualquer uma das reivindicações 1, 2, 3, 4,

5, 6, 7, 8, 9, 10 ou 11, caracterizado por ser para uso como medicamento.

15. Método de tratar um paciente, caracterizado por compreender administrar ao paciente uma quantidade 5 terapêuticamente eficaz da composição da reivindicação 12.

16. Uso de ácido nucléico, polipeptídeo ou anticorpo, de qualquer uma das reivindicações 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 ou 11, caracterizado por ser na fabricação de um medicamento para tratar ou prevenir doença 10 e/ou infecção causada por *Neisseria meningitidis*.

17. Método, de acordo com a reivindicação 15, ou o uso, de acordo com a reivindicação 16, caracterizado pelo fato de ser para prevenir meningite meningocócica.

POLIPEPTÍDEOS A PARTIR DE NEISSERIA MENINGITIDIS

São reveladas várias proteínas meningocócicas específicas. A invenção provê polipeptídeos relacionados, ácidos nucléicos, anticorpos e métodos. Todos esses podem
5 ser utilizados na medicina para tratar ou prevenir doença e/ou infecção causada por meningococo, como meningite bacteriana.