

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: <b>1999.04.30</b>	(73) Titular(es): <b>NOVARTIS AG</b> <b>LICHTSTRASSE 35 4056 BASEL</b> <b>J. CRAIG VENTER INSTITUTE, INC.</b>	<b>CH</b> <b>US</b>
(30) Prioridade(s): <b>1998.05.01 US 83758 P</b> <b>1998.07.31 US 94869 P</b> <b>1998.09.02 US 98994 P</b> <b>1998.09.02 US 99062 P</b> <b>1998.10.09 US 103749 P</b>	(72) Inventor(es): <b>CESIRA GALEOTTI</b> <b>RINO RAPPUOLI</b> <b>GUIDO GRANDI</b> <b>VEJA MASIGNANI</b> <b>CLAIRE FRASER</b>	<b>IT</b> <b>IT</b> <b>IT</b> <b>IT</b> <b>US</b>
(43) Data de publicação do pedido: <b>2008.07.16</b>	(74) Mandatário: <b>LUÍSA MARIA FERREIRA GUERREIRO</b> <b>RUA RAUL PROENÇA, 3 2820-478 CHARNECA DA</b> <b>CAPARICA</b>	<b>PT</b>
(45) Data e BPI da concessão: <b>2015.03.04</b> <b>134/2015</b>		

(54) Epígrafe: **ANTIGÉNIOS E COMPOSIÇÕES DA NEISSERIA MENINGITIDIS**

(57) Resumo:

A INVENÇÃO PROPORCIONA PROTEÍNAS DE NEISSERIA MENINGITIDIS, INCLUINDO AS SEQUÊNCIAS DE AMINOÁCIDOS E AS CORRESPONDENTES SEQUÊNCIAS DE NUCLEÓTIDOS. AS PROTEÍNAS ESTÃO PREVISTAS PARA SER ANTIGÉNIOS ÚTEIS PARA VACINAS E/OU DIAGNÓSTICOS.

## RESUMO

### **ANTIGÉNIOS E COMPOSIÇÕES DA *NEISSERIA MENINGITIDIS***

A invenção proporciona proteínas de *Neisseria meningitidis*, incluindo as sequências de aminoácidos e as correspondentes sequências de nucleótidos. As proteínas estão previstas para ser antígenos úteis para vacinas e/ou diagnósticos.

## DESCRIÇÃO

### ANTIGÉNIOS E COMPOSIÇÕES DA *NEISSERIA MENINGITIDIS*

#### Campo da Invenção

Esta invenção diz respeito a antígenos das espécies bacterianas *Neisseria meningitidis* e *Neisseria gonorrhoeae*.

#### Antecedentes da invenção

O patógeno humano *Neisseria meningitidis* é um diplococo gram negativo imóvel. Coloniza a faringe causando meningite e, ocasionalmente, septicemia na ausência de meningite. Este patógeno assemelha-se bastante a *N.gonorrhoeae*, apesar de uma característica que diferencia claramente o meningococo do gonococo ser a presença de uma cápsula polisacarídica que está presente em todos os meningococos patogênicos.

A *N.meningitidis* causa tanto doença endêmica quanto epidêmica. Nos EUA a incidência atinge 0,6-1 por 100,000 pessoas/ano, e pode ser muito superior durante os surtos de doença (ver Lieberman et al. (1996), Safety and Immunogenicity of a Serogroups A/C *Neisseria meningitidis* Oligosaccharide-Protein Conjugate Vaccine in Young Children. JAMA 275(19):1499-1503; Schuchat et al (1997) Bacterial Meningitis in the United States in 1995. N Engl J Med 337(14):970-976). Nos países em vias de desenvolvimento, as taxas de incidência de doença endêmica são muito superiores e durante períodos de doença epidêmica as taxas de incidência podem alcançar 500 casos por 100,000 pessoas/ano. A mortalidade é extremamente elevada, atingindo 10-20% nos EUA e sendo bastante mais elevada nos países em vias de desenvolvimento. Após a introdução da vacina conjugada contra *Haemophilus influenzae*, a infecção por *N.meningitidis* constitui agora a maior causa de meningite

bacteriana em todas as idades, nos EUA (Schuchat et al (1997) supra).

Foram já identificados 12 serogrupos de *N.meningitidis* com base no polissacárido capsular do organismo. O grupo A é o patógeno mais frequentemente implicado na doença epidémica na África Subsaariana. Os serogrupos B e C são responsáveis pela grande maioria dos casos nos EUA. e na maioria dos países desenvolvidos. Os serogrupos W135 e Y são responsáveis pelos casos restantes nos EUA. e nos países desenvolvidos. A vacina meningocócica usada presentemente é uma vacina polisacarídica tetravalente composta pelos serogrupos A, C, Y e W135. Apesar de eficaz em adolescentes e adultos, conduz a uma fraca resposta imunitária e a uma duração de proteção curta e não pode ser usada em crianças [p.ex. Morbidity and Mortality weekly report, Vol.46, No. RR-5 (1997)]. Isto deve-se ao facto de os polissacáridos serem antígenos independentes das células T que induzem uma resposta imunitária fraca e que não pode ser reforçada através de imunização repetida. No seguimento do sucesso obtido com as vacinas utilizadas contra *H.influenzae*, têm vindo a ser desenvolvidas vacinas conjugadas contra os serogrupos A e C, as quais estão em fase final de ensaios clínicos (Zollinger WD, "New and Improved Vaccines Against Meningococcal Disease". Em: New Generation Vaccines, supra, pp. 469-488; Lieberman et al (1996), supra; Costantino et al (1992), Development and phase I clinical testing of a conjugate vaccine against meningococcus A and C. Vaccine 10:691-698).

No entanto, o meningococo B (menB) continua a constituir um problema. Este serotipo é presentemente responsável por aproximadamente 50% do total de casos de meningite nos EUA., Europa e América do Sul. A abordagem polisacarídica não pode ser utilizada porque o polissacárido capsular do menB é um

polímero de ácido N-acetil-neuramínico ligado em  $\alpha(2-8)$  que também está presente no tecido dos mamíferos. Isto resulta em tolerância ao antígeno; na verdade, se uma resposta imunitária fosse desencadeada, seria contra o próprio organismo e consequentemente indesejável. De modo a evitar a indução de autoimunidade e a induzir uma resposta imunitária protetora, o polissacárido capsular foi, por exemplo, quimicamente modificado, substituindo-se os grupos de N-acetil por grupos de N-propionil e deixando a antigenicidade específica inalterada (Romero & Outschoom (1994), Current status of Meningococcal group B vaccine candidates: capsular or non-capsular? Clin Microbiol Rev 7(4):559-575).

Certas abordagens alternativas às vacinas contra o menB têm utilizado misturas complexas de proteínas da membrana externa (OMPs), que contêm OMPs isoladas ou OMPs enriquecidas com porinas, ou a que foram retiradas OMPs da classe 4 que se acredita poderem induzir anticorpos que bloqueiam a atividade bactericida. Esta abordagem produz vacinas que ainda não estão bem caracterizadas. Estas vacinas têm a capacidade de proteger contra a estirpe homóloga mas não são geralmente eficazes quando existem muitas variantes antigénicas das proteínas da membrana externa. Para ultrapassar a variabilidade antigénica, foram já construídas vacinas multivalentes contendo até nove porinas diferentes (por exemplo, Poolman JT (1992), Development of a meningococcal vaccine. Infect. Agents Dis. 4: 13-28). As proteínas adicionais para utilização em vacinas de membrana externa têm sido as proteínas opa e opc, mas nenhuma destas abordagens tem sido capaz de ultrapassar a variabilidade antigénica (p.ex. Ala'Aldeen & Borriello (1996), The meningococcal transferrin-binding proteins 1 and 2 are both surface exposed and generate bactericidal antibodies capable of killing homologous and heterologous strains. Vaccine 14 (1): 49-53).

Encontra-se disponível um certo volume de dados de sequência para os genes e proteínas do meningococo e gonococo (por exemplo, EP-A-0467714, WO96/29412), mas a sequenciação não está de modo algum completa. Outras proteínas de menB poderão incluir as listadas em WO97/28273, WO96/29412, WO95/03413, US 5439808, US 5879686, Rokbi et al. (1997), Infect Immun. 65:55-63 e WO97/13860.

O fornecimento de mais sequências pode traduzir-se numa oportunidade de identificar proteínas segregadas ou proteínas expostas à superfície que constituam presumíveis alvos do sistema imunitário e que não sejam antigenicamente variáveis. Por exemplo, algumas das proteínas identificadas poderão ser componentes de vacinas eficazes contra o meningococo B, outras poderão ser componentes de vacinas contra todos os serotipos do meningococo, e outras ainda poderão ser componentes de vacinas contra todas as *Neisseriae* patogénicas, incluindo *Neisseria meningitidis* ou *Neisseria gonorrhoeae*. As sequências específicas de *N.meningitidis* ou *N.gonorrhoeae* que são mais conservadas constituirão as sequências preferidas.

Constitui deste modo um objetivo da invenção fornecer sequências de ADN de *Neisseria* que codificam para proteínas que são antigénicas ou imunogénicas.

#### Breve descrição das Figuras

A Fig. 1 ilustra os produtos de expressão e purificação de proteínas obtidos a partir do ORF 287 previsto, após clonagem e expressão em *E. coli*.

A Fig. 2 ilustra o gráfico de hidrofiliicidade, o índice antigénico e as regiões AMPHI dos produtos de expressão de proteínas obtidos a partir do ORF 287 previsto, após clonagem e expressão em *E. coli*.

A Fig. 3 apresenta uma comparação dos alinhamentos das sequências de aminoácidos do ORF 287 para diferentes estirpes de *Neisseria*. O sombreado mais escuro indica regiões de homologia, e o sombreado mais claro indica a conservação de aminoácidos com características similares. A Figura demonstra um elevado grau de conservação entre as diversas estirpes, confirmando-se deste modo a sua utilidade como antigénio para vacinas e meios de diagnóstico.

### **Descrição da invenção**

A invenção fornece proteínas compreendendo as sequências de aminoácidos de *N.meningitidis* e as sequências de aminoácidos de *N.gonorrhoeae* que se encontram expostas no Exemplo.

A invenção fornece ainda proteínas compreendendo sequências homólogas (i.e., que apresentam identidade de sequência) com as sequências de aminoácidos de *N.meningitidis* revelados no exemplo. Dependendo da sequência em particular, o grau de homologia (identidade de sequência) é preferencialmente superior a 90% (por exemplo, 90%, 95%, 99% ou mais). Estas proteínas incluem mutantes e variantes alélicas das sequências reveladas no exemplo. Tipicamente, uma identidade de 50% ou mais entre duas proteínas é considerada como uma indicação de equivalência funcional. A identidade entre proteínas é preferencialmente determinada por meio do algoritmo de pesquisa de homologia de Smith-Waterman, tal como implementado no programa MPSRCH (Oxford Molecular), usando uma pesquisa de gaps de afinidade com parâmetros de *gap open penalty* = 12 e *gap extension penalty* = 1.

As proteínas da invenção poderão, como é evidente, ser preparadas por diversos métodos (por exemplo, expressão recombinante, purificação a partir de cultura celular, síntese química, etc.) e sob variadas formas (por exemplo, forma

nativa, proteínas de fusão, etc.). As proteínas serão preferencialmente preparadas sob uma forma substancialmente pura ou isolada (i.e., substancialmente isenta de outras proteínas da célula hospedeira de *N.meningitidis* ou de *N.gonorrhoeae*).

De acordo com um outro aspeto, a invenção fornece anticorpos que se ligam a essas proteínas. Estes anticorpos poderão ser policlonais ou monoclonais e a sua produção poderá ser efetuada através de qualquer meio adequado.

De acordo com um aspeto adicional, a invenção fornece um ácido nucleico compreendendo as sequências de nucleótidos de *N.meningitidis* e as sequências de nucleótidos de *N.gonorrhoeae* que se encontram expostas no Exemplo.

De acordo com ainda outro aspeto da invenção, é fornecido um ácido nucleico que se hibridiza com as sequências aqui fornecidas. As condições de hibridização são aqui apresentadas.

De acordo com um outro aspeto da invenção, é fornecido um ácido nucleico codificando para as proteínas e fragmentos de proteínas da invenção.

A invenção fornece ainda um ácido nucleico compreendendo sequências complementares às acima descritas (por exemplo, para uso como sequências anti-senso ou de sonda).

O ácido nucleico de acordo com a invenção poderá, evidentemente, ser preparado de diversas formas (por exemplo, por síntese química de parte ou do seu todo, a partir de bibliotecas genómicas ou de cADN, a partir do próprio organismo, etc.) e poderá encontrar-se sob diversas formas



(por exemplo cadeia simples, cadeia dupla, vetores, sondas, etc.).

Adicionalmente, o termo "ácido nucleico" inclui ADN e ARN e ainda os seus análogos, tais como aqueles que contêm estruturas primárias modificadas, além de ácidos nucleicos peptídicos (PNAs), etc..

De acordo com um aspeto adicional, a invenção fornece vetores que compreendem as sequências nucleotídicas da invenção (por exemplo vetores de expressão) e células hospedeiras transformadas com tais vetores.

Segundo um outro aspeto da invenção, são fornecidas composições compreendendo uma proteína, anticorpo e/ou ácido nucleico de acordo com o invento. Estas composições poderão ser utilizadas como vacinas, por exemplo, ou como reagentes para diagnóstico, ou como composições imunogénicas.

A invenção fornece igualmente um ácido nucleico, proteína ou anticorpo de acordo com a invenção para uso como medicamentos (por exemplo, como vacinas) ou como reagentes para diagnóstico. A invenção proporciona ainda o uso do ácido nucleico, proteína ou anticorpo de acordo com a invenção para o fabrico de: (i) um medicamento para prevenção de uma infeção devida a *Neisseria*; (ii) um reagente para diagnóstico, destinado a detetar a presença de *Neisseria* ou de anticorpos dirigidos contra *Neisseria* ou (iii) para estimular a produção de anticorpos. A referida *Neisseria* poderá pertencer a qualquer das espécies ou estirpes de *Neisseria* (como *N.gonorrhoeae*) mas corresponde preferencialmente a *N.meningitidis*, especialmente da estirpe B ou estirpe C.

De acordo com outros aspetos, a invenção proporciona diversos processos.

É fornecido um processo para a produção das proteínas da invenção, compreendendo o passo de cultura de uma célula hospedeira de acordo com a invenção sob condições que induzem a expressão de proteínas.

É fornecido um processo para a detecção de polinucleótidos do invenção, compreendendo os passos de: (a) colocação em contacto de uma sonda de ácido nucleico de acordo com a invenção com uma amostra biológica, sob condições de hibridização, para formar duplexes; e (b) detecção dos ditos duplexes.

É fornecido um processo para a detecção das proteínas da invenção, compreendendo os passos de: (a) colocação em contacto de um anticorpo de acordo com a invenção com uma amostra biológica sob condições adequadas à formação de complexos de anticorpo-antigénio; e (b) detecção dos ditos complexos.

Tendo acabado de descrever a invenção de forma geral, o mesmo será melhor compreendido tomando como referência os exemplos que se seguem, os quais são fornecidos com fins ilustrativos e não pretendem limitar o âmbito do presente invenção, a não ser que de outra forma seja especificado.

## **Metodologia - Resumo dos procedimentos e técnicas padronizadas**

### Generalidades

Esta invenção fornece sequências nucleotídicas de *Neisseria meningitidis* menB e as sequências de aminoácidos por estas codificadas. A partir das sequências expostas poderão

produzir-se ensaios com sondas de ácido nucleico e cassetes e vetores de expressão. Os vetores de expressão podem ser transformados em células hospedeiras de modo a produzir proteínas. Os polipéptidos purificados ou isolados (que também podem ser sintetizados quimicamente) podem ser usados para produzir anticorpos destinados à detecção de proteínas de menB. Ainda, as células hospedeiras ou extratos podem ser utilizadas em ensaios biológicos para isolar agonistas ou antagonistas. Adicionalmente, com estas sequências é possível implementar buscas para identificar moldes abertos de leitura e identificar sequências de aminoácidos. As proteínas podem também ser usadas em composições imunogénicas, composições antigénicas e como componentes de vacinas.

A colocação em prática da presente invenção utilizará, a não ser que indicado de outro modo, técnicas convencionais de biologia molecular, microbiologia, ADN recombinante e imunologia, as quais são bem conhecidas neste campo técnico. Tais técnicas encontram-se descritas em detalhe na literatura, por exemplo, em Sambrook, *Molecular Cloning: A Laboratory Manual*, 2<sup>a</sup> Ed. (1989); *ADN Cloning*, Volumes I e II (ed. D.N. Glover, 1985); *Oligonucleotide Synthesis* (ed. M.J. Gait, 1984); *Nucleic Acid Hybridization* (eds. B.D. Hames & S.J. Higgins, 1984); *Transcription and Translation* (eds. B.D. Hames & S.J. Higgins, 1984); *Animal Cell Culture* (ed. R.I. Freshney, 1986); *Immobilized Cells and Enzymes* (IRL Press, 1986); B. Perbal, *A Practical Guide to Molecular Cloning* (1984); a série *Methods in Enzymology* (Academic Press, Inc.), especialmente os volumes 154 e 155; *Gene Transfer Vectors for Mammalian Cells* (eds. J.H. Miller e M.P. Calos, 1987, Cold Spring Harbor Laboratory); Mayer e Walker, eds. (1987), *Immunochemical Methods in Cell and Molecular Biology* (Academic Press, London); Scopes (1987), *Protein Purification: Principles and Practice*, 2<sup>a</sup> Ed. (Springer-Verlag, N.Y.) e *Handbook of*

Experimental Immunology, Volumes I-IV (eds. D.M. Weir e C.C. Blackwell, 1986).

Serão usadas nesta especificação as abreviaturas convencionais para designação de nucleótidos e aminoácidos.

Todas as publicações, patentes e pedidos de patentes aqui citados são incorporados na sua totalidade para fins de referência.

### Sistemas de expressão

As sequências nucleotídicas de *Neisseria menB* poderão ser expressas Numa variedade de sistemas de expressão diferentes; por exemplo, os usados com células de mamífero, células de plantas, baculovírus, bactérias e leveduras.

#### i. Sistemas de Mamíferos

Os sistemas de expressão de mamíferos são bem conhecidos neste campo técnico. Um promotor de mamífero corresponde a qualquer sequência de ADN que tem a capacidade de se ligar a uma ARN polimerase de mamífero e iniciar a transcrição a jusante (3') de uma sequência codificante (por exemplo, um gene estrutural) para mRNA. Um promotor terá uma região de iniciação da transcrição, que geralmente está colocada em posição proximal relativamente à extremidade 5' da sequência codificante, e uma TATA box, geralmente localizada 25-30 pares de bases (pb) a montante do local de iniciação da transcrição. Pensa-se que a TATA box dirija a ARN polimerase II para iniciar a síntese de ARN no local correto. Um promotor de mamífero conterá igualmente um elemento promotor a montante, geralmente localizado entre 100 a 200 pb a montante da TATA box. Um elemento promotor a montante determina a rapidez à qual se inicia a transcrição e pode atuar em ambas as orientações (Sambrook et al. (1989) "Expression of Cloned Genes in

Mammalian Cells", em *Molecular Cloning: a Laboratory Manual*, 2ª ed.).

Os genes de vírus que infetam mamíferos são frequentemente expressos em frequência elevada e possuem uma larga variedade de hospedeiros; assim, as sequências que codificam genes de vírus que infetam mamíferos proporcionam sequências promotoras particularmente úteis. Como exemplos destes promotores encontram-se o promotor precoce de SV40, o promotor LTR do vírus do tumor mamário de ratinho, o promotor principal tardio de adenovírus (Ad MLP) e o promotor do vírus herpes simples. Adicionalmente, as sequências derivadas de genes não virais, tal como o gene de metalotioneína murino, proporcionam igualmente sequências promotoras úteis. A expressão poderá ser constitutiva ou regulada (indutível). Dependendo do promotor selecionado, muitos promotores podem ser indutíveis utilizando substratos conhecidos, tal como sucede usando o promotor do vírus do tumor mamário de ratinho (MMTV), em associação ao elemento de resposta a glucocorticóides (GRE) que é induzido por glucocorticóides em células transformadas hormono-responsivas (ver por exemplo a Patente dos EUA N°. 5,783,681).

A presença de um elemento potenciador, combinado com os elementos promotores acima descritos, aumentará geralmente os níveis de expressão. Um potenciador é uma sequência reguladora de ADN que tem a capacidade de estimular a transcrição em até 1000 vezes quando ligada a promotores homólogos ou heterólogos, sendo a síntese iniciada no local normal de iniciação do ARN. Os potenciadores são igualmente ativos quando colocados a montante ou a jusante do local de iniciação da transcrição, na orientação normal ou invertida, ou distanciados em mais do que 1000 nucleótidos do promotor (Maniatis et al. (1987) *Science* 236:1237; Alberts et al. (1989) *Molecular Biology of the Cell*, 2ª ed.). Os elementos

potenciadores derivados de vírus poderão ser particularmente úteis, uma vez que estes possuem geralmente uma gama mais alargada de hospedeiros. Os exemplos incluem o potenciador do gene precoce de SV40 (Dijkema et al. (1985) EMBO J. 4:761) e o potenciador/promotores derivados da longa sequência repetitiva terminal (LTR) do vírus do sarcoma de Rous (Gorman et al. (1982b) Proc. Natl. Acad. Sci. 79:6777) e do citomegalovírus humano (Boshart et al. (1985) Cell 41:521). Adicionalmente, alguns potenciadores são reguláveis e tornam-se ativos apenas na presença de um indutor, tal como uma hormona ou um ião metálico (Sassone-Corsi e Borelli (1986) Trends Genet 2:215; Maniatis et al. (1987) Science 236:1237).

Uma molécula de ADN poderá ser expressa intracelularmente em células de mamíferos. Uma sequência promotora poderá encontrar-se diretamente ligada à molécula de ADN, caso em que o primeiro aminoácido na extremidade N da proteína recombinante será sempre uma metionina, a qual é codificada pelo codão de iniciação ATG. Se desejado, a extremidade N será cortada da proteína por incubação *in vitro* com brometo de cianogénio.

Alternativamente, as proteínas estranhas poderão também ser segregadas pela célula para o meio de cultura através da criação de moléculas quiméricas de ADN codificando para uma proteína de fusão composta por um fragmento de sequência leader que proporcione a secreção da proteína estranha em células de mamífero. De preferência, existirão locais de processamento codificados entre o fragmento leader e o gene estranho que poderão ser submetidos a corte *in vivo* ou *in vitro*. O fragmento de sequência leader codifica geralmente para um péptido de sinal composto por aminoácidos hidrofóbicos que dirigem a secreção da proteína pela célula. O leader tripartido de adenovírus constitui um exemplo de uma sequência

leader que proporciona a secreção de uma proteína estranha por células de mamífero.

Geralmente, as sequências de terminação da transcrição e de poliadenilação que são reconhecidas pelas células de mamíferos correspondem a regiões reguladoras localizadas a 3' relativamente ao codão de paragem da tradução, e que deste modo flanqueiam, juntamente com os elementos promotores, a sequência codificante. A extremidade 3' do mRNA maduro é formada, após a transcrição, por corte e poliadenilação específicos de local (Birnstiel et al. (1985) *Cell* 41:349; Proudfoot e Whitelaw (1988) "Termination and 3' end processing of eukaryotic RNA". Em: *Transcription and splicing* (eds. B.D. Hames e D.M. Glover); Proudfoot (1989), *Trends Biochem. Sci.* 14:105). Estas sequências dirigem a transcrição de um mRNA que pode ser traduzido para o polipéptido codificado pelo ADN. Como exemplos de terminadores da transcrição/sinais de poliadenilação encontram-se os derivados do SV40 (Sambrook et al. (1989) "Expression of cloned genes in cultured mammalian cells". Em: *Molecular Cloning: a Laboratory Manual*).

Habitualmente, os componentes acima descritos, compreendendo um promotor, um sinal de poliadenilação e uma sequência de terminação da transcrição, são reunidos em construções de expressão. Os potenciadores, intrões com locais dadores e aceitadores de splice funcionais e sequências leader podem igualmente incluir-se numa construção de expressão, se desejado. As construções de expressão são frequentemente mantidas numa replicação, tal como um elemento extracromossómico (por exemplo, plasmídeos) que tem a capacidade de se manter de forma estável num hospedeiro, tal como células de mamíferos ou bactérias. Os sistemas de replicação em mamíferos incluem os derivados de vírus que infetam animais, os quais requerem fatores trans-actantes para a sua replicação. Por exemplo, os

plasmídeos contendo os sistemas de replicação de papovavírus, tal como o SV40 (Gluzman (1981) *Cell* 23:175) ou poliomavírus, replicam-se até um número extremamente elevado de cópias na presença do antigénio T viral apropriado. Como exemplos adicionais de replicões de mamífero encontram-se os derivados do papilomavírus bovino e do vírus de Epstein-Barr. Adicionalmente, o replicão poderá possuir dois sistemas de replicação, permitindo deste modo a sua manutenção, por exemplo, em células de mamíferos para a expressão e num hospedeiro procariótico para a clonagem e amplificação. Os exemplos de tais vetores de vaivém de mamíferos-bactérias incluem o pMT2 (Kaufman et al. (1989) *Mol. Cell Biol.* 9:946) e pHEBO (Shimizu et al. (1986) *Mol. Cell Biol.* 6:1074).

O procedimento de transformação usado depende do hospedeiro a ser transformado. Os métodos para introdução de polinucleótidos heterólogos em células de mamíferos são bem conhecidos da técnica e incluem os processos de transfecção mediada por dextrano, precipitação com fosfato de cálcio, transfecção mediada por polibreno, fusão de protoplastos, electroporação, encapsulação do(s) polinucleótido(s) em lipossomas e microinjeção direta do ADN nos núcleos.

As linhas celulares de mamífero disponíveis para uso como hospedeiros para expressão são conhecidas neste campo técnico e incluem muitas linhas celulares imortalizadas disponíveis através da American Type Culture Collection (ATCC), incluindo, mas de modo não limitativo, células de ovário de hamster chinês (CHO), células HeLa, células de rim de hamster bebé (BHK), células de rim de macaco (COS), células de carcinoma hepatocelular humano (por exemplo, Hep G2), e diversas outras linhas celulares.



## ii. Sistemas de Expressão Celular em Plantas

São já bem conhecidos neste campo técnico diversos tipos de sistemas de expressão genética em culturas de células de plantas e em plantas inteiras. Como exemplos de sistemas de expressão genética em células de plantas encontram-se os sistemas descritos em patentes como as patentes US 5693506, 5659122 e 5608143. Outros exemplos de expressão genética em culturas de células de plantas foram descritos por Zenk, *Phytochemistry* 30:3861-3863 (1991). Para além das referências acima apresentadas, poderão encontrar-se descrições de péptidos de sinal de proteínas de plantas em Vaulcombe et al., *Mol. Gen. Genet.* 209:33-40 (1987); Chandler et al., *Plant Molecular Biology* 3:407-418 (1984); Rogers, J. *Biol. Chem.* 260:3731-3738 (1985); Rothstein et al., *Gene* 55:353-356 (1987); Whittier et al., *Nucleic Acids Research* 15:2515-2535 (1987); Wirsal et al., *Molecular Microbiology* 3:3-14 (1989); Yu et al., *Gene* 122:247-253 (1992). Poderá encontrar-se uma descrição da regulação da expressão de genes de plantas pela fito-hormona ácido giberélico, e da secreção de enzimas induzida pelo ácido giberélico, em R.L. Jones e J. MacMillin, *Giberellins*, em: *Advanced Plant Physiology*, ed. Malcolm B. Wilkins (1984), Pitman Publishing Limited, London, págs. 21-52. Referências que descrevem outros genes de regulação metabólica: Sheen, *Plant Cell*, 2:1027-1038 (1990); Maas et al., *EMBO J.* 9:3447-3452 (1990); Benkel e Hickey, *Proc. Natl. Acad. Sci.* 84:1337-1339 (1987).

Tipicamente, por recurso a técnicas bem conhecidas neste campo, uma sequência polinucleotídica desejada é inserida numa cassette de expressão compreendendo elementos reguladores de genes concebidos para funcionamento em plantas. A cassette de expressão é inserida num vetor de expressão desejado juntamente com sequências acompanhantes a montante e a jusante da cassette de expressão adequada para expressão numa planta

hospedeira. As sequências acompanhantes terão origem plasmídica ou viral e proporcionam ao vetor as características necessárias para que este transfira o ADN de um hospedeiro de clonagem original, como uma bactéria, para uma planta hospedeira desejada. A construção básica do vetor de bactéria/planta proporcionará preferencialmente uma origem de replicação procariótica para uma larga gama de hospedeiros; um marcador procariótico de seleção; e, para as transformações com *Agrobacterium*, sequências de ADN T para a transferência mediada por *Agrobacterium* para os cromossomas da planta. Nos casos em que o gene heterólogo não se encontra facilmente acessível para detecção, a construção possuirá também, de preferência, um gene para um marcador de seleção apropriado para determinar se uma célula de planta foi transformada. Poderá encontrar-se uma exposição geral sobre marcadores de seleção, por exemplo para os membros da família da grama, em Wilmink e Dons, 1993, *Plant Mol. Biol. Repr.* 11(2):165-185.

É igualmente recomendável a presença de sequências adequadas para permitir a integração da sequência heteróloga no genoma da planta. Estas sequências poderão incluir sequências de transposição, e semelhantes, para a recombinação homóloga, assim como sequências Ti que permitem a inserção aleatória de uma cassete heteróloga de expressão num genoma de planta. Os marcadores procarióticos de seleção apropriados incluem a resistência a antibióticos como a ampicilina ou a tetraciclina. Poderão ainda encontrar-se presentes no vetor outras sequências de ADN que codifiquem para funções adicionais, tal como previsto neste campo técnico.

As moléculas de ácido nucleico da presente invenção poderão ser incluídas numa cassete de expressão de modo a possibilitar a expressão da(s) proteína(s) de interesse. Existirá habitualmente apenas uma cassete de expressão, se bem que a

existência de duas ou mais cassetes seja também possível. A cassette de expressão recombinante conterá, para além da sequência que codifica para a proteína heteróloga, os seguintes elementos: uma região promotora, sequências 5' não traduzidas da própria planta, codão de iniciação (dependendo de o gene estrutural se encontrar ou não equipado com este elemento), e uma sequência de terminação da transcrição e da tradução. A presença de locais específicos para enzimas de restrição a nível das extremidades 5' e 3' da cassette permitem a fácil inserção num vetor pré-existente.

Uma sequência codificante heteróloga poderá corresponder a qualquer das proteínas abrangidas pela presente invenção. A sequência codificando para a proteína de interesse codificará para um péptido de sinal que permita o processamento e translocação da proteína, segundo apropriado, e não conterá, em geral, qualquer sequência que possa resultar na ligação da proteína da invenção desejada a uma membrana. Uma vez que, na sua maior parte, a região de iniciação da transcrição dirá respeito a um gene que é expresso e translocado durante a germinação, ao se utilizar o péptido de sinal que proporciona a translocação será igualmente proporcionada a translocação da proteína de interesse. Deste modo, a(s) proteína(s) de interesse serão translocadas para o exterior das células nas quais são expressas e poderão ser eficientemente recolhidas. Tipicamente, as sementes segregam as proteínas através do aleurónio ou da camada escutelar do epitélio para o endosperma da semente. Apesar de não ser necessário que a secreção da proteína se processe para o exterior das células em que a proteína é produzida, tal facilita o isolamento e purificação da proteína recombinante.

Uma vez que a expressão final do produto genético desejado ocorrerá numa célula eucariótica, será desejável determinar se

qualquer porção do gene clonado contém sequências que serão processadas como intrões pelo equipamento de splicing do hospedeiro. Em caso afirmativo, poderá efetuar-se a mutagénese dirigida à região de "intrão" para evitar que se perca uma porção da mensagem genética sob a forma de falso código de intrão. Reed e Maniatis, *Cell* 41:95-105, 1985.

O vetor poderá ser microinjetado directamente nas células da planta, utilizando micropipetas para transferir mecanicamente o ADN recombinante. Crossway, *Mol. Gen. Genet.* 202:179-185, 1985. O material genético poderá igualmente ser transferido para a célula de planta utilizando polietilenoglicol; Krens et al., *Nature*, 296, 72-74, 1982. Outro método de introdução de segmentos de ácido nucleico consiste na penetração balística a alta velocidade por pequenas partículas que transportam o ácido nucleico no interior da matriz de pequenas esferas ou partículas ou à sua superfície; Klein et al., *Nature*, 327, 70-73, 1987 e Knudsen e Muller, 1991, *Planta*, 185:330-336, expõem o bombardeamento por partículas do endosperma de cevada para criar cevada transgénica. Outro método de introdução possível seria a fusão de protoplastos com outros elementos, como minicélulas, células, lipossomas ou outros corpos de superfície lipídica fundíveis; Fraley et al., *Proc. Natl. Acad. Sci. USA*, 79, 1859-1863, 1982.

O vetor poderá ainda ser introduzido nas células da planta por electroporação (Fromm et al., *Proc. Natl. Acad. Sci. USA* 82:5824, 1985). Nesta técnica, os protoplastos da planta são electroporados na presença de plasmídeos que contêm a construção genética. As biomembranas são reversivelmente permeabilizadas por impulsos elétricos de elevada força de campo, deste modo possibilitando a introdução dos plasmídeos. Os protoplastos electroporados da planta reconstroem a parede celular, dividem-se e formam o calo vegetal.

Todas as plantas a partir das quais é possível isolar e cultivar protoplastos para fornecer plantas inteiras regeneradas poderão ser transformadas por meio do presente invenção de modo a se recuperarem plantas completas que contenham o gene transferido. Sabe-se já que praticamente todas as plantas podem ser regeneradas a partir de células ou tecidos cultivados, incluindo, de modo não limitativo, todas as espécies principais de cana-de-açúcar, beterraba sacarina, algodão, árvores de fruto e outras árvores, legumes e vegetais hortícolas. Algumas das plantas apropriadas incluem, por exemplo, espécies pertencentes aos géneros *Fragaria*, *Lotus*, *Medicago*, *Onobrychis*, *Trifolium*, *Trigonella*, *Vigna*, *Citrus*, *Linum*, *Geranium*, *Manihot*, *Daucus*, *Arabidopsis*, *Brassica*, *Raphanus*, *Sinapis*, *Atropa*, *Capsicum*, *Datura*, *Hyoscyamus*, *Lycopersion*, *Nicotiana*, *Solanum*, *Petunia*, *Digitalis*, *Majorana*, *Cichorium*, *Helianthus*, *Lactuca*, *Bromus*, *Asparagus*, *Antirrhinum*, *Hererocallis*, *Nemesia*, *Pelargonium*, *Panicum*, *Pennisetum*, *Ranunculus*, *Senecio*, *Salpiglossis*, *Cucumis*, *Browaalia*, *Glycine*, *Lolium*, *Zea*, *Triticum* e *Sorghum*.

Os meios para regeneração variam consoante a espécie de planta considerada, mas em geral é fornecida em primeiro lugar uma suspensão de protoplastos transformados contendo cópias do gene heterólogo. É formado o tecido do calo, podendo induzir-se a formação de rebentos a partir do calo, os quais são subsequentemente enraizados. Alternativamente, a formação do embrião poderá ser induzida a partir da suspensão de protoplastos. Estes embriões germinam sob a forma de embriões naturais para formar plantas. O meio de cultura conterá geralmente vários aminoácidos e hormonas, tais como auxina e citoquininas. Será também vantajosa a adição de ácido glutâmico e de prolina ao meio, especialmente para as espécies como o milho e a alfafa. Os rebentos e as raízes desenvolvem-se normalmente em simultâneo. A eficiência da regeneração

dependerá do meio, do genótipo e da história da cultura. Se estas três variáveis forem controladas, a regeneração processar-se-á com reprodutibilidade e repetibilidade completas.

Em alguns sistemas de cultura de células de plantas, a proteína da invenção desejada poderá ser excretada ou, alternativamente, a proteína poderá ser extraída a partir da planta inteira. Quando a proteína da invenção desejada é segregada para o meio de cultura, esta poderá ser recolhida. Alternativamente, os embriões e a porção anembrionada das sementes, ou outros tecidos da planta, poderão ser submetidos a rutura mecânica para permitir a libertação de qualquer proteína segregada presente entre as células e os tecidos. A mistura poderá ser suspensa numa solução-tampão para se obterem as proteínas em forma solúvel. Serão então utilizados métodos convencionais de isolamento e purificação de proteínas para purificar a proteína recombinante. Os parâmetros de tempo, temperatura, pH, oxigénio e volumes serão ajustados, recorrendo a procedimentos de rotina, para otimizar a expressão e a recuperação da proteína heteróloga.

### iii. Sistemas de Baculovírus

O polinucleótido que codifica para a proteína poderá igualmente ser inserido num vetor adequado para expressão em células de inseto, estando ligado de modo operativo aos elementos de controlo que se encontram nesse vetor. A construção de vetores utiliza técnicas conhecidas neste campo. Geralmente, os componentes do sistema de expressão incluem um vetor de transferência, geralmente um plasmídeo bacteriano, que contém um fragmento de genoma do baculovírus e um local de restrição conveniente para inserção no gene ou genes heterólogo(s) a ser(em) expresso(s); um baculovírus de tipo selvagem com uma sequência homóloga ao fragmento específico de

baculovírus no vetor de transferência (tal permite a recombinação homóloga do gene heterólogo com o genoma do baculovírus); e as células hospedeiras de inseto e meio de cultura apropriados.

Após a inserção da sequência de ADN que codifica para a proteína no vetor de transferência, o vetor e o genoma viral de tipo selvagem são transfectados numa célula hospedeira de inseto, na qual se dá a recombinação do vetor e do genoma viral. O vírus recombinante empacotado é expresso e as placas recombinantes são identificadas e purificadas. Os materiais e métodos usados para os sistemas de expressão de baculovírus/células de inseto encontram-se comercialmente disponíveis sob a forma de kit, que, entre outras possibilidades, é fornecido pela empresa Invitrogen, San Diego, CA (kit "MaxBac"). Estas técnicas são geralmente conhecidas dos peritos neste campo e encontram-se descritas em detalhe em Summers e Smith, Texas Agricultural Experiment Station Bulletin No. 1555\_\_(1987) (referidos daqui em diante como "Summers e Smith").

Anteriormente à inserção da sequência de ADN codificando para a proteína no genoma do baculovírus, os componentes acima descritos, compreendendo um promotor, um leader (se desejado), a sequência codificante de interesse e uma sequência de terminação da transcrição, são geralmente reunidos numa construção de transcolocação intermédia (vetor de transferência). Esta construção poderá conter um único gene e elementos reguladores ligados de modo operativo; múltiplos genes, cada um com o seu próprio conjunto de elementos reguladores ligados de modo operativo; ou múltiplos genes, regulados pelo mesmo conjunto de elementos reguladores. As construções de transcolocação intermédia são frequentemente mantidas num replicação, tal como um elemento extracromossómico

(por exemplo, plasmídeos) com a capacidade de se manter de modo estável num hospedeiro, tal como uma bactéria. O replicação possuirá um sistema de replicação, deste modo permitindo a sua manutenção num hospedeiro adequado para clonagem e amplificação.

Atualmente, o vetor de transferência mais vulgarmente usado para a introdução de genes estranhos no AcNPV é o pAc373. Têm igualmente sido concebidos muitos outros vetores já conhecidos dos peritos da técnica. Estes incluem, por exemplo, o pVL985 (que altera o codão de iniciação da polihedrina de ATG para ATT, e que introduz um local de clonagem para BamHI 32 pares de bases a jusante do ATT; ver Luckow e Summers, *Virology* (1989) 17:31.

O plasmídeo geralmente contém também o sinal de poliadenilação da polihedrina (Miller et al. (1988) *Ann. Rev. Microbiol.* 42:177) juntamente com um gene procariótico de resistência à ampicilina (*amp*) e uma origem de replicação procariótica para seleção e propagação em *E. coli*.

Os vetores de transferência de baculovírus contém geralmente um promotor de baculovírus. Um promotor de baculovírus corresponde a qualquer sequência de ADN com a capacidade de se ligar a uma ARN polimerase de baculovírus e de iniciar a transcrição a jusante (5' para 3') de uma sequência codificante (p.ex., gene estrutural) para mARN. Um promotor terá uma região de iniciação da transcrição que é geralmente colocada em posição proximal relativamente à extremidade 5' da sequência codificante. Esta região de iniciação da transcrição inclui geralmente um local de ligação à ARN polimerase e um local de iniciação da transcrição. Um vetor de transferência de baculovírus poderá igualmente conter um segundo domínio designado por potenciador, o qual, se presente, se encontra



geralmente em posição distal relativamente ao gene estrutural. A expressão poderá ser regulada ou constitutiva.

Os genes estruturais, abundantemente transcritos em períodos tardios num ciclo de infeção viral, proporcionam sequências promotoras particularmente úteis. Como exemplos encontram-se as sequências derivadas do gene que codifica para a proteína do poliedro viral (Friesen et al. (1986), "The Regulation of Baculovirus Gene Expression", em: The Molecular Biology of Baculoviruses (ed. Walter Doerfler); Publs. EPO nos. 127 839 e 155 476) e o gene codificando para a proteína p10 (Vlak et al. (1988), J. Gen. Virol. 69:765).

O ADN codificando para sequências de sinal apropriadas poderá derivar de genes para proteínas segregadas por células de inseto ou por baculovírus, tal como o gene de polihedrina do baculovírus (Carbonell et al. (1988) Gene, 73:409). Alternativamente, uma vez que os sinais para as modificações pós-tradução em células de mamífero (tais como o corte do péptido de sinal, o corte proteolítico e a fosforilação) parecem ser reconhecidos pelas células de inseto, e que os sinais requeridos para a secreção e a acumulação no núcleo parecem igualmente ser conservados entre as células de invertebrados e as células de vertebrados, os leaders que não provêm de insetos, tais como os que derivam de genes codificando para o  $\alpha$ -interferão humano (Maeda et al. (1985), Nature 315:592), péptido humano de libertação da gastrina (Lebacqz-Verheyden et al. (1988) Molec. Cell Biol. 8:3129), IL-2 humana (Smith et al. (1985) Proc. Nat'l Acad. Sci USA 82:8404), IL-3 de ratinho (Miyajima et al. (1987) Gene 58:273), e glucocerebrosidase humana (Martin et al. (1988) ADN 7:99), poderão igualmente ser usados para possibilitar a secreção em insetos.

Um polipéptido ou poliproteína recombinante poderá ser expresso intracelularmente ou, se expresso com as sequências reguladoras apropriadas, poderá ser segregado. Uma boa expressão intracelular de proteínas estranhas não fundidas requer geralmente genes heterólogos que idealmente possuam uma curta sequência leader contendo sinais de iniciação da tradução adequados precedendo um sinal de iniciação ATG. Se desejado, a metionina na extremidade N poderá ser cortada da proteína madura através da incubação *in vitro* com brometo de cianogénio.

Alternativamente, as poliproteínas ou proteínas recombinantes que não são naturalmente segregadas poderão ser segregadas pelas células de inseto por meio da criação de moléculas quiméricas de ADN codificando para uma proteína de fusão composta por um fragmento de sequência leader que proporciona a secreção da proteína estranha em insetos. O fragmento de sequência leader codifica geralmente para um péptido de sinal composto por aminoácidos hidrofóbicos que dirige a translocação da proteína para o retículo endoplasmático.

Após a inserção da sequência de ADN e/ou do gene codificando para o produto de expressão precursor da proteína, uma célula hospedeira de inseto é co-transformada com o ADN heterólogo do vetor de transferência e com o ADN genómico do baculovírus de tipo selvagem - geralmente por co-transfecção. O promotor e a sequência de terminação da transcrição desta construção compreenderão geralmente uma secção de 2-5 kb do genoma do baculovírus. Os métodos para a introdução de ADN heterólogo no local desejado do baculovírus são bem conhecidos da técnica (ver Summers e Smith, *supra*; Ju et al. (1987); Smith et al., Mol. Cell Biol. (1983) 3:2156; e Luckow e Summers (1989)). Por exemplo, a inserção poderá efetuar-se num gene tal como o gene da polihedrina, por recombinação homóloga com crossover duplo;

a inserção poderá igualmente dar-se a nível de um local de restrição de uma enzima construído no gene de baculovírus desejado. Miller et al. (1989) Bioessays 4:91. A sequência de ADN, quando clonada em lugar do gene da polihedrina no vetor de expressão, é flanqueada em 5' e em 3' por sequências específicas de polihedrina e está posicionada a jusante do promotor da polihedrina.

O vetor de expressão de baculovírus recém-formado é subsequentemente empacotado num baculovírus recombinante infeccioso. A recombinação homóloga ocorre a uma baixa frequência (entre cerca de 1% e cerca de 5%); assim, a maior parte do vírus produzido após a co-transfecção corresponde ainda ao vírus de tipo selvagem. Deste modo, será necessário recorrer a um método de identificação dos vírus recombinantes. Uma vantagem do sistema de expressão é a de que uma inspeção visual permite a distinção dos vírus recombinantes. A proteína de polihedrina, a qual é produzida pelo vírus nativo, é produzida em níveis muito elevados no núcleo das células infetadas num período tardio após a infeção viral. A proteína de polihedrina acumulada forma corpos de oclusão que contêm também partículas incorporadas. Estes corpos de oclusão, que possuem uma dimensão de até 15  $\mu\text{m}$ , são muito refringentes, o que lhes dá uma aparência bastante brilhante que é facilmente visualizada ao microscópio. As células infetadas com vírus recombinantes não apresentam corpos de oclusão. Para distinguir os vírus recombinantes dos vírus de tipo selvagem, o sobrenadante da transfecção é plaqueado sobre uma monocamada de células de inseto recorrendo a técnicas conhecidas dos peritos neste campo. Nomeadamente, as placas são inspecionadas ao microscópio ótico para avaliar a presença (indicativa de vírus de tipo selvagem) ou ausência (indicativa de vírus recombinante) de corpos de oclusão. "Current Protocols in

Microbiology", Vol.2 (Ausubel et al., eds.) em 16.8 (supl. 10, 1990); Summers e Smith, *supra*; Miller et al. (1989).

Têm sido desenvolvidos vetores de expressão recombinantes de baculovírus para a infecção de diversas células de inseto. Por exemplo, foram desenvolvidos baculovírus recombinantes para as células de, entre outros: *Aedes aegypti*, *Autographa californica*, *Bombyx mori*, *Drosophila melanogaster*, *Spodoptera frugiperda* e *Trichoplusia ni* (Publ. PCT n° WO 89/046699; Carbonell et al. (1985) J. Virol. 56:153; Wright (1986) Nature 321:718; Smith et al. (1983) Mol. Cell Biol. 3:2156; e ver, genericamente, Fraser et al. (1989), In Vitro Cell. Dev. Biol. 25:225).

Encontram-se disponíveis no mercado células e meios de cultura celular destinados a expressão direta, e a expressão em produtos de fusão, de polipéptidos heterólogos num sistema de expressão de baculovírus; a tecnologia aplicada à cultura celular encontra-se difundida entre os peritos na técnica. Ver, p.ex., Summers e Smith, *supra*.

As células de inseto modificadas poderão então ser cultivadas num meio com nutrientes apropriados, o qual permite a manutenção estável do(s) plasmídeo(s) que se encontra(m) presente(s) no hospedeiro de inseto modificado. Quando o gene do produto de expressão se encontra sob controlo indutível, o hospedeiro poderá ser cultivado até uma elevada densidade e a expressão poderá ser induzida. Alternativamente, quando a expressão é constitutiva, o produto será expresso de modo contínuo para o meio e o meio com nutrientes terá de ser continuamente renovado, removendo o produto de interesse e adicionando os nutrientes consumidos. O produto poderá ser purificado por técnicas como a cromatografia, por exemplo, HPLC, cromatografia de afinidade, cromatografia de troca

iônica, etc.; electroforese; centrifugação em gradiente de densidade; extração por solvente, ou semelhantes. Quando apropriado, o produto poderá ser purificado de modo mais completo para remover substancialmente todas as proteínas de inseto que sejam igualmente segregadas para o meio, ou que resultem da lise das células de inseto, de forma a obter um produto que se encontra pelo menos substancialmente isento de resíduos do hospedeiro, p.ex., proteínas, lípidos e polissacáridos.

De modo a se obter a expressão das proteínas, as células recombinantes do hospedeiro que derivam dos transformantes são incubadas sob condições que possibilitam a expressão da sequência que codifica para a proteína recombinante. Estas condições variam, dependendo da célula de hospedeiro selecionada. No entanto, as condições são facilmente discerníveis pelos técnicos de perícia média, tomando como base os conhecimentos difundidos neste campo técnico.

#### iv. Sistemas Bacterianos

As técnicas de expressão bacteriana são já conhecidas da técnica. Um promotor bacteriano corresponde a qualquer sequência de ADN que tem a capacidade de se ligar a uma RNA polimerase bacteriana e iniciar a transcrição a jusante (3') de uma sequência codificante (por exemplo, um gene estrutural) para mRNA. Um promotor possuirá uma região de iniciação da transcrição que geralmente se encontra colocada em posição proximal relativamente à extremidade 5' da sequência codificante. Esta região de iniciação da transcrição inclui geralmente um local de ligação à ARN polimerase e um local de iniciação da transcrição. Um promotor bacteriano poderá igualmente conter um segundo domínio, designado por operador, que se poderá sobrepor a um local adjacente de ligação à ARN polimerase a nível do qual tem início a síntese de ARN. O

operador permite uma transcrição negativamente regulada (indutível), já que uma proteína repressora de um gene se poderá ligar ao operador e deste modo inibir a transcrição deste gene específico. A expressão constitutiva poderá ocorrer na ausência de elementos de regulação negativa, tal como o operador. Adicionalmente, a regulação positiva poderá ser atingida por meio de uma sequência de ligação a uma proteína ativadora do gene, a qual, quando presente, se encontra geralmente em posição proximal (5') relativamente à sequência de ligação à RNA polimerase. Como exemplo de uma proteína ativadora de gene encontra-se a proteína activadora de catabolitos (CAP), a qual ajuda a iniciar a transcrição do operão lac em *Escherichia coli* (*E. coli*) (Raibaud et al (1984) Annu. Rev. Genet. 18:173). A expressão poderá deste modo sofrer uma regulação positiva ou negativa, sendo a transcrição acentuada ou reduzida.

As sequências que codificam para enzimas de vias metabólicas fornecem sequências promotoras particularmente úteis. Os exemplos destas sequências incluem sequências promotoras derivadas de enzimas metabolizadoras de açúcares tais como a galactose, a lactose (*lac*) (Chang et al. (1977) Nature 198:1056) e a maltose. Outros exemplos incluem as sequências promotoras derivadas de enzimas biossintéticas como o triptofano (*trp*) (Goeddel et al. (1980) Nucl. Acids Res. 8:4057; Yelverton et al. (1981) Nucl. Acids Res. 9:731; Patente US 4738921; Publ. EPO Nos. 036 776 e 121 775). O sistema promotor da beta-lactamase (*bla*) (Weissman (1981) "The cloning of interferon and other mistakes", em *Interferon 3* (ed. I. Gresser)) e os sistemas promotores do bacteriófago lambda PL (Shimatake et al. (1981) Nature 292:128) e do T5 (Patente US 4689406) proporcionam igualmente sequências promotoras úteis.

Adicionalmente, certos promotores sintéticos que não ocorrem na natureza funcionam também como promotores em bactérias. Por exemplo, as sequências de ativação da transcrição de um promotor bacteriano ou de bacteriófago poderão ser reunidas com as sequências de operação de outro promotor bacteriano ou de bacteriófago, criando um promotor sintético híbrido (Patente dos EUA N° 4.551.433). Por exemplo, o promotor *tac* corresponde a um promotor híbrido *trp-lac*, composto pelo promotor *trp* e pelas sequências do operação *lac*, que é regulado pelo repressor *lac* (Amann et al. (1983) Gene 25:167; de Boer et al. (1983) Proc. Natl. Acad. Sci. 80:21). Ainda, um promotor bacteriano poderá incluir promotores de ocorrência natural com origem não bacteriana que possuam a capacidade de se ligar à ARN polimerase bacteriana e iniciar a transcrição. Um promotor de ocorrência natural e de origem não bacteriana poderá igualmente ser acoplado a uma ARN polimerase compatível para produzir elevados níveis de expressão de alguns genes em procariotas. O sistema de ARN polimerase de bacteriófago T7/promotor constitui um exemplo de um sistema promotor acoplado (Studier et al. (1986) J. Mol. Biol. 189:113; Tabor et al. (1985) Proc. Natl. Acad. Sci 82:1074). Adicionalmente, um promotor híbrido poderá igualmente ser composto por um promotor de bacteriófago e por uma região operadora de *E. coli* (Publ. EPO N° 267 851).

Para além de uma sequência promotora funcional, a presença de um local eficiente de ligação ao ribossoma é também útil para a expressão de genes estranhos em procariotas. Em *E. coli*, o local de ligação ao ribossoma é designado por sequência de Shine-Dalgarno (SD) e inclui um codão de iniciação (ATG) e uma sequência com 3-9 nucleótidos de extensão localizada 3-11 nucleótidos a montante do codão de iniciação (Shine et al. (1975) Nature 254:34). Pensa-se que a sequência SD promova a ligação do mRNA ao ribossoma através do emparelhamento de

bases entre a sequência SD e a extremidade 3' do rRNA 16S de *E. coli* (Steitz et al. (1979) "Genetic signals and nucleotide sequences in messenger RNA", em *Biological Regulation and Development: Gene Expression* (ed. R.F. Goldberger)). Para conseguir a expressão de genes eucarióticos e genes procarióticos com um fraco local de ligação ao ribossoma, é frequentemente necessário otimizar a distância entre a sequência SD e o codão ATG do gene eucariótico (Sambrook et al. (1989) "Expression of cloned genes in *Escherichia coli*", em "*Molecular Cloning: a Laboratory Manual*").

Uma molécula de ADN pode ser expressa intracelularmente. Poderá existir uma sequência promotora diretamente ligada à molécula de ADN, caso em que o primeiro aminoácido na extremidade N será sempre uma metionina, a qual é codificada pelo codão de iniciação ATG. Se desejado, a metionina na extremidade N poderá ser cortada da proteína por incubação *in vitro* com brometo de cianogénio ou por incubação *in vivo* ou *in vitro* com uma metionina N-terminal peptidase bacteriana (Publ. EPO n° 219 237).

As proteínas de fusão proporcionam uma alternativa à expressão direta. Geralmente, uma sequência de ADN que codifica para a porção N-terminal de uma proteína bacteriana endógena, ou outra proteína estável, é fundida à extremidade 5' de sequências codificantes heterólogas. Após a expressão, a construção proporcionará uma fusão das duas sequências de aminoácidos. Por exemplo, o gene celular do bacteriófago lambda poderá ser ligado à extremidade 5' de um gene estranho e expresso em bactérias. A proteína de fusão resultante retém de preferência um local para uma enzima de processamento (factor Xa), para separar a proteína de bacteriófago da proteína do gene estranho (Nagai et al. (1984) *Nature* 309:810). As proteínas de fusão poderão igualmente ser



produzidas a partir de sequências dos genes *lacZ* (Jia et al. (1987) *Gene* 60:197), *trpE* (Allen et al. (1987) *J. Biotechnol.* 5:93; Makoff et al. (1989) *J. Gen. Microbiol.* 135:11), e *Chey* (Publ. EPO n° 324 647). A sequência de ADN que se encontra na junção das duas sequências de aminoácidos poderá ou não codificar para um local suscetível de corte. Outro exemplo é o de uma proteína de fusão de ubiquitina. Uma tal proteína de fusão é produzida com a região da ubiquitina que retém de preferência um local para uma enzima de processamento (por exemplo, protease de processamento ubiquitina-específica) para separar a ubiquitina da proteína estranha. Usando este método, é possível isolar a proteína estranha nativa (Miller et al. (1989) *Bio/Technology* 7:698).

Alternativamente, as proteínas estranhas poderão igualmente ser segregadas pela célula através da criação de moléculas quiméricas de ADN que codificam para uma proteína de fusão compreendendo um fragmento de uma sequência de péptido de sinal que proporciona a secreção da proteína estranha em bactérias (Patente dos EUA N° 4.336.336). O fragmento de sequência de sinal codifica geralmente para um péptido de sinal composto por aminoácidos hidrofóbicos que dirigem a secreção da proteína a partir da célula. A proteína poderá ser segregada para o meio de crescimento (bactérias gram-positivas) ou para o espaço periplásmico, localizado entre as membranas celulares interna e externa (bactérias gram-negativas). Existirão de preferência locais de processamento, que poderão ser cortados *in vivo* ou *in vitro*, codificados entre o fragmento de péptido de sinal e o gene estranho.

O ADN codificando para sequências de sinal adequadas poderá derivar de genes para proteínas bacterianas segregadas, tal como o gene da proteína da membrana externa de *E. coli* (*ompA*) (Masui et al. (1983), em *Experimental Manipulation of Gene*

*Expression*; Ghayeb et al. (1984) EMBO J. 3:2437) e a sequência de sinal da fosfatase alcalina de *E. coli* (*phoA*) (Oka et al. (1985) Proc. Natl. Acad. Sci. 82:7212). Como exemplo adicional, a sequência de sinal do gene da alfa-amilase proveniente de diversas estirpes de *Bacillus* poderá ser usada para segregar proteínas heterólogas de *B. subtilis* (Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; Publ. EPO n° 244 042).

Geralmente, as sequências de terminação da transcrição que são reconhecidas por bactérias correspondem a regiões reguladoras localizadas em posição 3' relativamente ao codão de paragem da tradução, deste modo flanqueando, juntamente com o promotor, a sequência codificante. Estas sequências dirigem a transcrição de um mRNA que pode ser traduzido para o polipéptido codificado pelo ADN. As sequências de terminação da transcrição incluem frequentemente sequências de ADN com cerca de 50 nucleótidos que têm a capacidade de formar estruturas em gancho que auxiliam na terminação da transcrição. Os exemplos destas sequências incluem sequências de terminação de transcrição derivadas de genes com promotores fortes, tal como o gene *trp* de *E. coli*, assim como outros genes biossintéticos.

Geralmente, os componentes acima descritos, compreendendo um promotor, uma sequência de sinal (se desejado), a sequência codificante de interesse e uma sequência de terminação da transcrição, são reunidos em construções de expressão. As construções de expressão são frequentemente mantidas numa replicação, tal como um elemento extracromossómico (p.ex. plasmídeos), que tem a capacidade de se manter de forma estável num hospedeiro, tal como uma bactéria. O replicão possui um sistema de replicação, o que possibilita a sua manutenção num hospedeiro procariótico para expressão ou para clonagem e amplificação. Adicionalmente, um replicão poderá

corresponder a um plasmídeo de baixo número de cópias ou de elevado número de cópias. Um plasmídeo de elevado número de cópias possuirá geralmente um número de cópias que varia entre cerca de 5 e cerca de 200, e geralmente entre cerca de 10 e cerca de 150. Um hospedeiro contendo um plasmídeo de elevado número de cópias conterà preferencialmente pelo menos cerca de 10, e mais preferencialmente pelo menos cerca de 20 plasmídeos. Poderá selecionar-se um vetor com um elevado ou com um baixo número de cópias, dependendo do efeito que o vetor e a proteína estranha exercem sobre o hospedeiro.

Alternativamente, as construções de expressão poderão ser integradas no genoma bacteriano por meio de um vetor integrativo. Os vetores integrativos contêm geralmente pelo menos uma sequência homóloga ao cromossoma bacteriano que permite a integração do vetor. As integrações parecem resultar de recombinações entre o ADN homólogo do vetor e o cromossoma bacteriano. Por exemplo, os vetores integrativos construídos a partir de ADN proveniente de diversas estirpes de *Bacillus* integram-se no cromossoma de *Bacillus* (Publ. EPO n° 127 328). Os vetores integrativos poderão igualmente ser compostos por sequências de bacteriófago ou de transposição.

Geralmente, as construções de expressão extracromossômicas e integrativas conterão marcadores de seleção que permitam a seleção das estirpes bacterianas que foram transformadas. Os marcadores de seleção podem ser expressos pelo hospedeiro bacteriano e podem incluir genes que tornam as bactérias resistentes a fármacos como a ampicilina, cloranfenicol, eritromicina, canamicina (neomicina) e tetraciclina (Davies et al. (1978) Annu. Rev. Microbiol. 32:469). Os marcadores de seleção poderão igualmente incluir genes biossintéticos, tal como os genes das vias metabólicas da histidina, triptofano e leucina.

Alternativamente, alguns dos componentes acima descritos poderão ser reunidos em vetores de transformação. Os vetores de transformação incluem geralmente um marcador de seleção que poderá ser mantido numa replicação ou desenvolvido para um vetor integrativo, tal como acima descrito.

Têm sido desenvolvidos diversos vetores de expressão e de transformação, tanto sob a forma de replicões extracromossômicos como de vetores integrativos, para a transformação de um grande número de bactérias. Por exemplo, foram desenvolvidos vetores de expressão para as seguintes bactérias, entre outras: *Bacillus subtilis* (Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; Publ. EPO Nos. 036 259 e 063 953; Publ. PCT n° WO 84/04541), *Escherichia coli* (Shimatake et al. (1981) Nature 292:128; Amann et al. (1985) Gene 40:183; Studier et al. (1986) J. Mol. Biol. 189:113; Publ. EPO Nos. 036 776, 136 829 e 136 907), *Streptococcus cremoris* (Powell et al. (1988) Appl. Environ. Microbiol. 54:655); *Streptococcus lividans* (Powell et al. (1988) Appl. Environ. Microbiol. 54:655), *Streptomyces lividans* (Patente US 4745056).

Os métodos para a introdução de ADN exógeno em hospedeiros bacterianos são bem conhecidos neste campo técnico e incluem geralmente a transformação de bactérias tratadas com  $\text{CaCl}_2$  ou com outros agentes, tal como catiões divalentes e DMSO. O ADN poderá também ser introduzido nas células bacterianas por electroporação. Os procedimentos de transformação variam geralmente com a espécie bacteriana a transformar. Ver, por exemplo, o uso de *Bacillus*: Masson et al. (1989) FEMS Microbiol. Lett. 60:273; Palva et al. (1982) Proc. Natl. Acad. Sci. USA 79:5582; Publ. EPO Nos. 036 259 e 063 953; Publ. PCT n° WO 84/04541; o uso de *Campylobacter*: Miller et al. (1988) Proc. Natl. Acad. Sci. 85:856; Wang et al. (1990) J.

Bacteriol. 172:949; o uso de *Escherichia coli*: Cohen et al. (1973) Proc. Natl. Acad. Sci. 69:2110; Dower et al. (1988) Nucleic Acids Res. 16:6127; Kushner (1978) "An improved method for transformation of *Escherichia coli* with ColE1-derived plasmids. Em Genetic Engineering: Proceedings of the International Symposium on Genetic Engineering (eds. H.W. Boyer e S. Nicosia); Mandel et al. (1970) J. Mol. Biol. 53:159; Taketo (1988) Biochim. Biophys. Acta 949:318; o uso de *Lactobacillus*: Chassy et al. (1987) FEMS Microbiol. Lett. 44:173; o uso de *Pseudomonas*: Fiedler et al. (1988) Anal. Biochem. 170:38; o uso de *Staphylococcus*: Augustin et al. (1990) FEMS Microbiol. Lett. 66:203; o uso de *Streptococcus*: Barany et al. (1980) J. Bacteriol. 144:698; Harlander (1987) "Transformation of *Streptococcus lactis* by electroporation", em: Streptococcal Genetics (eds. J. Ferretti e R. Curtiss III); Perry et al. (1981) Infec. Immun. 32:1295; Powell et al. (1988) Appl. Environ. Microbiol. 54:655; Somkuti et al. (1987) Proc. 4<sup>th</sup> Evr. Cong. Biotechnology 1:412.

#### v. Expressão em Leveduras

Os sistemas de expressão em leveduras são igualmente familiares aos técnicos de perícia média neste campo. Um promotor de levedura corresponde a qualquer sequência de ADN que possua a capacidade de se ligar à ARN polimerase de levedura e de iniciar a transcrição a jusante (3') de uma sequência codificante (por exemplo, um gene estrutural) para mARN. Um promotor conterá uma região de iniciação da transcrição que é geralmente colocada em posição proximal relativamente à extremidade 5' da sequência codificante. Esta região de iniciação da transcrição inclui geralmente um local de ligação à ARN polimerase (a "TATA box") e um local de iniciação da transcrição. O promotor de levedura poderá igualmente conter um segundo domínio, designado por sequência ativadora a montante (UAS), que, se presente, se encontra

geralmente em posição distal relativamente ao gene estrutural. A UAS permite uma expressão regulada (indutível). A expressão constitutiva ocorre na ausência de uma UAS. A expressão regulada poderá ser positiva ou negativa, ou seja, a transcrição poderá ser aumentada ou reduzida.

Uma levedura é um organismo fermentativo com uma via metabólica ativa, pelo que as sequências que codifiquem para enzimas da via metabólica constituirão sequências promotoras particularmente úteis. Como exemplos destas enzimas encontram-se a álcool desidrogenase (ADH) (Publ. EPO nº 284 044), enolase, glucoquinase, glucose-6-fosfato isomerase, gliceraldeído-3-fosfato-desidrogenase (GAP ou GAPDH), hexoquinase, fosfofrutoquinase, 3-fosfoglicerato mutase e piruvatoquinase (PyK) (Publ. EPO nº 329 203). O gene *PHO5* de levedura, codificando para a fosfatase ácida, proporciona igualmente sequências promotoras úteis (Myanohara et al. (1983) Proc. Natl. Acad. Sci. USA 80:1).

Adicionalmente, certos promotores sintéticos que não ocorrem na natureza poderão também funcionar como promotores de levedura. Por exemplo, as sequências UAS de um promotor de levedura poderão ser reunidas com a região de ativação da transcrição de outro promotor de levedura, criando-se um promotor híbrido sintético. Os exemplos de tais promotores híbridos incluem a sequência reguladora da ADH ligada à região de ativação da transcrição da GAP (Patentes dos EUA Nos. 4.876.197 e 4.880.734). Outros exemplos de promotores híbridos incluem promotores que correspondem às sequências reguladoras dos genes *ADH2*, *GAL4*, *GAL10* ou *PHO5*, combinadas com a região de ativação da transcrição de um gene de uma enzima glicolítica como a GAP ou a PyK (Publ. EPO nº 164 556). Adicionalmente, um promotor de levedura poderá incluir promotores de ocorrência natural não derivados de leveduras

que possam a capacidade de se ligar à ARN polimerase de levedura e iniciar a transcrição. Os exemplos de tais promotores incluem, entre outros, (Cohen et al. (1980) Proc. Natl. Acad. Sci. USA 77:1078; Henikoff et al. (1981) Nature 283:835; Hollenberg et al. (1981) Current Topics Microbiol. Immunol. 96:119; Hollenberg et al. (1979) "The Expression of Bacterial Antibiotic Resistance Genes in the Yeast *Saccharomyces cerevisiae*", em: Plasmids of Medical, Environmental and Commercial Importance (eds. K.N. Timmis e A. Puhler); Mercerau-Puigalon et al. (1980) Gene 11:163; Panthier et al. (1980) Curr. Genet. 2:109).

Uma molécula de ADN poderá ser expressa intracelularmente em leveduras. Poderá existir uma sequência promotora diretamente ligada à molécula de ADN, caso em que o primeiro aminoácido localizado na extremidade N-terminal da proteína recombinante será sempre uma metionina, a qual é codificada pelo codão de iniciação ATG. Se desejado, a metionina na extremidade N poderá ser cortada da proteína por incubação *in vitro* com brometo de cianogénio.

As proteínas de fusão proporcionam uma alternativa aos sistemas de expressão em leveduras, assim como aos sistemas de expressão em mamíferos, plantas, baculovírus e bactérias. Geralmente, uma sequência de ADN que codifica para a porção N-terminal de uma proteína endógena de levedura, ou outra proteína estável, é fundida à extremidade 5' de sequências codificantes heterólogas. Após a expressão, esta construção proporcionará uma fusão das duas sequências de aminoácidos. Por exemplo, o gene da superóxido dismutase (SOD) humana ou de levedura poderá ser ligado à extremidade 5' de um gene estranho e expresso em leveduras. A sequência de ADN que se encontra na junção das duas sequências de aminoácidos poderá ou não conter um local suscetível de corte. Ver, por exemplo,

a Publ. EPO n° 196056. Outro exemplo é o de uma proteína de fusão de ubiquitina. Uma tal proteína de fusão é produzida com a região da ubiquitina que retém de preferência um local para uma enzima de processamento (por exemplo protease de processamento ubiquitina-específica) para separar a ubiquitina da proteína estranha. Usando este método, é possível isolar a proteína estranha nativa (ver, por exemplo, WO 88/024066).

Alternativamente, as proteínas estranhas poderão igualmente ser segregadas pela célula para o meio de cultura, recorrendo à criação de moléculas quiméricas de ADN que codifiquem para uma proteína de fusão composta por um fragmento de sequência leader que proporcione a secreção da proteína estranha pela levedura. De preferência, existirão locais de processamento codificados entre o fragmento leader e o gene estranho que poderão ser submetidos a corte *in vivo* ou *in vitro*. O fragmento de sequência leader codifica geralmente para um péptido de sinal composto por aminoácidos hidrofóbicos que dirigem a secreção da proteína pela célula.

O ADN codificando para sequências de sinal adequadas poderá derivar de genes de proteínas segregadas por leveduras, tal como o gene da invertase de levedura (Publ. EPO n° 012 873; Publ. JPO n° 62:096,086) e o gene do fator A (Patente dos EUA N° 4.588.684). Alternativamente, existem leaders não derivados de leveduras, tal como um leader de interferão, que proporcionam igualmente a secreção por leveduras (Publ. EPO n° 060 057).

Uma classe de leaders de secreção preferida é aquela dos que utilizam um fragmento do gene do fator alfa de levedura, o qual contém tanto uma sequência de sinal "pré" como uma região "pró". Os tipos de fragmentos de fator alfa que podem ser utilizados incluem o leader de fator alfa pré-pró de extensão



completa (com cerca de 83 resíduos de aminoácidos) assim como leaders de fator alfa truncados (geralmente com cerca de 25 a cerca de 50 resíduos de aminoácidos) (Patentes dos EUA Nos. 4.546.083 e 4.870.008; Publ. EPO nº 324 274). Outros leaders que utilizam um fragmento do leader de fator alfa para proporcionar a secreção incluem leaders híbridos de fator alfa construídos com uma pré-sequência de uma primeira levedura, mas uma pró-região proveniente de um segundo fator alfa de levedura (ver, por exemplo, Publ. PCT Nº WO 89/02463).

Geralmente, as sequências de terminação da transcrição que são reconhecidas por leveduras correspondem a regiões reguladoras localizadas em posição 3' relativamente ao codão de paragem da tradução, deste modo flanqueando, juntamente com o promotor, a sequência codificante. Estas sequências dirigem a transcrição de um mRNA que poderá ser traduzido para o polipéptido que é codificado pelo ADN. Existem diversos exemplos de sequências terminadoras da transcrição e de outras sequências de terminação reconhecidas por leveduras, tais como as que codificam para enzimas glicolíticas.

Geralmente, os componentes acima descritos, compreendendo um promotor, um leader (se desejado), a sequência codificante de interesse e uma sequência de terminação da transcrição, são reunidos em construções de expressão. As construções de expressão são frequentemente mantidas numa replicação, tal como um elemento extracromossómico (por exemplo, plasmídeos) com a capacidade de se manter de modo estável num hospedeiro, tal como uma levedura ou bactéria. O replicação poderá possuir dois sistemas de replicação, deste modo permitindo a sua manutenção, por exemplo, numa levedura para a expressão e num hospedeiro procariótico para clonagem e amplificação. Os exemplos de tais vetores de vaivém de levedura-bactéria incluem o YEpl24 (Botstein et al. (1979) Gene 8:17-24), pCl/1

(Brake et al. (1984) Proc. Natl. Acad. Sci USA 81:4642-4646) e YRp17 (Stinchcomb et al. (1982) J. Mol. Biol. 158:157). Adicionalmente, uma replicação poderá corresponder a um plasmídeo de elevado número de cópias ou a um plasmídeo de baixo número de cópias. Um plasmídeo de elevado número de cópias possuirá geralmente um número de cópias variando entre cerca de 5 e cerca de 200, e geralmente entre cerca de 10 e cerca de 150. Um hospedeiro contendo um plasmídeo de elevado número de cópias conterá preferencialmente pelo menos cerca de 10, e mais preferencialmente pelo menos cerca de 20 plasmídeos. Poderá selecionar-se um vetor com um elevado ou com um baixo número de cópias, dependendo do efeito que o vetor e a proteína estranha exercem sobre o hospedeiro. Ver, p.ex., Brake et al., *supra*.

Alternativamente, as construções de expressão poderão ser integradas no genoma de levedura através de um vetor integrativo. Os vetores integrativos contêm geralmente pelo menos uma sequência homóloga a um cromossoma de levedura que permite a integração do vetor, e de preferência contêm duas sequências homólogas flanqueando a construção de expressão. As integrações parecem resultar de recombinações entre o ADN homólogo no vetor e o cromossoma da levedura (Orr-Weaver et al. (1983) Methods in Enzymol. 101:228-245). Um vetor integrativo poderá ser dirigido para um locus específico na levedura através da seleção da sequência homóloga apropriada para inclusão no vetor. Ver Orr-Weaver et al., *supra*. Um ou mais construções de expressão poderão sofrer integração, possivelmente afetando os níveis de proteína recombinante produzidos (Rine et al. (1983) Proc. Natl. Acad. Sci. USA 80:6750)). As sequências cromossômicas incluídas no vetor poderão ocorrer sob a forma de um único segmento no vetor, o que resulta na integração do vetor completo, ou de dois segmentos homólogos a segmentos adjacentes no cromossoma e

flanqueando a construção de expressão no vetor, o que poderá resultar na integração estável de apenas a construção de expressão.

Usualmente, as construções de expressão extracromossômicas e integrativas poderão conter marcadores de seleção que permitam a seleção das estirpes de leveduras que foram transformadas. Os marcadores de seleção poderão incluir genes biossintéticos que podem ser expressos no hospedeiro de levedura, tal como os genes *ADE2*, *HIS4*, *LEU2*, *TRP1* e *ALG7*, e o gene de resistência a G418, que conferem resistência em células de leveduras à tunicamicina e a G418, respectivamente. Adicionalmente, um marcador de seleção apropriado poderá igualmente conferir à levedura a capacidade de se desenvolver na presença de compostos tóxicos, como metais. Por exemplo, a presença de *CUP1* permite à levedura crescer na presença de iões de cobre (Butt et al. (1987) Microbiol. Rev. 51:351).

Alternativamente, alguns dos componentes acima descritos poderão ser reunidos em vetores de transformação. Os vetores de transformação incluem geralmente um marcador de seleção que poderá ser mantido numa replicação ou desenvolvido para um vetor integrativo, tal como acima descrito.

Têm sido desenvolvidos diversos vetores de expressão e de transformação, tanto sob a forma de replicões extracromossômicos como de vetores integrativos, para a transformação de um grande número de leveduras. Por exemplo, foram desenvolvidos vetores de expressão e métodos para introdução de ADN exógeno em leveduras para as seguintes leveduras, entre outras: *Candida albicans* (Kurtz, et al. (1986) Mol. Cell. Biol. 6:142); *Candida maltosa* (Kunze, et al. (1985) J. Basic Microbiol. 25:141); *Hansenula polymorpha* (Gleeson, et al. (1986) J. Gen. Microbiol. 132:3459;

Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302); *Kluyveromyces fragilis* (Das, et al. (1984) J. Bacteriol. 158:1165); *Kluyveromyces lactis* (De Louvencourt et al. (1983) J. Bacteriol. 154:737; Van den Berg et al. (1990) Bio/Technology 8:135), *Pichia guillierimondii* (Kunze et al. (1985) J. Basic Microbiol. 25:141), *Pichia pastoris* (Cregg et al. (1985) Mol. Cell. Biol. 5:3376; Patentes dos EUA Nos. 4.837.148 e 4.929.555), *Saccharomyces cerevisiae* (Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163), *Schizosaccharomyces pombe* (Beach e Nurse (1981) Nature 300:706) e *Yarrowia lipolytica* (Davidow et al. (1985) Curr. Genet. 10:380471; Gaillardin et al. (1985) Curr. Genet. 10:49).

Os métodos para introdução de ADN exógeno em hospedeiros de levedura são bem conhecidos neste campo técnico, e incluem geralmente a transformação de esferoplastos, ou de células de levedura intactas, com tratamento por catiões alcalinos. Os procedimentos de transformação variam geralmente segundo a espécie de levedura a ser transformada. Ver, p.ex., [Kurtz et al. (1986) Mol. Cell. Biol. 6:142; Kunze et al. (1985) J. Basic Microbiol. 25:141; Candida]; [Gleeson et al. (1986) J. Gen. Microbiol. 132:3459; Roggenkamp et al. (1986) Mol. Gen. Genet. 202:302; Hansenula]; [Das et al. (1984) J. Bacteriol. 158:1165; De Louvencourt et al. (1983) J. Bacteriol. 154:1165; Van den Berg et al. (1990) Bio/Technology 8:135; Kluyveromyces]; [Cregg et al. (1985) Mol. Cell Biol. 5:3376; Kunze et al. (1985) J. Basic Microbiol. 25:141; Patentes dos EUA Nos. 4.837.148 e 4.929.555; Pichia]; [Hinnen et al. (1978) Proc. Natl. Acad. Sci. USA 75:1929; Ito et al. (1983) J. Bacteriol. 153:163; Saccharomyces]; [Beach e Nurse (1981) Nature 300:706; Schizosaccharomyces]; [Davidow et al. (1985) Curr. Genet. 10:39; Gaillardin et al. (1985) Curr. Genet. 10:49; Yarrowia].

### Definições

Uma composição contendo X encontra-se "substancialmente isenta de" Y quando pelo menos 85% em peso do total de X + Y na composição corresponde a X. De preferência, X compreende pelo menos cerca de 90% em peso do total de X + Y na composição, mais preferencialmente pelo menos cerca de 95% ou até 99% em peso.

Um fragmento de aminoácido ou proteína de *Neisseria* que é "conservado" estará presente numa proteína particular de *Neisseria* em pelo menos x% de *Neisseria*. O valor de x poderá corresponder a 50% ou mais, p.ex., 66%, 75%, 80%, 90%, 95%, ou mesmo 100% (i.e., o aminoácido é encontrado na proteína em questão em todas as *Neisseria*). De maneira a determinar se um aminoácido é "conservado" numa determinada proteína de *Neisseria*, é necessário comparar esse resíduo de aminoácido nas sequências da proteína em questão provenientes de uma grande variedade de diferentes *Neisseria* (uma população de referência). Essa população de referência pode incluir vários serogrupos diferentes de uma espécie em particular ou um único serogrupo. Uma população de referência preferida é constituída pelas 5 estirpes de *Neisseria* mais comuns.

O termo "heterólogo" refere-se a dois componentes biológicos que não são encontrados juntos na natureza. Os componentes poderão corresponder a células hospedeiras, genes ou regiões reguladoras, tais como promotores. Apesar de os componentes heterólogos não se encontrarem juntos na natureza, estes poderão funcionar em conjunto, tal como acontece quando um promotor heterólogo relativamente a um gene se encontra ligado de modo operativo a esse gene. Outro exemplo é dado por uma sequência de *Neisseria* que é heteróloga relativamente a uma célula hospedeira de ratinho.

"Epitopo" significa determinante antigénico, e pode provocar uma resposta celular e/ou humoral.

As condições de "elevada restringência" correspondem a 65 graus Celsius e uma solução de 0,1x SSC / 0,5% SDS.

Uma "origem de replicação" corresponde a uma sequência polinucleotídica que inicia e regula a replicação de polinucleótidos, tal como um vetor de expressão. A origem de replicação comporta-se como uma unidade autónoma de replicação polinucleotídica no interior de uma célula, tendo a capacidade de se replicar sob o seu próprio controlo. Poderá ser necessária a existência de uma origem de replicação para que um vetor se replique numa determinada célula hospedeira. Com certas origens de replicação, um vetor de expressão poderá reproduzir-se até um elevado número de cópias se as proteínas apropriadas se encontrarem presentes no interior da célula. Como exemplos de origens de replicação encontram-se as sequências de replicação autónoma, as quais são eficazes em leveduras; e o antígeno T viral, o qual é eficaz nas células COS-7.

Uma sequência "mutante" é definida como uma sequência de ADN, ARN ou aminoácidos que seja distinta de, mas possua homologia com, a sequência nativa ou apresentada. Dependendo da sequência particular considerada, o grau de homologia (identidade de sequência) entre a sequência nativa ou apresentada e a sequência mutante será preferencialmente superior a 50% (por exemplo, 60%, 70%, 80%, 90%, 95%, 99% ou mais), o que pode ser calculado tal como acima descrito. Tal como aqui é usado, o termo "variante alélica" de uma molécula, ou região, de ácido nucleico para a qual é aqui apresentada a sequência de ácido nucleico corresponde a uma molécula ou região de ácido nucleico que ocorre essencialmente no mesmo

locus do genoma de um outro ou segundo isolado, e que, devido à variação natural causada por, por exemplo, mutação ou recombinação, apresenta uma sequência de ácido nucleico que é semelhante mas não idêntica. Uma variante alélica de uma região codificante codifica tipicamente para uma proteína que possui uma atividade semelhante à da proteína codificada pelo gene com o qual se compara a variante. Uma variante alélica poderá igualmente compreender uma alteração nas regiões 5' ou 3' não traduzidas do gene, tais como as regiões reguladoras de controlo (ver, por exemplo, a Patente US 5753235).

### Anticorpos

Tal como aqui é usado, o termo "anticorpo" refere-se a um polipéptido ou grupo de polipéptidos composto por pelo menos um local de combinação do anticorpo. Um "local de combinação do anticorpo" corresponde ao espaço tridimensional de ligação possuindo uma conformação de superfície interna e uma distribuição de cargas que é complementar às características de um epítipo de um antígeno, o que permite a ligação entre o anticorpo e o antígeno. O termo "anticorpo" inclui, por exemplo, anticorpos de vertebrados, anticorpos híbridos, anticorpos quiméricos, anticorpos humanizados, anticorpos alterados, anticorpos univalentes, proteínas Fab e anticorpos de domínio único.

Os anticorpos dirigidos contra as proteínas da invenção encontram utilidade em procedimentos de cromatografia por afinidade, em imunoensaios e na destrição/identificação de proteínas de *Neisseria* menB. Os anticorpos dirigidos contra as proteínas do presente invenção ligam-se a polipéptidos antigénicos ou a proteínas, ou fragmentos de proteína, que estão presentes em, e especificamente associados a, estirpes menB de *Neisseria meningitidis*. Em alguns casos, estes antígenos poderão estar associados a estirpes específicas,

tal como os antigénios específicos das estirpes de menB. Os anticorpos da invenção podem ser imobilizados numa matriz e utilizados em imunoensaios ou em procedimentos de cromatografia em coluna por afinidade, de maneira a permitir a deteção e/ou separação de polipéptidos, proteínas ou fragmentos de proteína ou de células que incluem esses polipéptidos, proteínas ou fragmentos de proteína. Alternativamente, esses polipéptidos, proteínas ou fragmentos de proteína poderão ser imobilizados para que possam ser detetados anticorpos especificamente capazes de se ligarem a esses polipéptidos, proteínas ou fragmentos de proteína.

Os anticorpos dirigidos contra as proteínas da invenção, tanto policlonais como monoclonais, poderão ser preparados por recurso a métodos convencionais. Em geral, a proteína é inicialmente usada para imunizar um animal apropriado, de preferência um ratinho, um rato, um coelho ou uma cabra. Os coelhos e as cabras são preferidos para a preparação de soros policlonais devido ao volume de soro que é possível obter e à facilidade de obtenção de anticorpos marcados anti-coelho e anti-cabra. A imunização é geralmente realizada através da mistura ou emulsificação da proteína em soro fisiológico, preferencialmente num adjuvante como o adjuvante completo de Freund, e da injeção da mistura ou emulsão por via parentérica (em geral subcutânea ou intramuscular). Uma dose de 50-200 µg/injeção será tipicamente suficiente. Em geral, a imunização é reforçada 2-6 semanas mais tarde com uma ou mais injeções da proteína em soro fisiológico, preferencialmente usando o adjuvante incompleto de Freund. Poderão alternativamente gerar-se anticorpos por imunização *in vitro* usando métodos bem conhecidos da técnica, procedimento este que, para os fins do presente invenção, é considerado como equivalente à imunização *in vivo*. O antisoro policlonal é obtido sangrando o animal imunizado para um contentor de vidro ou de plástico e



incubando o sangue a 25°C durante uma hora, seguindo-se uma incubação a 4°C durante 2-18 horas. O soro é recuperado por centrifugação (por exemplo, a 1000g durante 10 minutos). Poderão obter-se cerca de 20-50 ml de soro por sangradura em coelhos.

Os anticorpos monoclonais são preparados utilizando o método padronizado de Kohler e Milstein (Nature (1975) 256:495-96), ou uma sua modificação. Tipicamente, um ratinho ou um rato é imunizado tal como acima descrito. No entanto, em lugar de sangrar o animal para extrair o soro, é removido o baço (e opcionalmente diversos nódulos linfáticos grandes), procedendo-se então à sua dissociação em células isoladas. Se desejado, os esplenócitos poderão ser selecionados (após a remoção das células aderentes não específicas) através da aplicação de uma suspensão celular a uma placa ou poço revestidos com o antígeno proteico. As células B que expressem uma imunoglobulina ligada à membrana específica para o antígeno ligam-se à placa e não são removidas por lavagem juntamente com o resto da suspensão. As células B resultantes, ou todos os esplenócitos dissociados, são então induzidos a fundir-se com células de mieloma para formar hibridomas, os quais são cultivados em meio seletivo (por exemplo, meio de hipoxantina, aminopterina e timidina, "HAT"). Os hibridomas resultantes são plaqueados por diluição limitante e são ensaiados quanto à produção de anticorpos que se ligam especificamente ao antígeno imunizante (e que não se ligam a antígenos não relacionados). Os hibridomas secretores de mAc selecionados são então cultivados *in vitro* (por exemplo, em frascos de cultura de tecidos ou em reatores de fibras ocas) ou *in vivo* (como ascites em ratinhos).

Se desejado, os anticorpos (policlonais ou monoclonais) poderão ser marcados recorrendo a técnicas convencionais. Os

marcadores adequados incluem fluoróforos, cromóforos, átomos radioativos (particularmente  $^{32}\text{P}$  e  $^{125}\text{I}$ ), reagentes densos a eletrões, enzimas e ligandos possuindo parceiros de ligação específica. As enzimas são tipicamente detetadas por meio da sua atividade. Por exemplo, a peroxidase de rábano é geralmente detetada em virtude da sua capacidade para converter o substrato 3,3',5,5'-tetrametilbenzidina (TMB) a um pigmento azul, quantificável através de um espectrofotómetro. O termo "parceiro de ligação específica" refere-se a uma proteína que possui a capacidade de se ligar a uma molécula de ligando com uma elevada especificidade, como acontece, por exemplo, no caso de um antigénio e de um anticorpo monoclonal que lhe é específico. Outros parceiros de ligação específica incluem a biotina e a avidina ou a estreptoavidina, IgG e proteína A, e os numerosos pares de ligando-receptor que se conhecem neste campo técnico. Deverá ter-se em conta que a descrição acima não pretende categorizar os vários marcadores em classes distintas, uma vez que o mesmo marcador poderá funcionar por diversos modos diferentes. Por exemplo, o  $^{125}\text{I}$  poderá funcionar como um marcador radioativo ou como um reagente denso a eletrões. A HRP poderá servir como uma enzima ou como um antigénio para um mAc. Adicionalmente, é possível combinar diversos marcadores para obter o efeito desejado. Por exemplo, os mAc e a avidina requerem igualmente marcadores quando usados no âmbito desta invenção; deste modo, poder-se-á marcar um mAc com biotina e detectar a sua presença através de avidina marcada com  $^{125}\text{I}$ , ou com um mAc anti-biotina marcado com HRP. Outras permutas e possibilidades se tornarão facilmente aparentes aos técnicos de perícia média neste campo, sendo todas estas possibilidades consideradas como equivalentes no âmbito da presente invenção.

Os antigénios, imunogénios, polipéptidos, proteínas ou fragmentos de proteína do presente invenção estimulam a

formação de anticorpos com parceiros específicos de ligação. Estes antigénios, imunogénios, polipéptidos, proteínas ou fragmentos de proteína do presente invenção compreendem composições imunogénicas do presente invenção. Essas composições imunogénicas podem ainda compreender ou incluir adjuvantes, veículos ou outras composições que promovam, intensifiquem ou estabilizem os antigénios, polipéptidos, proteínas ou fragmentos de proteína do presente invenção. Esses adjuvantes e veículos serão facilmente aparentes aos técnicos de perícia média neste campo.

#### Composições farmacêuticas

As composições farmacêuticas poderão compreender (incluir) polipéptidos, anticorpos ou o ácido nucleico da invenção. As composições farmacêuticas compreenderão uma quantidade terapêuticamente eficaz dos polipéptidos, anticorpos ou polinucleótidos da invenção reivindicado.

O termo "quantidade terapêuticamente eficaz", tal como aqui é usado, refere-se a uma quantidade de agente terapêutico que trata, melhora ou previne uma doença ou estado para os quais se deseja o efeito terapêutico, ou que exibe um efeito terapêutico ou profilático detetável. O efeito poderá ser detetado através de, por exemplo, marcadores bioquímicos ou concentrações de antigénio. Os efeitos terapêuticos incluem igualmente a redução dos sintomas físicos, tal como a redução da temperatura corporal, no caso de pacientes febris. A quantidade eficaz precisa a utilizar num indivíduo dependerá do tamanho do indivíduo e do seu estado de saúde, da natureza e gravidade do estado e da terapêutica ou combinação de terapêuticas que foram selecionadas para administração. Assim, não será útil especificar antecipadamente uma quantidade eficaz exata. No entanto, a quantidade eficaz para uma dada

situação poderá ser determinada através de experimentação de rotina e será decidida pelo clínico.

Para os propósitos do presente invenção, uma dose eficaz situar-se-á entre cerca de 0,01 mg/kg e 50 mg/kg, ou de 0,05 mg/kg a cerca de 10 mg/kg das construções de ADN para o indivíduo ao qual são administradas.

Uma composição farmacêutica poderá igualmente incluir um veículo farmaceuticamente aceitável. O termo "veículo farmaceuticamente aceitável" refere-se a um veículo para administração de um agente terapêutico, tal como anticorpos ou um polipéptido, genes e outros agentes terapêuticos. O termo abrange qualquer veículo farmacêutico que não induza por si mesmo a produção de anticorpos prejudiciais para o indivíduo que recebe a composição, e que possa ser administrado sem daí resultar uma toxicidade indevida. Os veículos adequados poderão corresponder a macromoléculas grandes de metabolização lenta, tais como proteínas, polissacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos e partículas virais inativas. Tais veículos são bem conhecidos dos técnicos de perícia média neste campo.

Poderão utilizar-se sais farmaceuticamente aceitáveis, por exemplo, sais de ácidos minerais como cloridratos, bromidratos, fosfatos, sulfatos e semelhantes; e os sais de ácidos orgânicos como acetatos, propionatos, malonatos, benzoatos e semelhantes. Encontra-se disponível uma exposição detalhada dos excipientes farmaceuticamente aceitáveis em Remington's Pharmaceutical Sciences (Mack Publ. Co., N.J., 1991).

Os veículos farmacologicamente aceitáveis que se encontram presentes em composições terapêuticas poderão conter líquidos como a água, soro fisiológico, glicerol e etanol. Adicionalmente, poderão encontrar-se presentes em tais veículos substâncias auxiliares, tais como agentes molhantes ou emulsificantes, substâncias tamponadoras de pH e semelhantes. Tipicamente, as composições terapêuticas são preparadas sob a forma de injetáveis, em soluções ou suspensões líquidas; poderão igualmente preparar-se formas sólidas adequadas para solução, ou suspensão, em veículos líquidos antes da injeção. Os lipossomas são incluídos no âmbito da definição de um veículo farmacologicamente aceitável.

#### Métodos de Administração

Uma vez formuladas, as composições da presente invenção poderão ser administradas diretamente ao indivíduo. Os indivíduos a tratar poderão ser animais e, em particular, seres humanos.

A administração direta das composições será geralmente conseguida através de injeção, subcutânea, intraperitoneal, intravenosa ou intramuscular, ou de libertação para o espaço intersticial de um tecido. As composições poderão ainda ser administradas a uma lesão. Outros modos de administração incluem a administração oral e pulmonar, supositórios e aplicações transdérmicas ou transcutâneas, agulhas e pistolas de gases ou hiposprays. A posologia poderá corresponder a um regime de dose única ou a um regime de múltiplas doses.

#### Vacinas

As vacinas produzidas de acordo com a invenção poderão ter fins profiláticos (i.e., de prevenção da infeção) ou terapêuticos (i.e., de tratamento da doença após a infeção).

Estas vacinas compreenderão antigénio(s) imunizante(s), imunogénio(s), polipéptido(s), proteína(s) ou fragmentos de proteína imunogénicos, ou ácidos nucleicos (por exemplo, ácido ribonucleico ou ácido desoxiribonucleico), geralmente em combinação com "veículos farmacêuticamente aceitáveis", os quais incluem qualquer veículo que não induza por si mesmo a produção de anticorpos prejudiciais para o indivíduo que recebe a composição. Os veículos adequados correspondem tipicamente a macromoléculas grandes de metabolização lenta, tais como proteínas, polissacáridos, ácidos polilácticos, ácidos poliglicólicos, aminoácidos poliméricos, copolímeros de aminoácidos, agregados lipídicos (tais como gotículas de óleo e lipossomas) e partículas virais inativas. Tais veículos são bem conhecidos dos técnicos de perícia média neste campo. Adicionalmente, estes veículos poderão funcionar como agentes imunoestimuladores ("adjuvantes"). Ainda, o imunogénio ou antigénio poderá estar conjugado com um toxoide bacteriano, tal como um toxoide dos agentes patogénicos da difteria, tétano ou cólera, de *H. pylori*, etc.

Os adjuvantes preferidos para acentuar a eficácia da composição incluem, de modo não limitativo: (1) sais de alumínio (alum) tais como hidróxido de alumínio, fosfato de alumínio, sulfato de alumínio, etc.; (2) formulações de emulsão óleo-em-água (com ou sem outros agentes imunoestimuladores específicos tais como muramil-péptidos (ver abaixo) ou componentes da parede celular bacteriana), tais como, por exemplo, (a) MF59 (Publ. PCT n° WO 90/14837), contendo 5% de esqualeno, 0,5% de Tween 80 e 0,5% de Span 85 (contendo opcionalmente quantidades diversas de MTP-PE (ver abaixo), se bem que tal não seja necessário), formulado em partículas submicrónicas por meio de um microfluidificador, tal como o microfluidificador Modelo 110Y (Microfluidics, Newton, MA), (b) SAF, contendo 10% de esqualeno, 0,4% de Tween

80, 5% de polímero L121 bloqueado por plurónico, e thr-MDP (ver abaixo), microfluidizado até uma emulsão submicrônica ou vortexado para gerar uma emulsão com partículas de maior tamanho, e (c) sistema adjuvante RIBI® (RAS) (Ribi Immunochem, Hamilton, MT) contendo 2% de esqualeno, 0,2% de Tween 80 e um ou mais componentes da parede celular bacteriana pertencentes ao grupo que consiste em monofosforil-lípido A (MPL), dimicolato de trealose (TDM) e esqueleto da parede celular (CWS), preferencialmente MPL+CWS (Detox®); (3) adjuvantes de saponinas, tal como o Stimulon® (Cambridge Bioscience, Worcester, MA), ou partículas criadas a partir destes, tal como os ISCOMS (complexos imunoestimuladores); (4) Adjuvante Completo de Freund (CFA) e Adjuvante Incompleto de Freund (IFA); (5) citocinas, tais como interleucinas (p.ex., IL-1, IL-2, IL-4, IL-5, IL-6, IL-7, IL-12, etc.), interferões (por exemplo, interferão gama), fator estimulador de colônias de macrófagos (M-CSF), fator de necrose de tumor (TNF), etc.; (6) mutantes destoxificados de uma toxina ADP-ribosilante bacteriana tal como uma toxina da cólera (CT), uma toxina pertússica (PT) ou uma toxina termolábil de E. coli (LT), particularmente LT-K63, LT-R72, CT-S109, PT-K9/G129; ver, por exemplo, W093/13302 e W092/19265; e (7) outras substâncias que atuam como agentes imunoestimuladores para aumentar a eficácia da composição. É dada preferência ao alum e ao MF59.

Tal como acima se mencionou, os muramil-péptidos incluem, mas de modo não limitativo, N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-normuramil-L-alanil-D-isoglutamina (nor-MDP), N-acetilmuramil-L-alanil-D-isoglutaminil-L-alanina-2-(1'-2'-dipalmitoil-*sn*-glicero-3-hidroxifosforiloxi)-etilamina (MTP-PE), etc.

As composições de vacinas compreendendo composições imunogénicas (por exemplo, que incluam o antigénio, o veículo

farmaceuticamente aceitável e o adjuvante) conterão tipicamente diluentes, tais como água, soro fisiológico, glicerol, etanol, etc. Adicionalmente, poderão encontrar-se presentes em tais veículos substâncias auxiliares, tais como agentes molhantes ou emulsificantes, substâncias tamponantes de pH e semelhantes. Alternativamente, as composições de vacinas podem compreender um antigénio, polipéptido, proteína, fragmento de proteína ou ácido nucleico num veículo farmaceuticamente aceitável.

Mais especificamente, as vacinas compreendendo composições imunogénicas compreendem uma quantidade imunologicamente eficaz dos polipéptidos imunogénicos, assim como de qualquer dos componentes acima mencionados, segundo requerido. Por "quantidade imunologicamente eficaz" pretende-se significar que a administração de tal quantidade a um indivíduo, em dose única ou como parte de uma série de doses, é eficaz para tratamento ou prevenção. Esta quantidade varia segundo o estado de saúde e as condições físicas do indivíduo a ser tratado, o grupo taxonómico do indivíduo a ser tratado (por exemplo, um primata não humano, um primata, etc.) a capacidade de produção de anticorpos evidenciada pelo sistema imunológico do indivíduo, o grau de proteção desejada, a formulação da vacina, a avaliação da situação clínica por parte do médico e outros fatores relevantes. Será de esperar que esta quantidade se encontre dentro de um intervalo relativamente alargado que poderá ser determinado através de ensaios de rotina.

Tipicamente, as composições de vacinas ou composições imunogénicas são preparadas sob a forma de injetáveis, em soluções líquidas ou suspensões; poderão ainda preparar-se formas sólidas adequadas para solução, ou suspensão, em veículos líquidos antes da injeção. A preparação poderá igualmente ser emulsificada ou encapsulada em lipossomas para



acentuar o seu efeito adjuvante, tal como acima se referiu para os veículos farmacêuticamente aceitáveis.

As composições imunogénicas são convencionalmente administradas por via parentérica, por exemplo, por injeção, subcutânea ou intramuscular. As formulações adicionais adequadas para outros modos de administração incluem as formulações orais e pulmonares, os supositórios e as aplicações transdérmicas e transcutâneas. A dosagem de tratamento poderá ser administrada num regime de dose única ou num regime de doses múltiplas. A vacina poderá ser administrada em conjunção com outros agentes imunoreguladores.

Como alternativa às vacinas baseadas em proteínas, poderá usar-se a vacinação com ADN (ver, p.ex., Robinson & Torres (1997) *Seminars in Immunology* 9:271-283; Donnelly et al. (1997) *Annu Rev Immunol* 15:617-648).

#### Veículos para Administração de Genes

Os veículos para terapia genética destinados à administração de construções, incluindo uma sequência codificante e um agente terapêutico da invenção, poderão ser administrados localmente ou sistemicamente a um mamífero para que se dê a expressão no mamífero. Estas construções poderão utilizar abordagens com vetores virais ou não virais em modalidades *in vivo* ou *ex vivo*. A expressão desta sequência codificante pode ser induzida por meio de promotores endógenos de mamífero ou de promotores heterólogos. A expressão da sequência codificante *in vivo* poderá ser constitutiva ou regulada.

A invenção inclui veículos para administração de genes que possuem a capacidade de expressar as sequências de ácido nucleico contempladas. O veículo para administração de genes corresponde preferencialmente a um vetor viral e, mais

preferencialmente, a um vetor retroviral, adenoviral, viral adeno-associado (AAV), de herpesvírus ou de alfavírus. O vetor viral poderá ainda corresponder a um astrovírus, coronavírus, ortomixovírus, papovavírus, paramixovírus, parvovírus, picornavírus, poxvírus ou togavírus. Ver, de modo geral, Jolly (1994) *Cancer Gene Therapy* 1:51-64; Kimura (1994) *Human Gene Therapy* 5:845-852; Connelly (1995) *Human Gene Therapy* 6:185-193; e Kaplitt (1994) *Nature Genetics* 6:148-153.

Os vetores de retrovírus são bem conhecidos neste campo técnico, incluindo os retrovírus de tipo B, C e D, os retrovírus xenotrópicos (por exemplo, NZB-X1, NZB-X2 e NZB9-1 (ver O'Neill (1985) *J. Virol.* 53:160), retrovírus politrópicos, p.ex., MCF e MCF-MLV (ver Kelly (1983) *J. Virol.* 45:291), spumavírus e lentivírus. Ver *RNA Tumor Viruses*, 2ª Ed., Cold Spring Harbor Laboratory, 1985.

As diversas porções de um vetor retroviral para terapia genética poderão derivar de diferentes retrovírus. Por exemplo, as LTRs do retrovetor poderão derivar do Vírus do Sarcoma Murino, um local de ligação ao tARN de um Vírus do Sarcoma de Rous, um sinal de empacotamento de um Vírus da Leucemia Murina, e uma origem para a síntese da segunda cadeia de um Vírus da Leucose Aviária.

Estes vetores retrovirais recombinantes poderão ser usados para gerar partículas de vetor retroviral competentes para transdução, recorrendo à sua introdução em linhas celulares de empacotamento apropriadas (ver a Patente US 5591624). Os vetores de retrovírus poderão ser construídos para integração específica de local no ADN da célula hospedeira através da incorporação de uma enzima integrase quimérica na partícula retroviral (ver W096/37626). Prefere-se que o vetor viral

recombinante corresponda a um vírus recombinante de replicação defetiva.

As linhas celulares de empacotamento adequadas para uso em conjunção com os vetores de retrovírus acima descritos são bem conhecidos da técnica, são facilmente preparados (ver W095/30763 e W092/05266) e podem ser usados para criar linhas celulares produtoras (também designadas por linhas celulares de vetor ou "LCVs") para a produção de partículas de vetor recombinante. Preferencialmente, as linhas celulares de empacotamento são produzidas a partir de células progenitoras humanas (por exemplo, células HT1080) ou linhas celulares progenitoras de arminho, o que elimina a inativação no soro humano.

Os retrovírus preferidos para a construção de vetores retrovirais para terapia genética incluem os Vírus da Leucose Aviária, Vírus da Leucemia Bovina, Vírus da Leucemia Murina, Vírus Indutor de Focos em Células de Arminho, Vírus do Sarcoma Murino, Vírus da Reticuloendoteliose e Vírus do Sarcoma de Rous. Os Vírus de Leucemia Murina particularmente preferidos incluem os vírus 4070A e 1504A (Hartley e Rowe (1976) J. Virol. 19:19-25), Abelson (N° ATCC VR-999), Friend (N° ATCC VR-245), Graffi, Gross (N° ATCC VR-590), Kirsten, Vírus do Sarcoma de Harvey e Rauscher (N° ATCC VR-998) e Vírus da Leucemia Murina de Moloney (N° ATCC VR-190). Tais retrovírus poderão ser obtidos a partir de depósitos ou colecções tais como a American Type Culture Collection ("ATCC") de Rockville, Maryland, ou ser isolados a partir de fontes conhecidas usando técnicas largamente difundidas neste campo.

Os exemplos de vetores retrovirais para terapia genética conhecidos e utilizáveis neste invenção incluem os descritos nos pedidos de patente GB2200651, EP0415731, EP0345242,

EP0334301, WO89/02468; WO89/05349, WO89/09271, WO90/02806, WO90/07936, WO94/03622. WO93/25698, WO93/25234, WO93/11230, WO93/10218, WO91/02805, WO91/02825, WO95/07994, US 5219740, US 4405712, US 4861719, US 4980289, US 4777127, US 5591624. Ver também Vile (1993) Cancer Res 53:3860-3864; Vile (1993) Cancer Res 53:962-967; Ram (1993) Cancer Res 53 83-88; Takamiya (1992) J Neurosci Res 33:493-503; Baba (1993) J Neurosurg 79:729-735; Mann (1983) Cell 33:153; Cane (1984) Proc Natl Acad Sci 81:6349; e Miller (1990) Human Gene Therapy 1.

Os vetores de adenovírus humanos para terapia genética são igualmente conhecidos neste campo técnico e utilizáveis nesta invenção. Ver, por exemplo, Berkner (1988) Biotechniques 6:616 e Rosenfeld (1991) Science 252:431, e as patentes WO93/07283, WO93/06223 e WO93/07282. Como exemplos de vetores adenovirais para terapia genética conhecidos e utilizáveis nesta invenção encontram-se os descritos nos documentos acima referidos e nas patentes WO94/12649, WO93/03769, WO93/19191, WO94/28938, WO95/11984, WO95/00655, WO95/27071, WO95/29993, WO95/34671, WO96/05320, WO94/08026, WO94/11506, WO93/06223, WO94/24299, WO95/14102, WO95/24297, WO95/02697, WO94/28152, WO94/24299, WO95/09241, WO95/25807, WO95/05835, WO94/18922 e WO95/09654. Alternativamente, poderá ser utilizada a administração de ADN ligado a adenovírus morto, tal como descrito em Curiel (1992) Hum. Gene Ther. 3:147-154. Os veículos para administração de genes de acordo com a invenção incluem vetores de vírus associado a adenovírus (AAV). Os principais, e preferidos, exemplos de tais vetores para uso nesta invenção correspondem aos vetores baseados em AAV-2 que são apresentados em Srivastava, WO93/09239. Os vetores AAV particularmente preferidos compreendem as duas sequências repetitivas terminais invertidas de AAV nas quais as sequências D nativas são modificadas por substituição de nucleótidos, de tal modo que, de preferência, pelo menos 5 nucleótidos nativos e até 18

nucleótidos nativos, mais preferencialmente pelo menos 10 nucleótidos nativos e até 18 nucleótidos nativos, e ainda mais preferencialmente 10 nucleótidos nativos são mantidos, sendo os restantes nucleótidos da sequência D retirados ou substituídos com nucleótidos não nativos. As sequências D nativas das sequências repetitivas terminais invertidas de AAV correspondem a sequências de 20 nucleótidos consecutivos em cada sequência repetitiva terminal invertida de AAV (i.e., existe uma sequência em cada extremidade) que não se encontram envolvidos na formação do HP. O nucleótido não nativo de substituição poderá corresponder a qualquer nucleótido que seja distinto do nucleótido que é encontrado na sequência D nativa na mesma posição. Como exemplos de outros vetores AAV utilizáveis encontram-se o pWP-19 e o pWN-1, ambos descritos em Nahreini (1993) Gene 124:257-262. O vetor psub201 constitui outro exemplo de um vetor AAV (ver Samulski (1987) J. Virol. 61:3096). Ainda outro exemplo de vetor AAV é dado pelo vetor ITR Double-D. A construção do vetor ITR Double-D é exposta na Patente dos EUA nº 5.478.745. Outros vetores encontram-se expostos nas patentes dos EUA nos 4.797.368 (de Carter), 5.139.941 (de Muzyczka), 5.474.935 (de Chatterjee) e WO94/288157 (de Kotin). Ainda outro exemplo de um vetor AAV utilizável nesta invenção corresponde ao vetor SSV9AFABTKneo, o qual contém o potenciador de AFP e o promotor de albumina e dirige a expressão predominantemente no fígado. A sua estrutura e construção são apresentadas em Su (1996) Human Gene Therapy 7:463-470. Outros vetores AAV para terapia genética encontram-se descritos nas patentes dos EUA nos 5.354.678, 5.173.414, 5.139.941 e 5.252.479.

Os vetores para terapia genética compreendendo sequências desta invenção incluem igualmente vetores de herpesvírus. Os exemplos principais e preferidos correspondem a vetores de vírus herpes simples que contêm uma sequência codificando para

um polipéptido de timidina quinase, tal como os descritos nas patentes US 5.288.641 e EP0176170 (Roizman). Outros exemplos de vetores de vírus herpes simplex incluem os vetores HFEM/ICP6-LacZ apresentadas na patente W095/04139 (Wistar Institute), pHSVlac, descrito em Geller (1988) Science 241:1667-1669, em W090/09441 e em W092/07945; HSV Us3::pgC-lacZ, descrito em Fink (1992) Human Gene Therapy 3:11-19; HSV 7134, 2 RH 105 e GAL4, descritos em EP 0453242 (Breakefield); e os vetores depositados na ATCC com os números de acesso ATCC VR-977 e ATCC VR-260.

Está igualmente contemplado neste invenção o uso de vetores de alfavírus para terapia genética. Os vetores de alfavírus preferidos são os vetores de vírus Sindbis, Togavírus, vírus de Semliki Forest (ATCC VR-67; ATCC VR-1247), vírus de Middleberg (ATCC VR-370), vírus de Ross River (ATCC VR-373; ATCC VR-1246), vírus da encefalite equina venezuelana (ATCC VR-923; ATCC VR-1250; ATCC VR-1249; ATCC VR-532) e os vírus descritos nas patentes US 5091309, 5217879 e W092/10578. Mais particularmente, são utilizáveis os vetores de alfavírus descritos na patente dos EUA com o nº de série 08/405.627, registada em 15 de Março de 1995, nas patentes W094/21792, W092/10578 e W095/07994 e nas patentes dos EUA nos 5.091.309 e 5.217.879. Tais alfavírus poderão ser obtidos a partir de depósitos ou coleções tais como a American Type Culture Collection ("ATCC") de Rockville, Maryland, ou ser isolados a partir de fontes conhecidas usando técnicas largamente difundidas neste campo. Serão preferencialmente usados vetores de alfavírus com reduzida citotoxicidade (ver USSN 08/679640).

Os sistemas vetores de ADN tais como os sistemas eucarióticos de expressão em camadas são igualmente úteis para a expressão dos ácidos nucleicos da invenção. Ver a W095/07994 para obter uma descrição detalhada dos sistemas eucarióticos de expressão

em camadas. Preferencialmente, os sistemas eucarióticos de expressão em camadas de acordo com esta invenção derivam de vetores de alfavírus e, mais preferencialmente, de vetores de vírus Sindbis.

Outros vetores virais adequados para uso no âmbito da presente invenção incluem os derivados de poliovírus, por exemplo o vírus ATCC VR-58 e os descritos em Evans, Nature 339 (1989) 385 e Sabin (1973) J. Biol. Standardization 1:115; o rinovírus, como por exemplo o ATCC VR-1110 e os descritos em Arnold (1990) J Cell Biochem L401; o poxvírus como o poxvírus de canário ou o vacciniavírus, por exemplo os vírus ATCC VR-111 e ATCC VR-2010 e os descritos em Fisher-Hoch (1989) Proc Natl Acad Sci 86:317; Flexner (1989) Ann NY Acad Sci 569:86, Flexner (1990) Vaccine 8:17 e nas patentes US 4.603.112, US 4.769.330 e WO89/01973; vírus SV40, por exemplo o ATCC VR-305 e os descritos em Mulligan (1979) Nature 277:108 e Madzak (1992) J Gen Virol 73:1533; o influenzavírus, por exemplo o ATCC VR-797 e os influenzavírus recombinantes produzidos por técnicas de genética reversa, tal como descrito na patente dos EUA 5.166.057 e em Enami (1990) Proc Natl Acad Sci 87:3802-3805; Enami & Palese (1991) J Virol 65:2711-2713 e Luytjes (1989) Cell 59:110 (ver também McMichael (1983) NEJ Med 309:13, e Yap (1978) Nature 273:238 e (1979) Nature 277:108); o vírus da imunodeficiência humana, tal como descrito na patente EP-0386882 e em Buchschacher (1992) J. Virol. 66:2731; o vírus do sarampo, por exemplo os ATCC VR-67 e VR-1247 e os descritos na EP-0440219; o vírus Aura, por exemplo o ATCC VR-368; o vírus Bebaru, por exemplo os ATCC VR-600 e ATCC VR-1240; o vírus Cabassou, por exemplo o ATCC VR-922; o vírus Chikungunya, por exemplo os ATCC VR-64 e VR-1241; o vírus de Fort Morgan, por exemplo o ATCC VR-924; o vírus Getah, por exemplo os ATCC VR-369 e ATCC VR-1243; o vírus Kyzylagach, por exemplo o ATCC VR-927; o vírus Mayaro, por exemplo o ATCC VR-

66; o vírus Mucambo, por exemplo os ATCC VR-580 e ATCC VR-1244; o vírus Ndumu, por exemplo o ATCC VR-371; o vírus Pixuna, por exemplo os ATCC VR-372 e ATCC VR-1245; o vírus Tonate, por exemplo o ATCC VR-925; o vírus Trinitite, por exemplo o ATCC VR-469; o vírus Una, por exemplo o ATCC VR-374; o vírus Whataroa, por exemplo o ATCC VR-926; o vírus Y-62-33, por exemplo o ATCC VR-375; o vírus O'Nyong, o vírus da encefalite do Oriente, por exemplo os ATCC VR-65 e ATCC VR-1242; o vírus da encefalite do Ocidente, por exemplo os ATCC VR-70, ATCC VR-1251, ATCC VR-622 e ATCC VR-1252; e coronavírus, por exemplo o ATCC VR-740 e os descritos em Hamre (1966) Proc Soc Exp Biol Med 121:190.

A administração das composições da invenção a células não se encontra limitada aos vetores virais acima mencionados. Poderão utilizar-se outros métodos e meios de libertação, tal como, por exemplo, vetores de expressão de ácidos nucleicos, ADN policatiónico condensado ligado ou não ligado a adenovírus morto (ver, por exemplo, a patente dos EUA com o n° de série 08/366.787, registada a 30 de Dezembro de 1994 e Curiel (1992) Hum Gene Ther 3:147-154), ADN ligado a ligando (ver, por exemplo, Wu (1989) J Biol Chem 264:16985-16987), veículos celulares de administração a células eucarióticas (ver, por exemplo, a patente dos EUA com o n° de série 08/240.030, de 9 de Maio de 1994, e a patente dos EUA com o n° de série 08/404.796), deposição de materiais de hidrogel fotopolimerizados, pistola de partículas manual para transferência de genes (descrita na Patente US 5149655), radiação ionizante (tal como descrito na Patente US 5206152 e em W092/11033), neutralização de cargas nucleicas ou fusão com membranas celulares. Encontram-se descrições de outras abordagens em Philip (1994) Mol Cell Biol 14:2411-2418 e em Woffendin (1994) Proc Natl Acad Sci 91:1581-1585.



Poderá utilizar-se a transferência de genes mediada por partículas; ver, por exemplo, a Patente US com o N° de série 60/023.867. Abreviadamente, a sequência poderá ser inserida em vetores convencionais que contêm sequências de controle convencionais para uma expressão em elevado grau, procedendo-se então a uma incubação com moléculas sintéticas para transferência de genes, tais como catiões poliméricos que se ligam ao ADN, como poli-lisina, protamina e albumina, ligadas a ligandos que se direcionam para as células, tal como o asialoorosomucóide, como se descreve em Wu & Wu (1987) J. Biol. Chem. 262:4429-4432, insulina, tal como descrito em Hucked (1990) Biochem Pharmacol 40:253-263, galactose, como se descreve em Plank (1992) Bioconjugate Chem 3:533-539, lactose ou transferrina.

Poderá igualmente utilizar-se ADN nu para transformar uma célula hospedeira. Encontram-se exemplos de métodos para introdução de ADN nu nas patentes W090/11092 e US 5.580.859. A eficiência de captação poderá ser melhorada usando esferas biodegradáveis de látex. As esferas de látex revestidas por ADN são transportadas de modo eficiente para o interior das células após a iniciação da endocitose pelas próprias esferas. O método poderá ainda ser melhorado por tratamento das esferas para aumentar a sua hidrofobicidade e, deste modo, facilitar a rutura do endossoma e a libertação do ADN para o citoplasma.

Encontram-se descrições de lipossomas que podem atuar como veículos para libertação de genes na patente US 5422120 e nas patentes W095/13796, W094/23697, W091/14445 e EP-524.968. Tal como se descreve na USSN 60/023.867, relativamente à administração por via não viral, as sequências de ácido nucleico codificando para um polipéptido poderão ser inseridas em vetores convencionais que contêm sequências de controle convencionais para permitir uma expressão em elevado grau,

procedendo-se então à incubação com moléculas sintéticas para transferência de genes tais como catiões poliméricos que se ligam ao ADN, como poli-lisina, protamina e albumina, ligadas a ligandos que se direccionam para as células, tal como o asialoorosomucóide, insulina, galactose, lactose ou transferrina. Outros sistemas de administração incluem o uso de lipossomas para a encapsulação de ADN compreendendo o gene sob o controlo de diversos promotores específicos de tecido ou ubiquitariamente ativos. Outros sistemas de administração não virais adequados incluem os sistemas de administração mecânica, como a abordagem descrita em Woffendin et al. (1994) Proc Natl Acad. Sci USA 91(24):11581-11585. Adicionalmente, a sequência codificante e o produto de expressão da mesma poderão ser administrados por meio de deposição de materiais de hidrogel fotopolimerizados. Outros métodos convencionais para administração de genes que poderão ser usados para a administração da sequência codificante incluem, por exemplo, o uso de uma pistola de partículas manual para transferência de genes, tal como se descreve na patente US 5149655; o uso de radiação ionizante para a ativação do gene transferido, tal como descrito na patente US 5206152 e em WO92/11033.

Como exemplos de veículos lipossómicos ou polycatiónicos para transferência de genes encontram-se os descritos nas patentes US 5422120 e 4762915; em WO95/13796; WO94/23697; WO91/1444; EP-0524968; e em Stryer, Biochemistry, págs. 236-240 (1975), W.H. Freeman, San Francisco; Szoka (1980) Biochem Biophys Acta 600:1; Bayer (1979) Biochem Biophys Acta 550:464; Rivnay (1987) Meth Enzymol 149:119; Wang (1987) Proc Natl Acad Sci 84:7851; Plant (1989) Anal Biochem 176:420.

Uma composição polinucleotídica poderá conter uma quantidade terapêuticamente eficaz de um veículo para terapia genética, de acordo com a definição do termo acima apresentada. Para os

fins do presente invenção, uma dose eficaz encontrar-se-á entre cerca de 0,01 mg/kg e 50 mg/kg, ou entre 0,05 mg/kg e cerca de 10 mg/kg, de construções de ADN para o indivíduo ao qual serão administradas.

#### Métodos de Administração

Uma vez formuladas, as composições polinucleotídicas da invenção serão administradas (1) diretamente ao indivíduo; (2) *ex vivo*, a células derivadas do indivíduo; ou (3) *in vitro*, para expressão das proteínas recombinantes. Os indivíduos a tratar poderão ser mamíferos ou aves. Poderão ainda tratar-se seres humanos.

A administração direta das composições será geralmente efetuada através de injeção, subcutânea, intraperitoneal, intravenosa ou intramuscular, ou de libertação para o espaço intersticial de um tecido. As composições poderão ainda ser administradas a um tumor ou lesão. Outros modos de administração incluem a administração oral e pulmonar, supositórios e aplicações transdérmicas, agulhas e pistolas de genes ou hiposprays. A posologia poderá corresponder a um regime de dose única ou a um regime de múltiplas doses.

Os métodos para administração *ex vivo* e reimplantação das células transformadas num indivíduo são já conhecidas da técnica e encontram-se descritas em, por exemplo, WO93/14778. Os exemplos de células úteis para aplicações *ex vivo* incluem, por exemplo, células estaminais, particularmente células hematopoiéticas, células linfáticas, macrófagos, células dendríticas ou células tumorais.

De modo geral, a administração de ácidos nucleicos através de aplicações *ex vivo* e *in vitro* poderá ser conseguida recorrendo, por exemplo, aos seguintes procedimentos bem

conhecidos neste campo técnico: transfecção mediada por dextrano, precipitação com fosfato de cálcio, transfecção mediada por polibreno, fusão de protoplastos, electroporação, encapsulação do(s) polinucleótido(s) em lipossomas e microinjeção direta do ADN nos núcleos celulares.

#### Composições farmacêuticas de polinucleótidos e polipéptidos

Para além dos sais e veículos farmacêuticamente aceitáveis acima descritos, poderão ainda usar-se os agentes que se seguem em conjunção com composições de polinucleótidos e/ou polipéptidos.

##### A. Polipéptidos

Como exemplos destes polipéptidos incluem-se, sem limitação: asialoorosomucóide (ASOR); transferrina; asialoglicoproteínas; anticorpos; fragmentos de anticorpos; ferritina; interleucinas; interferões, factor estimulador de colónias de granulócitos e macrófagos (GM-CSF), factor estimulador de colónias de granulócitos (G-CSF), factor estimulador de colónias de macrófagos (M-CSF), factor de células estaminais e eritropoietina. Poderão igualmente usar-se antigénios virais, tal como proteínas de envelope. Adicionalmente, poderão usar-se proteínas provenientes de outros organismos invasivos, tal como o péptido de 17 aminoácidos, designado por RII, da proteína do circumsporozoíto de *Plasmodium falciparum*.

##### B. Hormonas, Vitaminas, etc.

Outros grupos que poderão ser incluídos correspondem, por exemplo, a: hormonas, esteroides, androgénios, estrogénios, hormona da tiroide, vitaminas, ácido fólico.

#### C. Polialquilenos, Polissacáridos, etc.

Adicionalmente, poderá incluir-se um polialquilenoglicol em associação aos polinucleótidos ou polipéptidos desejados. Numa forma de realização preferida, o polialquilenoglicol corresponde ao polietilenoglicol. Adicionalmente, poderão incluir-se mono-, di- ou polissacáridos. Numa forma de realização preferida segundo este aspeto, o polissacárido corresponde a dextrano ou a DEAE-dextrano. Poderão ainda utilizar-se quitosano e poli(lactido-co-glicólido).

#### D. Lípidos e Lipossomas

O polinucleótido ou polipéptido desejado poderá igualmente ser encapsulado em lípidos ou empacotado em lipossomas antes de ser administrado ao indivíduo ou a células derivadas deste.

A encapsulação em lípidos é geralmente conseguida usando lipossomas que tenham a capacidade de se ligar a ácido nucleico, ou a capturá-lo e retê-lo, de modo estável. A razão de polinucleótido ou polipéptido condensado para preparação de lípido poderá variar, mas situa-se geralmente à volta de 1:1 (mg ADN:micromoles lípido), ou maior quantidade de lípido. Para consulta de uma revisão sobre o uso de lipossomas como veículos para libertação de ácidos nucleicos, ver Hug e Sleight (1991) *Biochim. Biophys. Acta* 1097:1-17; Straubinger (1983) *Meth. Enzymol.* 101:512-527.

As preparações de lipossomas destinadas a uso na presente invenção incluem preparações catiónicas (carregadas positivamente), aniónicas (carregadas negativamente), e neutras. Verificou-se que os lipossomas catiónicos medeiam a libertação intracelular de ADN plasmídico (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (Malone (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); e fatores de

transcrição purificados (Debs (1990) J. Biol. Chem. 265:10189-10192), em forma funcional.

Os lipossomas catiónicos são facilmente adquiríveis no mercado. Por exemplo, os lipossomas de N[1-2,3-dioleiloxi)propil]-N,N,N-trietilamónio (DOTMA) encontram-se disponíveis sob o nome comercial de Lipofectin, da empresa Gibco BRL, Grand Island, NY (ver também Felgner, supra). Outros lipossomas comercialmente disponíveis incluem o Transfectace (DDAB/DOPE) e o DOTAP/DOPE (Boehringer). Poderão preparar-se outros lipossomas catiónicos a partir de materiais prontamente disponíveis utilizando técnicas bem conhecidas neste campo. Ver, por exemplo, Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198, e W090/111092, para obter uma descrição da síntese de lipossomas de DOTAP (1,2-bis(oleiloxi)-3-(trimetilamónio)propano).

De forma similar, os lipossomas aniónicos e neutros encontrar-se-ão facilmente disponíveis no mercado, a partir de empresas como a Avanti Polar Lipids (Birmingham, AL), ou poderão ser facilmente preparados a partir de materiais prontamente disponíveis. Tais materiais incluem a fosfatidil-colina, colesterol, fosfatidil-etanolamina, dioleilfosfatidil-colina (DOPC), dioleilfosfatidil-glicerol (DOPG), dioleilfosfatidil-etanolamina (DOPE), entre outros. Estes materiais poderão igualmente ser misturados com os materiais de partida para DOTMA e DOTAP em razões apropriadas. Os métodos para o fabrico de lipossomas utilizando estes materiais são bem conhecidos da técnica.

Os lipossomas poderão compreender vesículas multilamelares (MLVs), vesículas unilamelares pequenas (SUVs), ou vesículas unilamelares grandes (LUVs). Os diversos complexos de lipossoma-ácido nucleicos são preparados usando métodos

conhecidos neste campo. Ver, por exemplo, Straubinger (1983) Meth. Immunol. 101:512-527; Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198; Papahadjopoulos (1975) Biochim. Biophys. Acta 394:483; Wilson (1979) Cell 17:77); Deamer & Bangham (1976) Biochim. Biophys. Acta 443:629; Ostro (1977) Biochem. Biophys. Res. Commun. 76:836; Fraley (1979) Proc. Natl. Acad. Sci. USA 76:3348); Enoch & Strittmatter (1979) Proc. Natl. Acad. Sci. USA 76:145; Fraley (1980) J. Biol. Chem. (1980) 255:10431; Szoka & Papahadjopoulos (1978) Proc. Natl. Acad. Sci. USA 75:145; e Schaefer-Ridder (1982) Science 215:166.

#### E. Lipoproteínas

Adicionalmente, poderão incluir-se lipoproteínas na composição de polinucleótido ou polipéptido a ser administrada. Os exemplos de lipoproteínas que podem ser utilizadas incluem: quilomicra, HDL, IDL, LDL e VLDL. Poderão igualmente usar-se mutantes, fragmentos ou fusões destas proteínas. Poderão ainda usar-se modificações de lipoproteínas de ocorrência natural, tal como LDL acetilada. Estas lipoproteínas podem dirigir a administração de polinucleótidos para células que expressem receptores para lipoproteínas. De preferência, se as lipoproteínas se encontrarem associadas aos polinucleótidos a administrar, não será incluído outro ligando direcionador na composição.

As lipoproteínas de ocorrência natural compreendem um lípido e uma porção de proteína. As porções de proteína são designadas por apoproteínas. Até ao presente, foram já isoladas e identificadas as apoproteínas A, B, C, D e E. Pelo menos duas destas contêm diversas proteínas, designadas por números romanos: AI, AII, AIV; CI, CII, CIII.

Uma lipoproteína poderá conter mais do que uma apoproteína. Por exemplo, as quilomicra de ocorrência natural compreendem as apoproteínas A, B, C e E; ao longo do tempo, estas lipoproteínas perdem a apoproteína A e adquirem as apoproteínas C e E. A lipoproteína VLDL compreende as apoproteínas A, B, C e E; a LDL compreende a apoproteína B e a HDL compreende as apoproteínas A, C e E.

Os aminoácidos destas apoproteínas são conhecidos e encontram-se descritos em, por exemplo, Breslow (1985) *Annu Rev. Biochem* 54:699; Law (1986) *Adv. Exp Med. Biol.* 151:162; Chen (1986) *J Biol Chem* 261:12918; Kane (1980) *Proc Natl Acad Sci USA* 77:2465; e Utermann (1984) *Hum Genet* 65:232.

As lipoproteínas contêm uma variedade de lípidos, incluindo triglicéridos, colesterol (livre e esterificado) e fosfolípidos. A composição dos lípidos varia nas lipoproteínas de ocorrência natural. Por exemplo, as quilomicra compreendem principalmente triglicéridos. Poderá encontrar-se uma descrição mais detalhada do conteúdo em lípidos das lipoproteínas de ocorrência natural em, por exemplo, *Meth. Enzymol.* 128 (1986). A composição dos lípidos é escolhida de modo a auxiliar a conformação da apoproteína relativamente à sua atividade de ligação ao recetor. A composição em lípidos poderá igualmente ser escolhida de modo a facilitar a interação hidrofóbica e a associação com a molécula de ligação ao polinucleótido.

As lipoproteínas de ocorrência natural poderão ser isoladas a partir do soro através de ultracentrifugação, por exemplo. Tais métodos encontram-se descritos em *Meth. Enzymol.* (supra); Pitas (1980) *J. Biochem.* 255:5454-5460 e Mahey (1979) *J Clin. Invest* 64:743-750.



As lipoproteínas podem igualmente ser produzidas por métodos *in vitro* ou recombinantes, através da expressão dos genes de apoproteínas numa célula hospedeira desejada. Ver, por exemplo, Atkinson (1986), *Annu Rev Biophys Chem* 15:403 e Radding (1958) *Biochim Biophys Acta* 30:443.

As lipoproteínas poderão ainda ser adquiridas a fornecedores como a empresa Biomedical Technologies, Inc., Stoughton, Massachusetts, EUA.

Poderá encontrar-se uma descrição mais detalhada sobre lipoproteínas em Zuckermann et al., pedido de patente PCT n° US97/14465.

#### F. Agentes Policatiónicos

Poderão incluir-se agentes policatiónicos, com ou sem lipoproteínas, numa composição com o polinucleótido ou polipéptido a administrar.

Os agentes policatiónicos exibem tipicamente uma carga positiva bruta a um pH fisiologicamente relevante e têm a capacidade de neutralizar a carga elétrica dos ácidos nucleicos, facilitando deste modo a sua administração a um local desejado. Estes agentes são utilizáveis em aplicações *in vitro*, *ex vivo* ou *in vivo*. Os agentes policatiónicos poderão ser usados para administrar ácidos nucleicos a um indivíduo vivo através das vias intramuscular, subcutânea, etc.

Como exemplos de polipéptidos úteis como agentes policatiónicos encontram-se os seguintes: polilisina, poliarginina, poliornitina e protamina. Outros exemplos incluem histonas, protaminas, albumina sérica humana, proteínas de ligação ao ADN, proteínas cromossómicas não-histónicas, proteínas de revestimento de vírus ADN, tais como

a proteína X174; os fatores transcricionais contêm igualmente domínios que se ligam ao ADN e que, deste modo, poderão ser úteis como agentes condensadores de ácidos nucleicos. Abreviadamente, os fatores transcricionais como C/CEBP, c-jun, c-fos, AP-1, AP-2, AP-3, CPF, Prot-1, Sp-1, Oct-1, Oct-2, CREP e TFIID contêm domínios básicos que se ligam a sequências de ADN.

Os agentes orgânicos policationicos incluem: espermina, espermidina e putrescina.

As dimensões e propriedades físicas de um agente policationico poderão ser extrapoladas a partir da lista acima tendo em vista a construção de outros agentes policationicos polipeptídicos ou a produção de agentes policationicos sintéticos.

#### Agentes Policationicos Sintéticos

Os agentes policationicos sintéticos com utilidade incluem, por exemplo, DEAE-dextrano e polibreno. Lipofectin® e LipofectAMINE® são monómeros que formam complexos policationicos quando combinados com polinucleótidos ou polipéptidos.

#### Ensaio para Imunodiagnóstico

Os antígenos de *Neisseria* da invenção poderão ser usados em imunoensaios para detetar as concentrações de anticorpo (ou, inversamente, os anticorpos anti-*Neisseria* podem ser usados para detetar as concentrações de antígeno). Poderão desenvolver-se imunoensaios baseados em antígenos recombinantes bem definidos para substituir os métodos invasivos de diagnóstico. Será possível detetar anticorpos contra proteínas de *Neisseria* em amostras biológicas, incluindo, por exemplo, amostras de sangue ou de soro. A

conceção dos imunoensaios abrange um grande intervalo de possibilidades, sendo bastantes destas variações já bem conhecidas neste campo técnico. Os protocolos para o imunoensaio poderão basear-se, por exemplo, em ensaios de competição, de reação direta ou de sanduíche. Os protocolos poderão também usar, por exemplo, suportes sólidos ou o processo de imunoprecipitação. A maioria dos ensaios envolve o uso de anticorpo ou polipéptido marcados; os marcadores poderão ser, por exemplo, fluorescentes, quimioluminescentes, radioativos ou moléculas de corante. Conhecem-se igualmente ensaios que amplificam os sinais provenientes da sonda; como exemplos destes ensaios encontram-se os que utilizam biotina e avidina, e os imunoensaios marcados com, e mediados por, enzimas, tais como os ensaios de ELISA.

Os kits adequados para imunodiagnóstico e contendo os reagentes marcados apropriados são construídos embalando os diversos materiais, incluindo as composições do presente invenção, em contentores apropriados, juntamente com os restantes reagentes e materiais (por exemplo, tampões adequados, soluções salinas, etc.) que são requeridos para a execução do ensaio, assim como um conjunto apropriado de instruções para o ensaio.

#### Hibridização de Ácidos Nucleicos

O termo "hibridização" refere-se à associação de duas sequências de ácido nucleico entre si por meio de ligações de hidrogénio. Tipicamente, uma das sequências estará fixada a um suporte sólido e a outra encontrar-se-á livre em solução. Seguidamente, as duas sequências serão colocadas em contacto uma com a outra sob condições que favoreçam o estabelecimento de ligações de hidrogénio. Os fatores que afetam esta ligação incluem: tipo e volume do solvente; temperatura de reação; tempo de hibridização; agitação; agentes para bloqueio da

ligação não específica da sequência presente na fase líquida ao suporte sólido (reagente de Denhart ou BLOTTO); concentração das sequências; uso de compostos para aumentar a taxa de associação das sequências (sulfato de dextrano ou polietilenoglicol); e a restringência das condições de lavagem após a hibridação. Ver Sambrook et al. [supra], Volume 2, capítulo 9, págs. 9.47 a 9.57.

O termo "restringência" refere-se a condições, numa reação de hibridização, que favorecem a associação de sequências muito similares relativamente à de sequências divergentes. Por exemplo, a combinação de temperatura e de concentração de sais deverá ser escolhida de modo a que a temperatura seja inferior em aproximadamente 120 a 200°C à do  $T_m$  calculado do híbrido em estudo. A temperatura e condições salinas poderão frequentemente ser determinadas de modo empírico através de experiências preliminares no decurso das quais se hibridizam amostras de ADN genómico imobilizadas sobre filtros com a sequência de interesse, procedendo-se então a lavagens sob condições de diferentes restringências. Ver Sambrook et al., pág. 9.50.

As variáveis a considerar para a realização de, por exemplo, um Southern blot são (1) a complexidade do ADN a ser submetido a blot e (2) a homologia entre a sonda e as sequências a detectar. A quantidade total do(s) fragmento(s) a estudar poderá variar numa magnitude de 10, desde 0,1 a 1  $\mu\text{g}$  para um produto de digestão de plasmídeo ou fago até  $10^{-9}$  a  $10^{-8}$  g para um gene de uma única cópia num genoma eucariótico altamente complexo. Para polinucleótidos de complexidade inferior poderão usar-se tempos de blot, hibridização e exposição substancialmente mais curtos, uma quantidade menor de polinucleótidos de partida e uma menor atividade específica das sondas. Por exemplo, um gene de levedura de uma só cópia

poderá ser detetado com um tempo de exposição de apenas 1 hora, partindo de 1 µg de ADN de levedura e efetuando um blot de 2 horas e uma hibridização durante 4-8 horas com uma sonda com  $10^8$  cpm/µg. Para um gene de mamífero de uma só cópia, uma abordagem conservadora poderá partir de 10 µg de ADN, seguindo-se um blot de 12 horas e uma hibridização durante 12 horas na presença de sulfato de dextrano a 10%, usando uma sonda com mais do que  $10^8$  cpm/µg, o que resulta num tempo de exposição de ~ 24 horas.

Diversos factores poderão afetar a temperatura de fusão ( $T_m$ ) de um híbrido ADN-ADN entre a sonda e o fragmento de interesse e, consequentemente, as condições apropriadas para hibridização e lavagem. Em muitos casos a sonda não é 100% homóloga relativamente ao fragmento. Outras variáveis a ter em conta incluem a extensão e conteúdo total em G+C das sequências de hibridização e a força iónica e conteúdo em formamida do tampão de hibridização. Os efeitos de todos estes fatores poderão ser englobados por uma só equação:

$$T_m = 81 + 16,6(\log_{10} C_i) + 0,4[\%(G+C)] - 0,6(\% \text{formamida}) - 600/n - 1,5, \\ (\% \text{ não correspondência})$$

em que  $C_i$  corresponde à concentração salina (iões monovalentes) e  $n$  corresponde à extensão do híbrido em pares de bases (fórmula ligeiramente modificada a partir de Meinkoth e Wahl (1984) Anal. Biochem. 138:267-284).

Na conceção de um protocolo experimental de hibridização, será possível alterar de modo conveniente alguns dos fatores que afetam a hibridização dos ácidos nucleicos. A temperatura de hibridização e das lavagens, e a concentração em sais durante as lavagens, são os fatores de ajuste mais simples. À medida que a temperatura de hibridização aumenta (i.e., que a

restringência aumenta), vai-se tornando menos provável a ocorrência de hibridização entre cadeias não homólogas e, como resultado, o efeito de fundo diminui. Se a sonda radiomarcada não for completamente homóloga relativamente ao fragmento imobilizado (tal como acontece frequentemente em experiências de hibridização de genes em famílias e interespecies), a temperatura de hibridização terá de ser reduzida, o que resulta no aumento do efeito de fundo. A temperatura das lavagens afeta de modo similar a intensidade da banda de hibridização e a intensidade do efeito de fundo. A restringência das lavagens é igualmente aumentada quando se reduzem as concentrações de sais.

De modo geral, as temperaturas de hibridização convenientes na presença de 50% de formamida correspondem a 42°C para uma sonda que possua 95% a 100% de homologia com o fragmento-alvo, 37°C para uma homologia de 90% a 95% e 32°C para uma homologia de 85% a 90%. Com homologias inferiores, o conteúdo em formamida deverá ser reduzido e a temperatura adequadamente ajustada, usando a equação acima apresentada. Se a homologia entre a sonda e o fragmento-alvo não for conhecida, a abordagem mais simples corresponderá a começar com condições de hibridização e lavagem não restritivas. Se se verificar a ocorrência de bandas não específicas ou um elevado efeito de fundo após autoradiografia, o filtro poderá ser lavado em condições de elevada restringência e reexposto. Se o tempo requerido para a exposição tornar esta abordagem impraticável, deverão testar-se em paralelo diversas restringências de hibridização e/ou lavagem.

#### Ensaaios com Sondas de Ácidos Nucleicos

Os métodos como a PCR, os ensaios de sonda de ADN ramificada ou as técnicas de blotting usando sondas de ácido nucleico de acordo com o invenção têm a capacidade de determinar a

presença de cADN ou mARN. Diz-se que uma sonda se “hibridiza” com uma sequência do invencão se formar com esta um duplex ou um complexo de dupla cadeia que seja suficientemente estável para detecção.

As sondas de ácido nucleico hibridizar-se-ão com as sequências nucleotídicas de *Neisseria* da invencão (incluindo tanto a cadeia de senso como a de anti-senso). Se bem que muitas sequências nucleotídicas diferentes codifiquem para a sequência de aminoácidos, a sequência nativa de *Neisseria* é preferida por corresponder à sequência que de facto se encontra presente nas células. O mARN representa uma sequência codificante, pelo que uma sua sonda deverá ser complementar à sequência codificante; o cADN de cadeia simples é complementar ao mARN, pelo que uma sonda de cADN deverá ser complementar à sequência não codificante.

A sequência da sonda não necessita de ser idêntica à sequência de *Neisseria* (ou ao seu complemento) - alguma variação de sequência e extensão poderá conduzir ao aumento da sensibilidade do ensaio se a sonda de ácido nucleico tiver a capacidade de formar um duplex com os nucleótidos-alvo, o qual poderá ser detetado. Adicionalmente, a sonda de ácido nucleico poderá incluir nucleótidos adicionais para estabilizar o duplex formado. Poderá tornar-se útil o recurso a uma sequência adicional de *Neisseria* como marcador para a detecção do duplex formado. Por exemplo, poderá ligar-se uma sequência nucleotídica não complementar à extremidade 5' da sonda, sendo a parte restante da sequência da sonda complementar a uma sequência de *Neisseria*. Alternativamente, poderão intercalar-se na sonda bases não complementares ou sequências mais longas, desde que a sequência da sonda exiba complementaridade suficiente com a sequência de *Neisseria* para que com esta se

hibridize, deste modo formando um duplex que poderá ser detetado.

A extensão e sequência exatas da sonda dependerão das condições de hibridização como a temperatura, condições salinas, etc. Por exemplo, para aplicações em diagnóstico, dependendo da complexidade da sequência do analito, a sonda de ácido nucleico conterá tipicamente pelo menos 10-20 nucleótidos, preferencialmente 15-25, e mais preferencialmente pelo menos 30 nucleótidos, se bem que possa ser mais curta. Os iniciadores curtos requerem geralmente temperaturas mais baixas para que possam formar complexos híbridos suficientemente estáveis com o molde.

As sondas poderão ser produzidas por procedimentos de síntese, tal como o método de triéster de Matteucci et al. [J. Am. Chem. Soc. (1981) 103:3185], ou de acordo com Urdea et al. [Proc. Natl. Acad. Sci. USA (1983) 80:7461], ou usando sintetizadores automatizados de oligonucleótidos comercialmente disponíveis.

A natureza química da sonda poderá ser selecionada segundo pretendido. Para certas aplicações, o ADN ou o RNA são adequados. Para outras aplicações, poderão incorporar-se algumas modificações - p.ex. modificações da estrutura primária, como a formação de fosforotioatos ou metilfosfonatos - para aumentar a semi-vida *in vivo*, alterar a afinidade para o ARN, aumentar a resistência às nucleases, etc. [ver, p.ex., Agrawal e Iyer (1995) Curr Opin Biotechnol 6:12-19; Agrawal (1996) TIBTECH 14:376-387]; poderão também ser utilizados análogos, como por exemplo ácidos nucleicos peptídicos também podem ser utilizados [ver, p.ex., Corey (1997) TIBTECH 15:224-229; Buchardt et al. (1993) TIBTECH 11:384-386].



Um exemplo de um ensaio de hibridização de nucleótidos foi descrito por Urdea et al. no pedido de patente internacional W092/02526 [ver igualmente a patente dos EUA 5.124.246].

Alternativamente, a reação em cadeia por polimerase (PCR) constitui um outro meio bem conhecido para a detecção de pequenas quantidades de ácidos nucleicos alvo. O ensaio encontra-se descrito em Mullis et al. (Meth. Enzymol. (1987) 155:335-350) e nas patentes dos EUA nos. 4.683.195 e 4.683.202. Dois nucleótidos "iniciadores" hibridizam-se com os ácidos nucleicos alvo e são usados como iniciadores para a reação. Os iniciadores poderão compreender uma sequência que não se hibridize com a sequência do alvo de amplificação (ou com o seu complemento) mas que ajude a estabilizar o duplex ou, por exemplo, que incorpore um local de restrição conveniente. Tipicamente, uma tal sequência flanqueará a sequência de *Neisseria* desejada.

Uma polimerase termoestável cria cópias dos ácidos nucleicos alvo a partir dos iniciadores usando os ácidos nucleicos alvo originais como molde. Quando a quantidade de ácidos nucleicos alvo gerados pela polimerase ultrapassa o limiar de detecção, estes poderão ser detetados através de métodos mais tradicionais, tais como o método de Southern Blot. Pelo método de Southern Blot, a sonda marcada hibridizar-se-á com a sequência de *Neisseria* (ou o seu complemento).

Adicionalmente, o mRNA ou o cADN poderão ser detetados através de técnicas convencionais de blotting como as descritas em Sambrook et al. [*supra*]. O mRNA, ou o cADN gerado a partir de mRNA usando uma enzima polimerase, poderá ser purificado e separado usando a electroforese em gel. Os ácidos nucleicos no gel são então transferidos por blot para um suporte sólido, tal como nitrocelulose. O suporte sólido é exposto a uma sonda

marcada e seguidamente lavado para remover qualquer sonda não hibridizada. Seguidamente, os duplexes contendo a sonda marcada são detetados. Tipicamente, a sonda é marcada com uma porção molecular radioativa.

### Exemplo

O exemplo que se segue descreve sequências de ácidos nucleicos que foram identificados em *N.meningitidis* e em *N.gonorrhoeae*, bem como os seus respetivos produtos de tradução deduzidos. Nem todas as sequências de ácido nucleico estão completas, i.e., estas codificam para menos do que a extensão completa da proteína de tipo selvagem.

O exemplo encontra-se de modo geral no seguinte formato:

- Uma sequência de nucleótidos que foi identificada em *N.meningitidis*
- O produto de tradução deduzido para a dita sequência de *N.meningitidis*
- Uma análise por computador do dito produto de tradução, com base em comparações entre bases de dados
- Uma sequência de nucleótidos correspondente que foi identificada em *N.gonorrhoeae*
- O produto de tradução deduzido para a dita sequência de *N.gonorrhoeae*
- Uma comparação de percentagens de identidade entre o produto de tradução da sequência de *N.meningitidis* e o produto de tradução da sequência de *N.gonorrhoeae*
- Uma sequência de nucleótidos correspondente que foi identificada no serogrupo A de *N.meningitidis*.
- O produto de tradução deduzido para a dita sequência do serogrupo A de *N.meningitidis*
- Uma comparação de percentagens de identidade entre o produto de tradução da sequência de *N.meningitidis* e o produto de tradução da sequência de *N.gonorrhoeae*

- Uma descrição das características da proteína que indica que esta poderá ser adequadamente antigénica ou imunogénica.

As comparações de sequência foram realizadas no NCBI (<http://www.ncbi.nlm.nih.gov>), usando os algoritmos BLAST, BLAST2, BLASTn, BLASTp, tBLASTn, BLASTx, e tBLASTx [p.ex. ver também Altschul et al. (1997) Gapped BLAST and PSI-BLAST: a new generation or protein database search programs. *Nucleic Acids Research* 25:2289-3402]. As pesquisas foram realizadas recorrendo às seguintes bases de dados: sequências não-redundantes GenBank+EMBL+DDBJ+PDB e produtos de tradução não-redundantes GenBank CDS +sequências PDB+SwissProt+SPupdate+PIR.

Para a dedução da presença de domínios hidrofóbicos, procedeu-se ao varrimento das sequências de nucleótidos em todos os seis moldes de leitura usando um algoritmo baseado nos estudos estatísticos de Esposti et al. [Critical evaluation of the hydropathy of membrane proteins (1990) *Eur J Biochem* 190:207-219]. Estes domínios representam potenciais regiões transmembranares ou sequências leader hidrofóbicas.

Os moldes abertos de leitura foram deduzidos a partir de fragmentos de sequências nucleotídicas, usando o programa ORFFINDER (NCBI).

As sequências de aminoácidos sublinhadas indicam possíveis domínios transmembranares ou sequências leader nos ORF, conforme previsto por recurso ao algoritmo PSORT (<http://www.psort.nibb.ac.jp>). Também os domínios funcionais foram deduzidos recorrendo ao programa MOTIFS (GCG Wisconsin & PROSITE).

Para o exemplo seguinte: tomando como base a presença de uma presumível sequência leader e/ou de vários presumíveis domínios transmembranares (a sublinhado simples) na proteína do gonococo, prevê-se que as proteínas de *N.meningitidis* e de *N.gonorrhoeae*, e os seus respetivos epitopos, possam ser úteis como antigénios ou composições imunogénicas para aplicação em vacinas ou como meios de diagnóstico.

Encontra-se acima um resumo das técnicas e procedimentos padronizados que se poderão utilizar para a concretização da invenção (por exemplo, para usar as sequências apresentadas para fins de diagnóstico ou de vacinação). Este resumo não pretende limitar a invenção, mas sim oferecer exemplos que poderão ser utilizados, embora tal não seja imperativo.

Os métodos a seguir descritos foram utilizados especialmente para expressar, purificar e caracterizar bioquimicamente as proteínas deste invenção.

#### Preparação do ADN Cromossómico

A estirpe 2996 de *N.meningitidis* foi cultivada até à fase exponencial em 100 ml de meio GC, recolhida por centrifugação e ressuspensa em 5 ml de tampão (Sacarose a 20% (p/v), Tris-HCl 50 mM e EDTA 50 mM, pH 8,0). Passados 10 minutos de incubação no gelo, as bactérias foram lisadas através da adição de 10 ml de solução de lise (NaCl 50 mM, Na-Sarkosyl a 1% e Proteinase K a 50 µg/ml), e a suspensão foi incubada a 37°C durante 2 horas. Foram realizadas duas extrações com fenol (equilibrado a pH 8) e uma com CHCl<sub>3</sub>/álcool isoamílico (24:1). O ADN foi precipitado por adição de acetato de sódio 0,3M e de 2 volumes de etanol, sendo então recolhido por centrifugação. O pellet foi lavado uma vez com etanol a 70% (v/v) e foi redissolvido em 4 ml de tampão TE (Tris-HCl 10 mM,

EDTA 1mM, pH 8,0). Mediu-se a concentração de ADN através de leitura da DO a 260 nm.

### Conceção de oligonucleótidos

Foram concebidos iniciadores oligonucleotídicos sintéticos com base na sequência codificante de cada ORF, usando (a) a sequência de meningococo B, quando disponível; (b) a sequência de gonococo/meningococo A, adaptada ao uso do códon de preferência do meningococo, segundo necessário. Ao conceber os iniciadores 5' de modo a que se sequenciem imediatamente a jusante da sequência leader prevista, foi possível omitir todos os péptidos de sinal previstos.

Para a maior parte dos ORFS, os iniciadores 5' incluíram dois locais de reconhecimento para enzimas de restrição (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, ou *Eco*RI-*Nde*I ou *Eco*RI-*Nhe*I), dependendo do padrão de restrição do gene de interesse. Os iniciadores 3' incluíram um local de restrição para *Xho*I ou para *Hind*III (tabela 1). Este procedimento foi estabelecido de modo a dirigir a clonagem de cada produto de amplificação (correspondendo a cada ORF) em dois sistemas de expressão diferentes: pGEX-KG (usando *Bam*HI-*Xho*I, *Bam*HI-*Hind*III, *Eco*RI-*Xho*I ou *Eco*RI-*Hind*III), e pET21b+ (usando *Nde*I-*Xho*I, *Nhe*I-*Xho*I, *Nde*I-*Hind*III ou *Nhe*I-*Hind*III).

5'-end primer tail:	<u>CGCGGATCCCATATG</u>	( <i>Bam</i> HI- <i>Nde</i> I)
	<u>CGCGGATCCGCTAGC</u>	( <i>Bam</i> HI- <i>Nhe</i> I)
	<u>CCGGAATTCTACATATG</u>	( <i>Eco</i> RI- <i>Nde</i> I)
	<u>CCGGAATTCTAGCTAGC</u>	( <i>Eco</i> RI- <i>Nhe</i> I)
3'-end primer tail:	<u>CCCGCTCGAG</u>	( <i>Xho</i> I)
	<u>CCCGCTCGAG</u>	( <i>Hind</i> III)

Para a clonagem dos ORFs no vetor pGEX-His, os iniciadores 5' e 3' continham apenas um local de restrição (*EcoRI*, *KpnI* ou *SalI* para os iniciadores 5' e *PstI*, *XbaI*, *SphI* ou *SalI* para os iniciadores 3'). De novo, os locais de restrição foram escolhidos de acordo com o padrão de restrição particular do gene (tabela 1).

5'-end primer tail:	(AAA)AAAGAATTC	( <i>EcoRI</i> )
	(AAA)AAAGGTACC	( <i>KpnI</i> )
3'-end primer tail:	(AAA)AAACTGCAG	( <i>PstI</i> )
	(AAA)AAATCTAGA	( <i>XbaI</i> )
	AAAGCATGC	( <i>SphI</i> )
5' or 3'-end primer tail:	AAAAAAGTCGAC	( <i>SalI</i> )

Os iniciadores incluíram não só as sequências de reconhecimento para enzimas de restrição como também nucleótidos que se hibridizaram com a sequência a amplificar. A temperatura de fusão dependeu do número e do tipo dos nucleótidos hibridizantes no iniciador completo, e foi determinado para cada iniciador usando a seguinte fórmula:

$$T_m = 4 (G+C) + 2 (A+T) \text{ (cauda excluída)}$$

$$T_m = 64,9 + 0,41 (\%GC) - 600/N \text{ (iniciador completo)}$$

As temperaturas médias de fusão dos oligonucleótidos selecionados foram usualmente de 65-70°C para o oligonucleótido completo e de 50-55°C para a região de hibridização apenas.

A Tabela 1 apresenta os iniciadores dianteiro e reverso usados para cada amplificação. Em certos casos, a sequência do iniciador não corresponde exatamente à sequência do ORF previsto. Tal acontece porque, quando foram realizadas as

amplificações iniciais, as sequências 5' e/ou 3' completas para alguns ORFs de meningococo B não eram conhecidas. No entanto, as sequências correspondentes já haviam sido identificadas no gonococo ou no meningococo A. Assim, nos casos em que a sequência do meningococo B estava incompleta ou era incerta, foram usadas as sequências do gonococo ou do meningococo A como base para a concepção do iniciador. Estas sequências foram alteradas de acordo com o códon de preferência. Deve ter-se em conta que, uma vez identificada a sequência completa, esta abordagem já não será necessária.

Os oligonucleótidos foram sintetizados recorrendo a um sintetizador de ADN/RNA Perkin Elmer 394, eluídos das colunas usando 2 ml  $\text{NH}_4\text{OH}$ , e desprotegidos por incubação durante 5 horas a 56 °C. Os oligonucleótidos foram precipitados através da adição de acetato de Na 0,3M e 2 volumes de etanol. As amostras foram então centrifugadas e os pellets ressuspensos em 100µl ou em 1 ml de água. A  $\text{DO}_{260}$  foi determinada recorrendo a um espectrofotómetro Perkin Elmer Lambda Bio e a concentração foi ajustada para 2-10 pmol/µl.

### Amplificação

Foi usado o seguinte protocolo padronizado de PCR: utilizou-se 50-200 ng de ADN genómico como molde na presença de 20-40 µM de cada primer oligonucleotídico, dNTPs em solução a 400-800 µM, 1 x tampão PCR (incluindo  $\text{MgCl}_2$  1,5 mM), 2,5 unidades de ADN polimerase *TaqI* (usando AmpliTaq da Perkin-Elmer, Platinum da GIBCO, ADN polimerase Pwo, ou Taq polimerase Tahara Shuzo). Em alguns casos, a reacção de PCR foi optimizada por adição de 10µl de DMSO ou de 50 µl de betáina 2M.

Após um passo inicial a quente (adicionando-se a polimerase durante uma incubação preliminar de toda a mistura durante 3 minutos a 95°C), cada amostra foi submetida a uma amplificação

em dois passos. Os primeiros 5 ciclos foram realizados usando a temperatura de hibridização que corresponde aos primers sem cauda para enzimas de restrição (ver acima). Seguiram-se então 30 ciclos efetuados à temperatura de hibridização calculada para os oligonucleótidos de extensão completa. Os ciclos foram completados por um passo final de extensão de 10 minutos a 72°C.

Os ciclos-padrão foram realizados da seguinte forma:

	<b>Desnaturação</b>	<b>Hibridização</b>	<b>Alongamento</b>
<b>Primeiros 5 ciclos</b>	30 segundos 95°C	30 segundos 50-55°C	30-60 segundos 72°C
<b>Últimos 30 ciclos</b>	30 segundos 95°C	30 segundos 65-70°C	30-60 segundos 72°C

O tempo de alongamento variou de acordo com o comprimento do ORF a amplificar. As amplificações foram realizadas usando os sistemas de PCR GeneAmp 9600 ou 2400 da Perkin Elmer. Para conferir os resultados, 1/10 do volume de amplificação foi carregado num gel de agarose a 1-1,5% (p/v) e a dimensão de cada fragmento amplificado foi então comparada com um marcador de peso molecular de ADN.

O ADN amplificado foi carregado directamente num gel de agarose a 1% ou, alternativamente, foi em primeiro lugar precipitado com etanol e ressuspenso num volume adequado para ser carregado num gel de agarose a 1,0%. O fragmento de ADN correspondente à banda de tamanho apropriado foi purificado usando o Gel Extraction Kit da Qiagen de acordo com as instruções do fabricante. Os fragmentos de ADN foram eluídos num volume de 30µl ou de 50µl de H<sub>2</sub>O ou de Tris 10mM, pH 8,5.



#### Digestão dos fragmentos de PCR

O ADN purificado que corresponde ao fragmento amplificado foi submetido a dupla digestão com as enzimas de restrição apropriadas para: clonagem em pET-21b+ e expressão subsequente da proteína como produto de fusão C-terminal com His; para clonagem em pGEX-KG e expressão subsequente da proteína como produto de fusão N-terminal com GST; e para clonagem em pGEX-His e expressão da proteína como produto de fusão N-terminal de GST-His.

Cada fragmento purificado de ADN foi incubado a 37°C, por um período de entre 3 a 12 horas, com 20 unidades da enzima de restrição apropriada (New England Biolabs) num volume final de 30 ou 40 µl e na presença do tampão de digestão apropriado. Os fragmentos digeridos foram purificados utilizando o Kit de Purificação para PCR QIAquick (de acordo com as indicações do fabricante), procedendo-se então à sua eluição num volume de 30 µl ou de 50 µl de H<sub>2</sub>O ou de Tris-HCl 10mM, pH 8,5. A concentração do ADN foi determinada por meio de electroforese quantitativa em gel de agarose (gel a 1,0%), na presença de um marcador de peso molecular titulado.

#### Digestão dos vetores de clonagem (pET22B, pGEX-KG, pTRC-His A, pET21b+, pGEX-KG e pGEX-His)

O vetor pGEX-His é um vetor pGEX-2T modificado possuindo uma região que codifica para seis resíduos de histidina a montante do local de corte da trombina e que contém o local múltiplo de clonagem do vetor pTRC99 (Pharmacia). Submeteu-se 10 µg do plasmídeo a digestão dupla com 50 unidades de cada enzima de restrição num volume de reação de 200 µl e na presença do tampão apropriado, durante um período de incubação de 12 horas a 37°C. Após o carregamento de todo o produto de digestão num gel de agarose a 1%, a banda correspondente ao vetor digerido foi purificada a partir do gel utilizando o QIAquick Gel

Extraction Kit da Qiagen e o ADN foi eluído em 50 µl de Tris-HCl 10 mM, pH 8,5. A concentração de ADN foi avaliada através da medição da  $DO_{260}$  da amostra, sendo então ajustada a 50 µg/µl. Utilizou-se 1 µl de plasmídeo por cada procedimento de clonagem.

Submeteu-se 10 µg do vetor plasmídico a digestão dupla com 50 unidades de cada enzima de restrição num volume de reação de 200 µl e na presença do tampão apropriado, durante um período de incubação de 12 horas a 37°C. O produto de digestão foi carregado sobre um gel de agarose a 1,0% e a banda correspondente ao vetor digerido foi purificada usando o QIAquick Gel Extraction Kit da Qiagen. O ADN foi eluído em 50 µl de Tris-HCl 10 mM, pH 8,5. A concentração de ADN foi avaliada por medição da  $DO_{260nm}$  da amostra, sendo esta concentração ajustada a 50 µg/µl. Utilizou-se 1 µl de plasmídeo por cada procedimento de clonagem.

### Clonagem

Para alguns ORFs, os fragmentos correspondentes a cada ORF, previamente digeridos e purificados, foram ligados em pET22b bem como em pGEX-KG. Num volume final de 20 µl, o procedimento de ligação efetuou-se com uma razão molar fragmento/vetor de 3:1 e utilizando 0,5 µl de ADN ligase NEB T4 (400 unidades/µl), na presença do tampão fornecido pelo fabricante. A reação foi incubada à temperatura ambiente durante 3 horas. Em algumas experiências, a ligação foi realizada utilizando o "Rapid Ligation Kit", da Boehringer, de acordo com as indicações do fabricante.

De modo a introduzir o plasmídeo recombinante numa estirpe adequada, 100 µl de células competentes de *E. coli* DH5 foram incubadas com a solução de reação de ligase em gelo durante 40 minutos, seguindo-se uma incubação a 37°C durante 3 minutos, e

seguidamente, após a adição de 800 µl de caldo LB, uma nova incubação a 37°C durante 20 minutos. As células foram então centrifugadas a velocidade máxima num microcentrífuga Eppendorf e ressuspensas em aproximadamente 200 µl do sobrenadante. A suspensão foi então plaqueada em LB-ampicilina (100 mg/ml).

A deteção dos clones recombinantes foi realizada recolhendo 5 colónias escolhidas ao acaso, as quais foram cultivadas por um período de 12 horas a 37°C em 2 ml (clones pGEX ou pTC) ou em 5ml (clones pET) de caldo LB + ampicilina a 100 µg/ml. Foi então preparado um pellet das células e o ADN foi extraído utilizando o QIAprep Spin Miniprep Kit da Qiagen, de acordo com as indicações do fabricante, até ser atingido um volume final de 30 µl. 5 µl de cada "miniprep" individual (aproximadamente 1g) foram digeridos com *NdeI/XhoI* ou com *BamHI/XhoI* e todo o produto de digestão foi então carregado num gel de agarose a 1-1,5% (dependendo do tamanho esperado para o fragmento inserido), em paralelo com o marcador de peso molecular (ADN de cadeia dupla de 1 Kb, GIBCO). A deteção dos clones positivos foi realizada com base no tamanho correto do fragmento inserido.

Para outros ORFs, o fragmento correspondente a cada ORF, previamente digerido e purificado, foi ligado ao pET21b+ e ao pGEX-KG. Foi usada uma razão molar fragmento/vetor de 3:1 num volume final de 20 µl, incluindo 0,5 µl de ADN ligase T4 (400 unidades/µl, NEB) e tampão de ligação fornecido pelo fabricante. A reação foi efetuada à temperatura ambiente durante 3 horas. Em algumas experiências, a ligação foi efetuada usando o "Rapid Ligation Kit" da Boehringer segundo o protocolo do fabricante.

O plasmídeo recombinante foi transformado em 100 µl de células competentes de *E. coli* DH5 ou HB101 por incubação da solução de reação da ligase e da bactéria durante 40 minutos em gelo, seguindo-se um passo a 37°C durante 3 minutos. Este passo foi seguido pela adição de 800 µl de caldo LB e por uma incubação a 37°C durante 20 minutos. As células foram centrifugadas a velocidade máxima numa microcentrífuga Eppendorf, sendo ressuspensas em aproximadamente 200 µl do sobrenadante e plaqueadas em agar LB-ampicilina (100 mg/ml).

A pesquisa de clones recombinantes foi efetuada cultivando 5 colónias selecionadas ao acaso durante 12 horas a 37°C em 2,0 ml (clones de pGEX-KG) ou 5,0 ml (clones de pET) de caldo LB + 100 µg/ml de ampicilina. Preparou-se um pellet das células e o ADN plasmídico foi extraído usando o kit QIAprep Spin Miniprep da Qiagen, seguindo as instruções do fabricante. Aproximadamente 1 µg de cada miniprep individual foi digerido com as enzimas de restrição apropriadas e o produto de digestão foi carregado sobre um gel de agarose a 1-1,5% (dependendo da dimensão esperada para o fragmento inserido), em paralelo com o marcador de peso molecular (cadeia dupla de ADN de 1 kb, GIBCO). Os clones positivos foram selecionados com base na dimensão do fragmento inserido.

Os ORFs foram clonados no plasmídeo pGEX-His através de digestão dupla do produto de PCR e de ligação ao vetor digerido de modo similar. Após a clonagem, os plasmídeos recombinantes foram transformados na célula hospedeira *E. coli* W3110. Os clones individuais foram cultivados durante 12 horas a 37°C em caldo LB com 50 µg/ml de ampicilina.

Alguns ORFs podem ser clonados no vetor pGEX-HIS utilizando os locais de clonagem para *EcoRI-PstI*, ou *EcoRI-SalI*, ou *SalI-*

*Pst*I. Após a clonagem, os plasmídeos recombinantes podem ser introduzidos na célula hospedeira *E.coli* W3110.

### Expressão

Cada ORF clonado no vetor de expressão pode então ser transformado na estirpe adequada à expressão do produto proteico recombinante. Utilizou-se 1 µl de cada construção para transformar 30 µl de *E.coli* BL21 (vetor pGEX), *E.coli* TOP 10 (vetor pTRC) ou *E.coli* BL21-DE3 (vetor pET), de acordo com o acima descrito. No caso do vetor pGEX-His, foi utilizada a mesma estirpe de *E.coli* (W3110) quer para a clonagem inicial quer para a expressão. As colónias recombinantes isoladas foram inoculadas em 2ml de LB+Amp (100 µg/ml), tendo sido incubadas a 37°C por um período de 12 horas, e seguidamente diluídas a 1:30 em 20 ml de LB+Amp (100 µg/ml), em frascos de incubação com capacidade de 100 ml, tendo sido assegurado que a DO<sub>600</sub> desta diluição inicial se encontrava entre 0,1 e 0,15. Os frascos foram incubados a 30°C em agitadores giratórios de banho de água até que a DO indicasse um crescimento exponencial adequado para a indução da expressão (0,4-0,8 de DO para os vetores pET e pTRC; 0,8-1 de DO para os vetores pGEX e pGEX-His). A expressão de proteínas nos vetores pET, pTRC e pGEX-His foi induzida por adição de IPTG 1mM, enquanto no caso do sistema pGEX a concentração final de IPTG foi de 0,2 mM. Após um período de incubação de 3 horas a 30°C, a concentração final da amostra foi verificada por determinação da DO. Para a avaliação da expressão, 1ml de cada amostra foi recolhido e centrifugado em microcentrífuga, sendo o pellet resultante ressuspenso em PBS e analisado por SDS-PAGE a 12% corado por azul de Coomassie. Toda a amostra foi centrifugada a 6000g e o pellet foi ressuspenso em PBS para utilização subsequente.

### Purificação em larga escala de proteínas de fusão com GST

Para alguns ORFs, cultivou-se uma única colónia durante 12 horas a 37°C numa placa de agar LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido LB + Amp num agitador de banho de água e cultivadas durante 12 horas. As bactérias foram diluídas a 1:30 em 600 ml de meio fresco e foram cultivadas a temperatura ótima (20-37°C) até uma DO<sub>550</sub> de 0,8-1. A expressão das proteínas foi induzida com IPTG 0,2mM, seguindo-se um período de incubação de 3 horas. A cultura foi centrifugada a 8000 rpm a 4°C. O sobrenadante foi rejeitado e o pellet de bactérias foi ressuspensão em 7,5 ml de PBS frio.

As células foram lisadas por sonicação em gelo com um sonicador Branson B-15 durante 30 segundos a 40W, sendo então congeladas e descongeladas por duas vezes e novamente centrifugadas. O sobrenadante foi recolhido e misturado com 150µl de resina de Glutathione-Sefarose 4B (Pharmacia) (previamente lavada com PBS), sendo então incubado à temperatura ambiente durante 30 minutos. A amostra foi centrifugada a 700g durante 5 minutos a 4°C. A resina foi lavada por duas vezes com 10 ml de PBS frio durante 10 minutos, ressuspensão em 1ml de PBS frio, e carregada numa coluna descartável. A resina foi lavada por duas vezes com 2 ml de PBS frio até que o eluído atingisse uma DO<sub>280</sub> de 0,02-0,06. A proteína de fusão com GST foi eluída pela adição de 700µl de tampão de eluição de Glutathione frio (glutathione reduzida 10mM, Tris-HCl 50mM) e foram recolhidas fracções até que a DO<sub>280</sub> atingisse 0,1. Carregou-se um volume de 21µl de cada fracção num gel de SDS a 12%, utilizando como padrões os padrões de peso molecular para SDS-PAGE de larga gama da Biorad (M1) (200, 116,25 , 97,4 , 66,2 , 45, 31, 21,5 , 14,4 , 6,5 kDa) ou o Amersham Rainbow Marker (M") (220, 66, 46, 30, 21,5 , 14,3 kDa). Dado que o peso molecular de GST é de 26kDa,

este valor tem de ser adicionado ao PM de cada proteína de fusão com GST.

Para outros ORFs, para cada clone a ser purificado como uma proteína de fusão com GST, uma só colónia foi recolhida e colocada em cultura durante 12 horas a 37°C numa placa de agar LB/Amp (100 µg/ml). Uma colónia isolada desta placa foi inoculada em 20 ml de meio líquido LB/Amp (100 µg/ml) e colocada em cultura durante 12 h a 37°C com agitação. A cultura de 12 h foi diluída a 1:30 em 600 ml de meio líquido LB/Amp (100 µg/ml) e foi deixada em cultura à temperatura ótima (20-37°C) até que a  $DO_{550nm}$  atingisse 0,6-0,8. A expressão de proteína recombinante foi induzida pela adição de IPTG (concentração final 0,2 mM), e a cultura foi incubada por um período adicional de 3 horas. As bactérias foram recolhidas por centrifugação a 8000xg durante 15 min a 4°C.

O pellet de bactérias foi ressuspenso em 7,5 ml de PBS frio. As células foram lisadas por sonicação em gelo por quatro vezes durante 30 segundos a 40W usando um sonicador Branson 450, sendo então centrifugadas a 13000xg durante 30 min a 4°C. O sobrenadante foi recolhido e misturado com 150µl de resina de Glutathione-Sefarose 4B (Pharmacia), previamente equilibrada com PBS, sendo então incubado à temperatura ambiente durante 30 minutos com agitação suave. O lote de preparação foi centrifugado a 700xg durante 5 minutos a 4°C e o sobrenadante foi rejeitado. A resina foi lavada por duas vezes (por lote) com 10 ml de PBS frio durante 10 minutos, ressuspensa em 1 ml de PBS frio, e carregada numa coluna descartável. Prosseguiu-se a lavagem da resina com PBS frio até que o eluído atingisse uma  $DO_{280nm}$  de 0,02-0,01. A proteína de fusão com GST foi eluída pela adição de 700µl de tampão de eluição de glutathione frio (glutathione reduzida 10mM, Tris-HCl 50mM, pH 8,0), e foram recolhidas fracções até que a  $DO_{280nm}$  do eluído indicou

ter sido obtida toda a proteína recombinante. Foram analisadas alíquotas de 20 µl de cada fração de eluição por SDS-PAGE usando um gel a 12%. A massa molecular das proteínas purificadas foi determinada utilizando os padrões de peso molecular de larga gama da Biorad (M1) (200, 116, 97,4, 66,2, 45,0, 31,0, 21,5, 14,4 , 6,5 kDa) ou o Amersham Rainbow Marker (M") (220, 66,2 , 46,0 , 30,0 , 21,5 , 14,3 kDa). Os pesos moleculares das proteínas de fusão com GST são uma combinação da proteína GST de 26 kDa e do seu parceiro de fusão. As concentrações de proteína foram estimadas usando o ensaio de Bradford.

#### Purificação a larga escala de proteínas de fusão com His solúveis

Para alguns ORFs, cultivou-se uma única colônia durante 12 horas a 37°C numa placa de agar LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido de LB + Amp e incubadas durante 12 horas num agitador de banho de água. As bactérias foram diluídas a 1:30 em 600 ml de meio fresco e foram cultivadas a temperatura ótima (20-37°C) até uma DO<sub>550</sub> de 0,6-0,8. A expressão das proteínas foi induzida com IPTG 1mM, seguindo-se um período de incubação adicional de 3 horas. A cultura foi centrifugada a 8000 rpm a 4°C, o sobrenadante foi rejeitado e o pellet de bactérias foi ressuspenso em 7,5 ml de tampão de imidazol 10 mM frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 10 mM, pH 8). As células foram lisadas por sonicação em gelo com um sonicador Branson B-15 durante 30 segundos a 40W, sendo então congeladas e descongeladas por duas vezes e novamente centrifugadas. O sobrenadante foi recolhido e misturado com 150µl de resina de Ni<sup>2+</sup> (Pharmacia) (previamente lavada com tampão de imidazol 10 mM), sendo então incubado à temperatura ambiente, com agitação suave, durante 30 minutos. A amostra foi centrifugada a 700g durante 5 minutos a 4°C. A resina foi lavada por duas vezes com 10 ml de



tampão de imidazol 10 mM frio durante 10 minutos, ressuspensa em 1ml de tampão de imidazol 10 mM frio, e carregada numa coluna descartável. A resina foi lavada a 4°C com 2 ml de tampão de imidazol 10 mM frio até que o eluído atingisse uma DO<sub>280</sub> de 0,02-0,06. A resina foi lavada com 2 ml de tampão de imidazol 20 mM frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 20 mM, pH 8) até que o eluído atingisse uma DO<sub>280</sub> de 0,02-0,06. A proteína de fusão com His foi eluída pela adição de 700µl de tampão de imidazol 250 mM frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 250 mM, pH 8) e foram recolhidas frações até que a DO<sub>280</sub> atingisse 0,1. Carregou-se um volume de 21µl de cada fração num gel de SDS a 12%.

#### Purificação a larga escala de proteínas de fusão com His insolúveis

Cultivou-se uma única colónia durante 12 horas a 37°C numa placa de agar LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido LB + Amp num agitador de banho de água e cultivadas durante 12 horas. As bactérias foram diluídas a 1:30 em 600 ml de meio fresco e foram cultivadas a temperatura ótima (37°C) até uma DO<sub>550</sub> de 0,6-0,8. A expressão das proteínas foi induzida por adição de IPTG 1mM, seguindo-se um período de incubação adicional de 3 horas. A cultura foi centrifugada a 8000 rpm a 4°C. O sobrenadante foi rejeitado e o pellet de bactérias foi ressuspensado em 7,5 ml de tampão B (ureia 8 M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 8,8). As células foram lisadas por sonicação em gelo com um sonicador Branson B-15 durante 30 segundos a 40W, sendo então congeladas e descongeladas por duas vezes e novamente centrifugadas. O sobrenadante foi armazenado a -20°C, tendo os pellets sido ressuspensos em 2 ml de tampão de guanidina (cloridrato de guanidina 6M, tampão fosfato 100 mM, Tris-HCl 10 mM, pH 7,5) e tratados num homogeneizador durante 10 ciclos. O produto foi então centrifugado a 13000 rpm durante 40 minutos. O

sobrenadante foi misturado com 150µl de resina de Ni<sup>2+</sup> (Pharmacia) (previamente lavada com tampão B), sendo então incubado à temperatura ambiente, com agitação suave, durante 30 minutos. A amostra foi centrifugada a 700g durante 5 minutos a 4°C. A resina foi lavada por duas vezes com 10 ml de tampão B durante 10 minutos, ressuspensa em 1ml de tampão B, e carregada numa coluna descartável. A resina foi lavada à temperatura ambiente com 2 ml de tampão B até que o eluído atingisse uma DO<sub>280</sub> de 0,02-0,06. A resina foi lavada com 2 ml de tampão C (ureia 8M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 6,3) até que o eluído atingisse uma DO<sub>280</sub> de 0,02-0,06. A proteína de fusão com His foi eluída pela adição de 700µl de tampão de eluição (ureia 8M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 4,5) e foram recolhidas frações até que a DO<sub>280</sub> atingisse 0,1. Carregou-se um volume de 21µl de cada fração num gel de SDS a 12%.

#### Purificação de proteínas de fusão com His

Para cada clone a ser purificado como uma proteína de fusão com His, uma só colónia foi recolhida e colocada em cultura durante 12 horas a 37°C numa placa de agar LB/Amp (100 µg/ml). Uma colónia isolada desta placa foi inoculada em 20 ml de meio líquido LB/Amp (100 µg/ml) e colocada em cultura durante 12 h a 37°C com agitação. A cultura de 12 h foi diluída a 1:30 em 600 ml de meio líquido LB/Amp (100 µg/ml) e foi deixada em cultura à temperatura ótima (20-37°C) até que a DO<sub>550nm</sub> atingisse 0,6-0,8. A expressão de proteína recombinante foi induzida pela adição de IPTG (concentração final 1,0 mM), e a cultura foi incubada por um período adicional de 3 horas. As bactérias foram recolhidas por centrifugação a 8000xg durante 15 min a 4°C.

O pellet de bactérias foi ressuspenso em 7,5 ml de (i) tampão A frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 10 mM, pH

8,0) para as proteínas solúveis ou (ii) tampão B (ureia 8 M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 8,8) para as proteínas insolúveis. As células foram lisadas por sonicação em gelo por quatro vezes durante 30 segundos a 40W usando um sonicador Branson 450, sendo então centrifugadas a 13000xg durante 30 min a 4°C. Para as proteínas insolúveis, os pellets foram ressuspensos em 2,0 ml de tampão C (cloridrato de guanidina 6 M, tampão fosfato 100 mM, Tris-HCl 10 mM, pH 7,5) e tratados com um homogeneizador Dounce durante 10 ciclos. O homogenato foi centrifugado a 13000xg durante 40 min e o sobrenadante foi reservado.

Os sobrenadantes das preparações de proteínas solúveis e das de proteínas insolúveis foram misturados com 150µl de resina de Ni<sup>2+</sup> (Pharmacia) (previamente equilibrada com tampão A ou com tampão B, segundo apropriado), sendo então incubados à temperatura ambiente, com agitação suave, durante 30 minutos. A resina usada foi a Chelating Sepharose Fast Flow (Pharmacia), preparada de acordo com o protocolo do fabricante. O lote de preparação foi centrifugado a 700xg durante 5 min a 4°C e o sobrenadante foi rejeitado. A resina foi lavada por duas vezes (por lote) com 10 ml de tampão A ou de tampão B durante 10 min, ressuspensa em 1,0 ml de tampão A ou B e carregada sobre uma coluna descartável. Prosseguiu-se a lavagem da resina com (i) tampão A a 4°C ou (ii) tampão B à temperatura ambiente, até que o eluído atingisse uma DO<sub>280nm</sub> de 0,02-0,01. A resina foi ainda lavada com (i) tampão C frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 20 mM, pH 8,0) ou (ii) tampão D (ureia 8M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 6,3) até que o eluído atingisse uma DO<sub>280nm</sub> de 0,02-0,01. A proteína de fusão com His foi eluída pela adição de 700µl de (i) tampão de eluição A frio (NaCl 300 mM, tampão fosfato 50 mM, imidazol 250 mM, pH 8,0) ou (ii) tampão de eluição B (ureia 8M, Tris-HCl 10 mM, tampão fosfato 100 mM, pH 4,5) e

foram recolhidas frações até que a  $DO_{280nm}$  do eluído indicasse ter sido obtida toda a proteína recombinante. Foram analisadas alíquotas de 20  $\mu$ l de cada fração de eluição por SDS-PAGE usando um gel a 12%. As concentrações de proteína foram estimadas usando o ensaio de Bradford.

#### Renaturação de proteínas de fusão com His

Nos casos em que foi necessária a desnaturação para solubilizar as proteínas, foi utilizado um passo de renaturação antes da imunização. Adicionou-se glicerol às frações desnaturadas acima obtidas para proporcionar uma concentração final de 10% (v/v). As proteínas foram diluídas até 200  $\mu$ g/ml recorrendo ao tampão de diálise I (glicerol 10% (v/v), arginina 0,5M, tampão fosfato 50mM, glutathiona reduzida 5,0mM, glutathiona oxidada 0,5mM, ureia 2M, pH 8.8) e foram submetidas a diálise contra o mesmo tampão a 4°C durante 12-14 horas. As proteínas foram novamente submetidas a diálise contra o tampão de diálise II (glicerol 10% (v/v), arginina 0,5M, tampão fosfato 50mM, glutathiona reduzida 5,0mM, glutathiona oxidada 0,5mM, pH 8.8) durante 12-14 horas a 4°C.

Alternativamente, adicionou-se glicerol a 10% às proteínas desnaturadas. As proteínas foram então diluídas até 20  $\mu$ g/ml recorrendo ao tampão de diálise I (glicerol a 10%, arginina 0,5M, tampão fosfato 50mM, glutathiona reduzida 5mM, glutathiona oxidada 0,5mM, ureia 2M, pH 8.8) e foram submetidas a diálise contra o mesmo tampão a 4°C durante 12-14 horas. As proteínas foram novamente dialisadas contra o tampão de diálise II (glicerol a 10%, arginina 0,5M, tampão fosfato 50mM, glutathiona reduzida 5mM, glutathiona oxidada 0,5mM, pH 8.8) durante 12-14 horas a 4°C.

A concentração de proteína foi determinada através da seguinte fórmula:

$$\text{Proteína (mg/ml)} = (1,55 \times \text{DO}_{280}) - (0,76 \times \text{DO}_{260})$$

### Purificação de proteínas

Para analisar a solubilidade, os pellets obtidos a partir de 3,0 ml das culturas foram ressuspensos em 500 µl de tampão M1 (PBS, pH 7,2). Adicionou-se 25 µl de lisozima (10 mg/ml) e as bactérias foram incubadas durante 15 min a 4°C. As células foram lisadas por sonicação em gelo por quatro vezes durante 30 segundos a 40W usando um sonicador Branson 450, sendo então centrifugadas a 13000xg durante 30 min a 4°C. O sobrenadante foi recolhido e o pellet foi ressuspenso em tampão M2 [ureia 8 M, NaCl 0,5 M, imidazol 20 mM e NaH<sub>2</sub>PO<sub>4</sub> 0,1 M], sendo então incubado durante 3 a 4 horas a 4°C. Após centrifugação, o sobrenadante foi recolhido e o pellet foi ressuspenso em tampão M3 [cloridrato de guanidínio 6 M, NaCl 0,5 M, imidazol 20 mM e NaH<sub>2</sub>PO<sub>4</sub> 0,1 M] durante 12 h a 4°C. Os sobrenadantes de todos os passos foram analisados por SDS-PAGE. Verificou-se que algumas das proteínas são solúveis em PBS e que outras necessitam de ureia ou de cloridrato de guanidínio para a solubilização.

Para efetuar purificações à escala preparativa, foram induzidas culturas de 500 ml e as proteínas de fusão foram solubilizadas em tampão M1, M2 ou M3 usando o procedimento acima descrito. Os extratos em bruto foram carregados sobre uma coluna de superfluxo Ni-NTA (Qiagen) equilibrada com tampão M1, M2 ou M3, dependendo do tampão de solubilização utilizado. O material não ligado foi eluído por lavagem da coluna com o mesmo tampão. A proteína de fusão recombinante foi eluída com o tampão correspondente contendo imidazol 500 mM, sendo então dialisada contra o mesmo tampão na ausência de imidazol.

### Imunização de ratinhos

Utilizaram-se 20µg de cada proteína purificada para imunizar ratinhos por via intraperitoneal. No caso de alguns ORFs, imunizaram-se ratinhos Balb-C utilizando Al(OH)<sub>3</sub> como adjuvante aos dias 1, 21 e 42, e as respostas imunitárias foram monitorizadas em amostras recolhidas ao dia 56. Para outros ORFs, imunizaram-se ratinhos CD1 usando o mesmo protocolo. Alternativamente, misturou-se 20 µg de cada proteína purificada com adjuvante de Freund e usou-se esta mistura para imunizar ratinhos CD1 por via intraperitoneal. Para muitas das proteínas, a imunização foi efetuada aos dias 1, 21 e 35, e a resposta imunitária foi monitorizada nas amostras recolhidas aos dias 34 e 49. Para algumas proteínas, a terceira imunização foi efetuada ao dia 28, e não ao dia 35, e a resposta imunitária foi medida aos dias 20 e 42, e não aos dias 34 e 49.

### Ensaio ELISA (análise de soros)

A estirpe acapsulada M7 de MenB foi semeada em placas de agar de chocolate e incubada durante 12 horas a 37°C. As colónias de bactérias foram recolhidas das placas de agar utilizando uma zaragatoa de dracon esterilizada, sendo então inoculadas em 7 ml de Caldo de Mueller-Hinton (Difco) contendo 0,25% de Glucose. O crescimento das bactérias foi monitorizado a cada 30 minutos seguindo a DO<sub>620</sub>. As bactérias foram deixadas em crescimento até que o valor de DO atingisse valores de 0,3-0,4. A cultura foi centrifugada durante 10 minutos a 10000 rpm. O sobrenadante foi rejeitado e as bactérias foram lavadas uma vez com PBS, ressuspensas em PBS contendo 0.025% de formaldeído e incubadas durante 2 horas à temperatura ambiente, seguindo-se um período de incubação de 12 horas a 4°C com agitação. Adicionaram-se 100µl de células bacterianas a cada poço de uma placa Greiner de 96 poços, incubando-se a placa durante 12 horas a 4°C. Os poços foram então lavados por

três vezes com tampão de lavagem PBT (Tween-20 a 0,1% em PBS). Adicionaram-se então 200 µl de tampão de saturação (Polivinilpirrolidona 10 a 2.7% em água) a cada poço e as placas foram incubadas durante 2 horas a 37°C. Os poços foram lavados por três vezes com PBT. Adicionaram-se 200 µl de soros diluídos (tampão de diluição: BSA a 1%, Tween-20 a 0,1%, NaN<sub>3</sub> a 0,1% em PBS) a cada poço e as placas foram incubadas durante 90 minutos a 37°C. Os poços foram lavados três vezes com PBT. Adicionaram-se a cada poço 100 µl de soro de coelho anti-ratinho conjugado com HRP (Dako) diluído a 1:2000 em tampão de diluição, e as placas foram incubadas durante 90 minutos a 37°C. Os poços foram lavados três vezes com tampão de PBT. Adicionaram-se 100 µl de tampão de substrato para HRP (25 ml de tampão de citrato pH 5, 10 mg of O-fenildiamina e 10 µl de H<sub>2</sub>O<sub>2</sub>) a cada poço e as placas foram deixadas à temperatura ambiente durante 20 minutos. Adicionaram-se 100 µl de H<sub>2</sub>SO<sub>4</sub> a cada poço e mediu-se a DO<sub>490</sub>. O teste ELISA foi considerado positivo quando se encontraram DO<sub>490</sub> 2,5 vezes superiores às determinadas com os respectivos soros de pré-imunização.

Alternativamente, a estirpe acapsulada M7 de MenB foi semeada em placas de agar de chocolate e incubada durante 12 horas a 37°C. As colónias de bactérias foram recolhidas das placas de agar utilizando uma zaragatoa de dracon esterilizada, sendo inoculadas em Caldo de Mueller-Hinton (Difco) contendo 0,25% de Glucose. O crescimento das bactérias foi monitorizado a cada 30 minutos seguindo a DO<sub>620</sub>. As bactérias foram deixadas em crescimento até que o valor de DO atingisse valores de 0,3-0,4. A cultura foi centrifugada durante 10 minutos a 10000 rpm. O sobrenadante foi rejeitado e as bactérias foram lavadas uma vez com PBS, ressuspensas em PBS contendo 0,025% de formaldeído e incubadas durante 1 hora a 37°C, seguindo-se um período de incubação de 12 horas a 4°C com agitação. Adicionaram-se 100µl de células bacterianas a cada poço de uma

placa Greiner de 96 poços, incubando-se a placa durante 12 horas a 4°C. Os poços foram então lavados por três vezes com tampão de lavagem PBT (Tween-20 a 0,1% em PBS). Adicionaram-se então 200 µl de tampão de saturação (Polivinilpirrolidona 10 a 2,7% em água) a cada poço e as placas foram incubadas durante 2 horas a 37°C. Os poços foram lavados por três vezes com PBT. Adicionaram-se 200 µl de soros diluídos (tampão de diluição: BSA a 1%, Tween-20 a 0,1%,  $\text{NaN}_3$  a 0,1% em PBS) a cada poço e as placas foram incubadas durante 2 horas a 37°C. Os poços foram lavados três vezes com PBT. Adicionaram-se a cada poço 100 µl de soro de coelho anti-ratinho conjugado com HRP (Dako) diluído a 1:2000 em tampão de diluição, e as placas foram incubadas durante 90 minutos a 37°C. Os poços foram lavados três vezes com tampão de PBT. Adicionaram-se 100 µl de tampão de substrato para HRP (25 ml de tampão de citrato pH 5, 10 mg of O-fenildiamina e 10 µl de  $\text{H}_2\text{O}_2$ ) a cada poço e as placas foram deixadas à temperatura ambiente durante 20 minutos. Adicionaram-se 100 µl de  $\text{H}_2\text{SO}_4$  a 12,5% a cada poço e mediu-se a  $\text{DO}_{490}$ . Os títulos de ELISA foram arbitrariamente calculados como correspondendo à diluição de soro que forneceu um valor de  $\text{DO}_{490}$  superior em 0,4 relativamente ao valor do soro pré-imune. O ensaio ELISA foi considerado positivo quando a diluição de soro com  $\text{DO}_{490}$  de 0,4 foi superior a 1:400.

#### Ensaio de Ligação e FACScan para Bactérias

A estirpe acapsulada M7 de MenB foi semeada em placas de agar de chocolate e incubada durante 12 horas a 37°C. As colónias de bactérias foram recolhidas das placas de agar utilizando uma zaragatoa de dracon esterilizada, sendo então inoculadas em 4 tubos, cada um contendo 8 ml de Caldo de Mueller-Hinton (Difco) contendo 0,25% de glucose. O crescimento das bactérias foi monitorizado a cada 30 minutos seguindo a  $\text{DO}_{620}$ . As bactérias foram deixadas em crescimento até que o valor de DO atingisse valores de 0,35-0,5. A cultura foi centrifugada



durante 10 minutos a 4000 rpm. O sobrenadante foi rejeitado e o pellet foi ressuspensão em tampão de bloqueio (BSA a 1% em PBS,  $\text{NaN}_3$  a 0,4%) e centrifugado durante 5 minutos a 4000 rpm. As células foram ressuspensas em tampão de bloqueio até se verificar uma  $\text{DO}_{620}$  de 0,07. Adicionaram-se 100  $\mu\text{l}$  de células bacterianas a cada poço de uma placa Costar de 96 poços. Adicionaram-se 100  $\mu\text{l}$  de soro diluído (1:100, 1:200, 1:400 em tampão de bloqueio) a cada poço e as placas foram incubadas durante 2 horas a 4°C. As células foram centrifugadas durante 5 minutos a 4000 rpm, tendo o sobrenadante sido aspirado; as células foram lavadas por adição de 200  $\mu\text{l}$  de tampão de bloqueio a cada poço. Adicionaram-se a cada poço 100  $\mu\text{l}$  de  $\text{F(ab)}_2$  de cabra anti-ratinho conjugado com R-Ficoeritrina, diluído a 1:100, e as placas foram incubadas durante 1 hora a 4°C. As células foram submetidas a centrifugação a 4000rpm durante 5 minutos e lavadas por adição de 200  $\mu\text{l}$  de tampão de bloqueio a cada poço. O sobrenadante foi aspirado e as células foram ressuspensas em 200  $\mu\text{l}$ /poço de PBS, 0,25% de formaldeído. As amostras foram transferidas para tubos de FACScan e lidas. As configurações do FACScan (potência de laser, 15 mW) foram : FL2 on; Limiar FSC-H: 92; Voltagem FSC PMT: E 01; SSC PMT: 474; Ganhos de Amp. 6.1; FL-2 PMT: 586; valores de compensação: 0.

#### Preparações de OMV

As bactérias foram deixadas em cultura durante um período de 12 horas em 5 placas de GC, sendo então recolhidas com uma ansa e ressuspensas em 10 ml de Tris-HCl 20mM. Efetuou-se uma inativação por calor a 56°C durante 30 minutos e as bactérias foram lisadas por sonicação em gelo durante 10 minutos (50% ciclo, 50% rendimento). As células que permaneceram inteiras foram removidas por centrifugação a 5000g durante 10 minutos e a fração total de membranas celulares foi recuperada por centrifugação a 50000g durante 75 minutos. Para extrair as

proteínas de membrana citoplasmática a partir do extracto bruto das membranas externas, a fração foi ressuspensa na totalidade em Sarkosyl (Sigma) a 2% e incubada à temperatura ambiente durante 20 minutos. A suspensão foi centrifugada a 10000g durante 10 minutos para remover agregados e o sobrenadante foi de novo centrifugado a 50000g durante 75 minutos para produzir pellets de membranas externas. As membranas externas foram ressuspensas em Tris-HCl 10mM, pH8 e a concentração das proteínas foi medida através de um ensaio Bio-Rad para proteínas utilizando BSA como padrão.

#### Preparação de Extratos Integrais

As bactérias foram cultivadas durante 12 horas numa placa de GC, sendo então recolhidas com uma ansa e ressuspensas em 1 ml de Tris- HCl 20mM. A inativação por calor foi realizada a 56°C durante 30 minutos.

#### Western Blot

As proteínas purificadas (500ng/aplicação), as vesículas de membranas externas (5 µg) e os extractos totais de células (25µg) derivados da estirpe 2996 de MenB foram carregados sobre um gel de SDS-poliacrilamida a 12% e transferidos para uma membrana de nitrocelulose. A transferência foi realizada durante 2 horas a 150mA a 4°C, num tampão de transferência (0,3 % de base Tris, 1,44 % glicina, metanol a 20% (v/v)). A membrana foi saturada por incubação durante 12 horas a 4°C em tampão de saturação (10% leite desnatado, Triton X100 a 0,1% em PBS). A membrana foi lavada por duas vezes com tampão de lavagem (3% leite desnatado, Triton X100 a 0,1% em PBS) e incubada por um período de duas horas a 37°C com soro de ratinho diluído a 1:200 em tampão de lavagem. A membrana foi lavada por duas vezes e incubada durante 90 minutos com uma diluição a de 1:2000 de Ig anti-ratinho marcada com peroxidase de rábano. A membrana foi lavada por duas vezes com Triton

X100 a 0,1% em PBS e foi revelada com o Opti-4CN Substrate Kit da Bio-Rad. A reação foi interrompida através da adição de água.

#### Ensaio Bactericida

As estirpes MC58 e 2996 foram cultivadas em placas de agar de chocolate durante 12 horas a 37°C. Foram recolhidas das placas de agar 5 a 7 colónias, as quais foram usadas para inocular 7 ml de Caldo de Mueller-Hinton. A suspensão foi incubada num agitador Nutator a 37°C e foi deixada em crescimento até que o valor de DO<sub>620</sub> atingisse valores de 0,5-0,8. A cultura foi dividida em alíquotas para tubos esterilizados Eppendorf de 1,5ml e centrifugada durante 20 minutos numa microcentrífuga a velocidade máxima. O pellet foi lavado uma vez em tampão de Gey (Gibco) e ressuspenso no mesmo tampão até se obter uma DO<sub>620</sub> de 0,5, sendo então diluído a 1:20000 em tampão de Gey e armazenado a 25°C.

Adicionaram-se 50µl de tampão de Gey com 1% BSA a cada um dos 96 poços de uma placa de cultura de tecidos. Adicionaram-se a cada poço 25µl de soro de ratinho diluído (1:100; tampão de diluição: tampão de Gey/BSA a 0,2%), tendo a placa sido incubada de seguida a 4°C. Adicionaram-se 25µl da suspensão de bactérias acima descrita a cada poço. Adicionaram-se a cada poço 25µl de complemento de coelho inativado pelo calor (banho de água a 56°C durante 30 minutos) ou de complemento normal de coelho (não inativado). Imediatamente após a adição do complemento de coelho, foram semeados 22µl de cada amostra/poço em placas de agar de Mueller-Hinton (tempo 0). A placa de 96 poços foi incubada durante 1 hora a 37°C com rotação; seguidamente 22µl de cada amostra/poço foram semeados em placas de agar Mueller-Hinton (tempo 1h). Depois de um período de incubação de 12 horas, as colónias correspondentes ao tempo 0 e ao tempo 1h foram contadas.

### EXEMPLO 1

Utilizando os procedimentos acima descritos, foram usados os iniciadores oligonucleotídicos que se seguem no ensaio de reação em cadeia por polimerase (PCR), de modo a clonar os ORFs segundo indicado:

**Tabela 1: Oligonucleótidos usados na PCR para o Exemplo 2**

	Reverse	CCCGCTCGAG-TTTAGAACCGCATTTGCC	XhoI
953	Forward	CGCGGATCCCATATG-GCCACCTACAAAGTGGAC	BamHI-NdeI

A seguinte sequência parcial de ADN foi identificada em *N.gonorrhoeae* <SEQ ID 2915>:

#### g953.seq

```
1 ATGAAAAAA TCATCTTCGC CGCGCTCGCA GCGGCAGCCG TCGGCACTGC
51 CTCCGCCACC TACAAAGTGG ACGAATATCA CGCCAACGTC CGTTTCGCCA
101 TCGACCACTT CAACACCAGC ACCAACGTCG GCGGTTTTTA CGGTCTGACC
151 GGTTCGTCG AGTTCGATCA AGCAAAACGC GACGGCAAAA TCGACATCAC
201 CATTCCCGTC GCCAACCTGC AAAGCGGTTT GCAACCCTTC ACCGGCCACC
251 TGAAATCCGC CGACATCTTC GATGCCGCTC AATATCCGGA CATCCGCTTC
301 GTTTCACCA AATTCAACTT CAACGGCAAA AAAGTTGTTT CCGTTGACGG
351 CAACCTGACC ATGCGCGGCA AAACCGCCCC CGTCAAATC AAAGCCGAAA
401 AATTCAACTG CTACCAAAGC CCGATGGCGG AAACCGAAGT TTGCGGCGGC
451 GACTTCAGCA CCACCATCGA CCGCACCAA TGGGGCGTGG ACTACCTCGT
501 TAACGCCGGT ATGACCAAAA ACGTCCGCAT CGACATCAA ATCGAAGCTG
551 CAAAACAATA A
```

Isto corresponde à sequência de aminoácidos <SEQ ID 2916; ORF 953.ng>:

#### g953.pep

```
1 MKKIIFAALA AAAVGTASAT YKVDEYHANV RFAIDHENTS TNVGGFYGLT
51 GSVEFDQAKR DGKIDITIPV ANLQSGSQPF TGHLSADIF DAAQYPDIRF
101 VSTKFNFNKG KLVSDGNLT MRGKTAPVKL KAEKFNCYQS PMAETEVCGG
151 DFSTTIDRTK WGVVDYLVNAG MTKNVRIDIQ IEAAKQ*
```

A seguinte sequência parcial de ADN foi identificada em *N.meningitidis* <SEQ ID 2917>:

# m953.seq

```

1  ATGAAAAAAA TCATCTTCGC CGCACTCGCA GCCGCCGCCA TCAGTACTGC
51 CTCCGCCGCC ACCTACAAAG TGGACGAATA TCACGCCAAC GCCCGTTTCG
101 CCATCGACCA TTTCAACACC AGCACCAACG TCGGCGGTTT TTACGGTCTG
151 ACCGGTTCCG TCGAGTTCGA CCAAGCAAAA CGCGACGGTA AAATCGACAT
201 CACCATCCCC ATTGCCAACC TGCAAAGCGG TTCGCAACAC TTTACCGACC
251 ACCTGAAATC AGCCGACATC TTCGATGCCG CCCAATATCC GGACATCCGC
301 TTTGTTTCCA CCAAATTCAA CTTCAACGGC AAAAACTGG TTTCCGTTGA
351 CGGCAACCTG ACCATGCACG GCAAAACCGC CCCCGTCAAA CTCAAAGCCG
401 AAAAATTCAA CTGCTACCAA AGCCCGATGG AGAAAACCGA AGTTTGTGGC
451 GCGGACTTCA GCACCACCAT CGACCGCACC AAATGGGGCA TGGACTACCT
501 CGTTAACGTT GGTATGACCA AAAGCGTCCG CATCGACATC CAAATCGAGG
551 CAGCCAAACA ATAA

```

Isto corresponde à sequência de aminoácidos <SEQ ID 2918; ORF 953>:

## m953.pep

```

1  MKKIIFAALA AAAISTASAA TYKVDEYHAN ARFAIDHFNT STNVGGFYGL
51  TGSVEFDQAK RDGKIDITIP IANLQSGSQH FTDHLKSADI FDAAQYPDIR
101 FVSTKFNENG KKLVSVDGNL TMHGKTAPVK LKAKEFNICY SPMEKTEVCG
151 GDFSTTIDRT KWGMDYLVNV GMTKSVRIDI QIEAAKQ*

```

A análise por computador desta sequência de aminoácidos forneceu os seguintes resultados: Homologia com um ORF previsto de *Neisseria gonorrhoeae* ORF 953 mostra 93% de identidade sobre 187 aa sobreposto com o ORF previsto (ORF 953) de *N. gonorrhoeae*.

m953/g953 93.0% identity in 187 aa overlap

	10	20	30	40	50	60
m953.pep	MKKIIFAALAAAIASTASAATYKVDEYHANARFAIDHFNTSTNVGGFYGLTGSVEFDQAK					
g953	MKKIIFAALAAAVGTASA-TYKVDEYHANVRFAIDHFNTSTNVGGFYGLTGSVEFDQAK					
	10	20	30	40	50	
	70	80	90	100	110	120
m953.pep	RDGKIDITIPIANLQSGSQHFTDHLKSADIFDAAQYPDIRFVSTKFNENGKKLVSVDGNL					
g953	RDGKIDITIPVANLQSGSQPFTGHLKSADIFDAAQYPDIRFVSTKFNENGKKLVSVDGNL					
	60	70	80	90	100	110
	130	140	150	160	170	180
m953.pep	TMHGKTAPVKLKAKEFNICYSPMEKTEVCGGDFSTTIDRTKWGMDYLVNVGMTKSVRIDI					
g953	TMRGKTAPVKLKAKEFNICYSPMAETEVCGGDFSTTIDRTKWGMDYLVNAGMTKNVRIDI					
	120	130	140	150	160	170
m953.pep	QIEAAKQX					
g953	QIEAAKQX					
	180					

A seguinte sequência parcial de ADN foi identificada em *N.meningitidis* <SEQ ID 2919>:

**a953.seq**

```
1 ATGAAAAAAAA TCATCATCGC CGCGCTCGCA GCAGCCGCCA TCGGCACTGC
51 CTCCGCCGCC ACCTACAAAG TGGACGAATA TCACGCCAAC GCCCGTTTCT
101 CTATCGACCA TTTCAACACC AGCACCAACG TCGGCGGTTT TTACGGTCTG
151 ACCGGTTCCG TTGAGTTCGA CCAAGCAAAA CGCGACGGTA AAATCGACAT
201 CACCATCCCC GTTGCCAACC TGCAAAGCGG TTCGCAACAC TTTACCGACC
251 ACCTGAAATC AGCCGACATC TTCGATGCCG CCCAATATCC GGACATCCGC
301 TTTGTTTCCA CCAAATTCAA CTTCAACGGC AAAAAACTGG TTTCCGTTGA

351 CGGCAACCTG ACCATGCACG GCAAACCGC CCCCGTCAA CTCAAAGCCG
401 AAAAATTCAA CTGCTACCAA AGCCCGATGT TGAAAACCGA AGTTTGCCGC
451 GCGGACTTCA GCACCACCAT CGACCGCACC AAATGGGGCA TGGACTACCT
501 CGTTAACGTT GGTATGACCA AAAGCGTCCG CATCGACATC CAAATCGAGG
551 CAGCCAAACA ATAA
```

Isto corresponde à sequência de aminoácidos <SEQ ID 2920; ORF 953.a>:

**a953.pep**

```
1 MKKIIIAALA AAAIGTASAA TYKVDEYHAN ARFSIDHFNT STNVGGFYGL
51 TGSVEFDQAK RDGKIDITIP VANLQSGSQH FTDHLKSADI FDAAQYPDIR
101 FVSTKFNENG KKLVSVDGNL TMHGKTAPVK LKAKEFNICYQ SPMLKTEVCG
151 GDFSTTIDRT KWGMDYLVNV GMTKSVRID I QIEAAKQ*
```

A análise por computador desta sequência de aminoácidos forneceu os seguintes resultados: Homologia com um ORF previsto de *N.meningitidis* ORF 953 mostra 97,3% de identidade sobre 187 aa sobreposto com o ORF previsto (ORF 953) de *N.meningitidis*.

a953/m953 97.3% identity in 187 aa overlap

	10	20	30	40	50	60
a953.pep	MKKIIIAALAAAAIGTASAATYKVDEYHANARFSIDHFNTSTNVGGFYGLTGSVEFDQAK					
	:     :     :     :     :     :     :     :     :					
m953	MKKIIFAALAAAAISTASAATYKVDEYHANARFAIDHFNTSTNVGGFYGLTGSVEFDQAK					
	10	20	30	40	50	60
	70	80	90	100	110	120
a953.pep	RDGKIDITIPVANLQSGSQHFTDHLKSADIFDAAQYPDIRFVSTKFNFNKGKLVSDGNL					
	:     :     :     :     :     :     :     :					
m953	RDGKIDITIPIANLQSGSQHFTDHLKSADIFDAAQYPDIRFVSTKFNFNKGKLVSDGNL					
	70	80	90	100	110	120
	130	140	150	160	170	180
a953.pep	TMHGKTAPVKLKA EKFN CYQSPMLKTEVCGGDFSTTIDRTKWGMDYLVNVGMTKSVRID I					
	:     :     :     :     :     :     :     :					
m953	TMHGKTAPVKLKA EKFN CYQSPMEKTEVCGGDFSTTIDRTKWGMDYLVNVGMTKSVRID I					
	130	140	150	160	170	180
a953.pep	QIEAAKQX					
m953	QIEAAKQX					

Os exemplos mencionados pretendem ilustrar e não limitar a invenção.

## **REIVINDICAÇÕES**

1. Composição compreendendo uma proteína isolada compreendendo:  
(i) uma sequência de aminoácidos selecionada a partir do grupo que consiste nas SEQ IDs 2918, 2916 e 2920; ou (ii) uma sequência de aminoácidos que tem 90% ou mais de identidade de sequência com uma sequência de aminoácidos selecionada a partir do grupo que consiste nas SEQ IDs 2918, 2916 e 2920.
2. Composição da reivindicação 1, incluindo um adjuvante.
3. Composição da reivindicação 2, em que o adjuvante é sais de alumínio ou MF59.
4. Anticorpo que se liga especificamente a uma sequência de aminoácidos selecionada a partir do grupo que consiste nas SEQ IDs 2918, 2916 e 2920.
5. Molécula de ácido nucleico que codifica uma proteína, tal como definido na reivindicação 1.
6. Molécula de ácido nucleico de acordo com a reivindicação 5, que compreende uma sequência de nucleótidos selecionada a partir do grupo que consiste nas SEQ IDs 2917, 2915 e 2919.
7. Molécula de ácido nucleico compreendendo uma sequência de nucleótidos complementar a uma molécula de ácido nucleico de acordo com a reivindicação 5 ou reivindicação 6.
8. Vetor compreendendo a molécula de ácido nucleico de qualquer uma das reivindicações 5 a 7.
9. Vetor de acordo com a reivindicação 8, que é um vetor de expressão.



10. Célula hospedeira transformada com o vetor de acordo com a reivindicação 8 ou a reivindicação 9.
11. Processo para a produção de uma proteína, tal como definido na reivindicação 1, compreendendo o passo de cultivar uma célula hospedeira de acordo com a reivindicação 10 sob condições que induzem a expressão da proteína.
12. Utilização de uma proteína como definido na reivindicação 1 no fabrico de um medicamento para o tratamento ou prevenção de infeção causada por bactérias *Neisseria*.