

C12N 15/31 (2007.10) **C07K 14/22** (2007.10)
C07K 16/12 (2007.10) **C12Q 1/68** (2007.10)
A61K 39/95 (2007.10) **G01N 33/50** (2007.10)

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **1999.04.30**

(30) Prioridade(s): **1998.05.01 US 83758 P**
1998.07.31 US 94869 P
1998.09.02 US 98994 P
1998.09.02 US 99062 P
1998.10.09 US 103749 P

(43) Data de publicação do pedido: **2006.04.12**

(45) Data e BPI da concessão: **2007.10.24**
024/2008

(73) Titular(es):

NOVARTIS VACCINES AND DIAGNOSTICS, INC.
4560 HORTON STREET EMERYVILLE, CA 94608

US

J. CRAIG VENTER INSTITUTE, INC.

US

(72) Inventor(es):

CLAIRE FRASER

US

CESIRA GALEOTTI

IT

GUIDO GRANDI

IT

ERIN HICKEY

US

VEGA MASIGNANI

IT

(74) Mandatário:

FRANCISCO NOVAES CUNHA BRITO SOTO MAIOR DE
ATAYDE

AV. DUQUE D'AVILA, N.º 32, 1º ANDAR 1000-141 LISBOA PT

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

(57) Resumo:

ANITIGÉNIOS DA NEISSERIA MENINGITIDIS E COMPOSIÇÕES.

DESCRIÇÃO

EPÍGRAFE:	"ANTIGÉNIOS DA <i>NEISSERIA MENINGITIDIS</i> E COMPOSIÇÕES"
-----------	--

Campo do Invento

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

Antecedentes do invento

O patogénio humano *Neisseria meningitidis* é um diplococo gram negativo imóvel. Coloniza a faringe causando meningite e, ocasionalmente, septicémia na ausência de meningite. Este patogénio 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 patogénicos.

A *N.meningitidis* causa tanto doença endémica quanto epidémica. Nos E.U.A. 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 endémica são muito superiores e durante períodos de incidência epidémica as taxas de incidência podem alcançar 500 casos por 100,000 pessoas/ano. A mortalidade é extremamente elevada, atingindo 10-20% nos E.U.A. 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 E.U.A. (Schuchat et al (1997) supra).

Foram já identificados 12 serogrupos de *N.meningitidis* com base no polisacárido capsular do organismo. O grupo A é o patogénio mais frequentemente implicado na doença epidémica na África Sub-Sahariana. Os serogrupos B e C são responsáveis pela grande maioria dos casos nos E.U.A. e na maioria dos países desenvolvidos. Os serogrupos W135 e Y são responsáveis pelos casos restantes nos E.U.A. 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 protecção 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 polisacáridos serem antigénios 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 E.U.A., Europa e América do Sul. A abordagem polisacarídica não pode ser utilizada porque o polisacárido capsular do menB é um

polímero de ácido N-acetil-neuramínico ligado em (2-8) que também está presente no tecido dos mamíferos. Isto resulta em tolerância ao antigénio; na verdade, se uma resposta imunitária fosse desencadeada, seria contra o próprio organismo e conseqüentemente indesejável. De modo a evitar a indução de autoimunidade e a induzir uma resposta imunitária protectora, o polisacárido capsular foi, por exemplo, quimicamente modificado, tendo os grupos de N-acetil sido substituídos por grupos de N-propionil, deixando a antigenicidade específica inalterada (Romero & Ootschoom (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 actividade 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 (p.ex. 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 (p.ex. EP-A-0467714, WO96/29412), mas a sequenciação não está de modo algum completa. Um mapa físico do cromossoma de uma estirpe do serogrupo A de *N.meningitidis* é descrito em Dempsey et al. (1995), J Bacteriol 177:6390-400, e em Moxon (1997), Lancet 350: 1240-44, é reportado que as sequências genómicas de diferentes patogénios, incluindo *N.meningitidis*, estão a ser determinadas. O fornecimento de mais sequências pode traduzir-se numa oportunidade de identificar proteínas secretadas 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.

É ainda objecto do invento fornecer sequências de DNA de *Neisseria* que codificam para proteínas que são antigénicas ou imunogénicas.

Descrição do invento

O invento fornece proteínas compreendendo as sequências de aminoácidos de *N.meningitidis* e as sequências de aminoácidos de *N.gonorrhoeae* que se encontram expostas nos Exemplos.

O invento fornece ainda proteínas compreendendo sequências homólogas (i.e., que apresentam identidade de sequência)

relativamente às sequências de aminoácidos de *N.meningitidis* expostas nos Exemplos e que são eficazes na prevenção da doença causada por *Neisseria*. Dependendo da sequência em particular, o grau de homologia (identidade de sequência) é preferencialmente superior a 50% (p.ex., 60%, 70%, 80%, 90%, 95%, 99% ou mais). Estas proteínas incluem mutantes e variantes alélicas das sequências descritas nos Exemplos. 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*.

O invento proporciona ainda proteínas compreendendo fragmentos das sequências de aminoácidos de *N.meningitidis* e das sequências de aminoácidos de *N. gonorrhoeae* que se encontram expostas nos Exemplos. Os fragmentos deverão compreender pelo menos *n* aminoácidos consecutivos destas sequências, em que, dependendo da sequência particular, *n* corresponde a 8 ou mais aminoácidos (p.ex., 10, 12, 14, 16, 18, 20 ou mais). Os fragmentos compreendem um epítipo da sequência.

As proteínas do invento poderão, como é evidente, ser preparadas por diversos métodos (p.ex., expressão recombinante, purificação a partir de cultura celular, síntese química, etc.) e sob variadas formas (p.ex., 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*).

Ainda, o invento fornece anticorpos que se ligam a estas sequências de aminoácidos. Estes anticorpos poderão ser

policlonais ou monoclonais e a sua produção poderá ser efectuada através de qualquer meio adequado.

Num seu aspecto adicional, o invento fornece um ácido nucleico compreendendo as sequências nucleotídicas de *N.meningitidis* e as sequências nucleotídicas de *N. gonorrhoeae* que se apresentam nos Exemplos.

Segundo um outro aspecto, são fornecidos ácidos nucleicos com identidade de sequência superior a 50% (p.ex., 60%, 70%, 80%, 90%, 95%, 99% ou mais) em relação às sequências de ácido nucleico aqui expostas. A identidade de sequência determina-se de acordo com o já descrito.

São igualmente fornecidos ácidos nucleicos compreendendo fragmentos destas sequências. Estes fragmentos deverão compreender pelo menos n nucleótidos consecutivos das sequências de *N.meningitidis* ou das sequências de *N.gonorrhoeae* e, dependendo da sequência particular considerada, n corresponderá a 12 ou mais (p.ex., 20, 25, 30, 35, 40 ou mais). De acordo com um outro aspecto do invento, é fornecido um ácido nucleico codificando para as proteínas e fragmentos de proteínas do invento.

O invento fornece ainda um ácido nucleico compreendendo sequências complementares às acima descritas (p.ex. para uso como sequências anti-senso ou de sonda).

O ácido nucleico de acordo com o invento poderá, evidentemente, ser preparado de diversas formas (p.ex. a por síntese química, de parte ou do seu todo, a partir de bibliotecas genómicas ou de cDNA, a partir do próprio organismo, etc.) e poderá encontrar-se sob diversas formas (p.ex. de cadeia simples, cadeia dupla, vectores, sondas, etc.).

Adicionalmente, o termo "ácido nucleico" inclui DNA e RNA e ainda os seus análogos, tais como aqueles que contêm

estruturas primárias modificadas, além de ácidos nucleicos peptídicos (PNAs), etc..

De acordo com um aspecto adicional, o invento fornece vectores que compreendem as sequências nucleotídicas do invento (p.ex. vectores de expressão) e células hospedeiras transformadas com tais vectores.

Segundo um outro aspecto do invento, 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.

O invento fornece igualmente um ácido nucleico, proteína ou anticorpo de acordo com o invento para uso como medicamentos (p. ex., como vacinas) ou como reagentes para diagnóstico. O invento proporciona ainda o uso do ácido nucleico, proteína ou anticorpo de acordo com o invento para o fabrico de: (i) um medicamento para prevenção de uma infecção devida a *Neisseria*; (ii) um reagente para diagnóstico, destinado a detectar a presença de *Neisseria* ou de anticorpos dirigidos contra *Neisseria* ou (iii) 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 do serogrupo B ou serogrupo C.

É fornecido um método para o tratamento de um paciente, podendo envolver a administração a um paciente de uma quantidade terapêuticamente eficaz de um ácido nucleico, proteína e/ou anticorpo de acordo com o invento.

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

Segue-se um resumo das técnicas e procedimentos padronizados que poderão ser utilizados para colocar em prática o invento (p.ex., para utilizar as sequências expostas com fins de vacinação ou diagnóstico), que está anexado como Apêndice a este pedido de patente. Este Resumo fornece exemplos que podem ser utilizados.

Tendo acabado de descrever o invento de forma geral, o mesmo será melhor compreendido tomando como referência os exemplos que se seguem, os quais são fornecidos com fins ilustrativos.

Metodologia - Resumo dos procedimentos e técnicas padronizados

Generalidades

Este invento fornece sequências nucleotídicas de *Neisseria meningitidis* menB e as sequências de aminoácidos por estas codificadas. A partir destas sequências poderão produzir-se ensaios com sondas de ácido nucleico e cassetes e vectores de expressão. Os vectores 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 extractos 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 do presente invento utilizará, a não ser que indicado de outro modo, técnicas convencionais de biologia molecular, microbiologia, DNA recombinante e

imunologia, as quais são bem conhecidas neste campo técnico. Tais técnicas encontram-se descritas em detalhe na literatura [p.ex., Sambrook, *Molecular Cloning: A Laboratory Manual*, 2ª Ed. (1989); *DNA 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*, 2ª 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 patente aqui citados são incorporados na sua totalidade para fins de referência.

Sistemas de expressão

As sequências nucleotídicas de *Neisseria menB* poderão ser expressas em uma 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 DNA que tem a capacidade de se ligar a uma RNA polimerase de mamífero e iniciar a transcrição a jusante (3') de uma sequência codificante (p.ex. 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 RNA polimerase II para iniciar a síntese de RNA no local correcto. Um promotor de mamífero conterà 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 actuar em ambas as orientações [Sambrook et al. (1989) "Expression of Cloned Genes in Mammalian Cells", em Molecular Cloning: a Laboratory Manual, 2^a ed.].

Os genes de vírus que infectam mamíferos são frequentemente expressos em frequência elevada e possuem uma larga variedade de hospedeiros; assim, as sequências que codificam genes de vírus que infectam mamíferos proporcionam sequências promotoras particularmente úteis. Como exemplos destes promotores encontram-se o promotor precoce de SV40, o promotor LTR viral do tumor mamário de ratinho, o promotor principal tardio de adenovírus (Ad MLP) e o promotor do vírus herpes simplex. dicionalmente, 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 seleccionado, muitos promotores podem ser indutíveis utilizando substratos conhecidos, tal como sucede usando o promotor do vírus de 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 nº 5,783,681).

A presença de um elemento enhancer, combinado com os elementos promotores acima descritos, aumentará geralmente os níveis de expressão. Um enhancer é uma sequência reguladora de DNA 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 RNA. Os enhancers são igualmente activos quando colocados a montante ou a jusante do local de iniciação da transcrição, na orientação normal ou invertida, ou a uma distância de mais do que 1000 nucleótidos de afastamento do promotor [Maniatis et al. (1987) Science 236:1237; Alberts et al. (1989) Molecular Biology of the Cell, 2^a ed.]. Os elementos enhancer derivados de vírus poderão ser particularmente úteis, uma vez que estes possuem geralmente uma gama mais alargada de hospedeiros. Os exemplos incluem o enhancer do gene precoce de SV40 [Dijkema et al. (1985) EMBO J. 4:761] e o enhancer/promotores derivados da longa sequência repetitiva terminal (LTR) do vírus de sarcoma Rous [Gorman et al. (1982b) Proc. Natl. Acad. Sci. 79:6777] e do citomegalovírus humano [Boshart et al. (1985) Cell 41:521]. Adicionalmente, alguns enhancers são reguláveis e tornam-se activos 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 DNA poderá ser expressa intracelularmente em células de mamíferos. Uma sequência promotora poderá encontrar-se directamente ligada à molécula de DNA, 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

separada da proteína por incubação *in vitro* com brometo de cianogénio.

Alternativamente, as proteínas estranhas poderão também ser secretadas pela célula para o meio de cultura através da criação de moléculas quiméricas de DNA que codificam 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 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, 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 DNA. 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 enhancers, introns 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 num replicon, tal como um elemento extracromossómico (p.ex., 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 infectam animais, os quais requerem factores 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 replicons de mamífero encontram-se os derivados do papilomavírus bovino e do vírus de Epstein-Barr. Adicionalmente, o replicon 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 vectores de vai-vem 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

liposomas lipossomas e microinjecção directa do DNA 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 (p.ex., 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 US 5.693.506, 5.659.122 e 5.608.143. 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); Wirsel 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 giberelínico, e da secreção de enzimas induzida pelo ácido giberelínico, 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 vector 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 vector as características necessárias para que este transfira o DNA de um hospedeiro de clonagem original, como uma bactéria, para uma planta hospedeira desejada. A construção básica do vector de bactéria/planta proporcionará preferencialmente uma origem de replicação procariótica para uma larga gama de hospedeiros; um marcador procariótico de selecção; e, para as transformações com *Agrobacterium*, sequências de DNA 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 selecção apropriado para determinar se uma célula de planta foi transformada. Poderá encontrar-se uma exposição geral sobre marcadores de selecção, por exemplo para os membros da família da grama, em Wilkink 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 cassette heteróloga de expressão num genoma de planta. Os marcadores procarióticos de selecção apropriados incluem a resistência a antibióticos como a ampicilina ou a tetraciclina. Poderão ainda encontrar-se presentes no vector outras sequências de DNA que codifiquem para funções adicionais, tal como previsto neste campo técnico.

As moléculas de ácido nucleico do presente invento poderão ser incluídas numa cassette de expressão de modo a possibilitar a expressão da(s) proteína(s) de interesse. Existirá habitualmente apenas uma cassette de expressão, se bem que a existência de duas ou mais cassettes 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 únicos de restrição enzimática a nível das extremidades 5' e 3' da cassette permitem a fácil inserção num vector pré-existente.

Uma sequência codificante heteróloga poderá corresponder a qualquer das proteínas abrangidas pelo presente invento. 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 do invento 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 facilmente recolhidas. Tipicamente, as sementes secretam 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 introns pelo equipamento de splicing do hospedeiro. Em caso afirmativo, poderá efectuar-se a mutagénesis dirigida à região de "intron" para evitar que se perca uma porção da mensagem genética sob a forma de falso código de intron. Reed e Maniatis, *Cell* 41:95-105, 1985.

O vector poderá ser microinjectado directamente nas células da planta, utilizando micropipetas para transferir mecanicamente o DNA 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 meio de 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 mini-células, células, lipossomas ou outros corpos de superfície lipídica

fundíveis; Fraley et al., Proc. Natl. Acad. Sci. USA, 79, 1859-1863, 1982.

O vector 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 eléctricos 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 invento 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-do-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 citocininas. 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 genotipo 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 do invento desejada poderá ser excretada ou, alternativamente, a proteína poderá ser extraída a partir da planta inteira. Quando a proteína do invento desejada é secretada 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 ruptura mecânica para permitir a libertação de qualquer proteína secretada 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 Baculovirus

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

Após a inserção da sequência de DNA que codifica para a proteína no vector de transferência, o vector e o genoma viral de tipo selvagem são transfectados numa célula hospedeira de insecto, na qual se dá a recombinação do vector 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 baculovirus/células de insecto 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 DNA codificando para a proteína no genoma do baculovirus, 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 (vector 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 mantidos num replicon, tal como um elemento extracromossómico (p.ex., plasmídeos) com a capacidade de se manter de modo estável num hospedeiro, tal como uma bactéria. O replicon possuirá um sistema de replicação, deste modo permitindo a sua manutenção num hospedeiro adequado para clonagem e amplificação.

Actualmente, o vector de transferência mais vulgarmente usado para a introdução de genes estranhos no AcNPV é o pAc373. Têm igualmente sido concebidos muitos outros vectores 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) e um gene procariótico de resistência à ampicilina (amp) e origem de replicação para selecção e propagação em *E. coli*.

Os vectores de transferência de baculovírus contêm geralmente um promotor de baculovírus. Um promotor de baculovírus corresponde a qualquer sequência de DNA com a capacidade de se ligar a uma RNA polimerase de baculovírus e

de iniciar a transcrição a jusante (5' para 3') de uma sequência codificante (p.ex., gene estrutural) para mRNA. 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 RNA polimerase e um local de iniciação da transcrição. Um vector de transferência de baculovírus poderá igualmente conter um segundo domínio designado por enhancer, 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 infecção 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 DNA codificando para sequências de sinal apropriadas poderá derivar de genes para proteínas secretadas por células de insecto 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 insecto, 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 insectos, tais como os que derivam de genes

codificando para o α -interferão humano (Maeda et al. (1985), Nature 315:592), o 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) DNA 7:99), poderão igualmente ser usados para possibilitar a secreção em insectos.

Um polipéptido ou poliproteína recombinante poderá ser expresso intracelularmente ou, se expresso com as sequências reguladoras apropriadas, poderá ser secretado. 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 secretadas poderão ser secretadas pelas células de insecto por meio da criação de moléculas quiméricas de DNA codificando para uma proteína de fusão composta por um fragmento de sequência leader que proporciona a secreção da proteína estranha em insectos. 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 DNA e/ou do gene codificando para o produto de expressão precursor da proteína, uma célula hospedeira de insecto é co-transformada com o DNA heterólogo do vector de transferência e com o DNA 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 DNA 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 efectuar-se num gene tal como o gene da polihedrina, por recombinação homóloga com crossover duplo; a inserção poderá igualmente dar-se num local de restrição de um enzima construído no gene de baculovírus desejado. Miller et al. (1989) *Bioessays* 4:91. A sequência de DNA, quando clonada em lugar do gene da polihedrina no vector de expressão, é flanqueada em 5' e em 3' por sequências específicas de polihedrina e está posicionada a jusante do promotor da polihedrina.

O vector 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 inspecção 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 infectadas num período tardio após a infecção 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 μm , são muito refringentes, o que lhes dá uma aparência bastante brilhante que é facilmente visualizada ao microscópio. As células infectadas 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 insecto recorrendo a técnicas conhecidas dos peritos neste campo. Nomeadamente, as placas são inspeccionadas ao microscópio óptico 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 vectores de expressão recombinantes de baculovírus para a infecção de diversas células de insecto. 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* (WO 89/046699; (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 directa, 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 insecto 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 insecto 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 aumentando os nutrientes consumidos. O produto poderá ser purificado por técnicas como a cromatografia, p.ex., HPLC, cromatografia de afinidade, cromatografia de troca iónica, etc.; electroforese; centrifugação em gradiente de densidade; extracção por solvente, ou semelhantes. Quando apropriado, o produto poderá ser purificado de modo mais completo para remover substancialmente todas as proteínas de insecto que sejam igualmente secretadas para o meio, ou que resultem da lise das células de insecto, de forma a obter um produto que se encontra pelo menos substancialmente isento de resíduos do hospedeiro, p.ex., proteínas, lípidos e polisacáridos.

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 seleccionada. 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 DNA que tem a capacidade de se ligar a uma RNA polimerase bacteriana e iniciar a transcrição a jusante (3') de uma sequência codificante (p.ex. 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 RNA 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 RNA polimerase a nível do qual tem início a síntese de RNA. 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 activadora 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 activadora 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 dos EUA N° 4.738.921; 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. Grosser)] e os sistemas promotores do bacteriófago lambda PL [Shimatake et al. (1981) Nature 292:128] e do T5 [Patente dos EUA N° 4.689.406] 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 operão *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 RNA polimerase bacteriana e iniciar a transcrição. Um promotor de ocorrência natural e de origem não bacteriana poderá igualmente ser acoplado a uma RNA polimerase compatível para produzir elevados níveis de expressão de alguns genes em procariotas. O sistema de RNA 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)]. Acerca da expressão de genes eucarióticos e genes procarióticos com um fraco local de ligação ao ribossoma, ver Sambrook et al. (1989) "Expression of cloned genes in *Escherichia coli*", em "Molecular Cloning: a Laboratory Manual".

Uma molécula de DNA pode ser expressa intracelularmente. Poderá existir uma sequência promotora directamente ligada à molécula de DNA, 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 directa. Geralmente, uma sequência de DNA 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 o corte da proteína de bacteriófago a partir da proteína estranha [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 DNA que se encontra na junção das duas sequências de aminoácidos poderá ou não codificar para um local susceptível 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 (p.ex., protease de processamento ubiquitina-específica) para cortar a ubiquitina a partir 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 secretadas pela célula através da criação de moléculas quiméricas de DNA 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 secretada 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 DNA codificando para sequências de sinal adequadas poderá derivar de genes para proteínas bacterianas secretadas, tal como o gene da proteína da membrana externa de *E. coli* (*ompA*) [Masui et al. (1983), em *Experimental Manipulation of*

Gene Expression; Ghrayeb 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 secretar 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 DNA. As sequências de terminação da transcrição incluem frequentemente sequências de DNA 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 num replicon, 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 replicon possui um sistema de replicação, o que possibilita a sua manutenção num hospedeiro procariótico para expressão ou para clonagem e amplificação. Adicionalmente, um replicon 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 seleccionar-se um vector com um elevado ou com um baixo número de cópias, dependendo do efeito que o vector e a proteína estranha exercem sobre o hospedeiro.

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

Geralmente, as construções de expressão extracromossômicas e integrativas conterão marcadores de selecção que permitam a selecção das estirpes bacterianas que foram transformadas. Os marcadores selectivos podem expressar-se no hospedeiro bacteriano e podem incluir genes que tornam as bactérias resistentes a ampicilina, cloranfenicol, eritromicina, canamicina (neomicina) e tetraciclina (Davies et al. (1978) *Annu. Rev. Microbiol.* 32:469).

Os marcadores de selecção poderão ser expressos pelo hospedeiro bacteriano e poderão incluir genes que tornem a bactéria resistente a fármacos como a ampicilina, o cloranfenicol, a eritromicina, a kanamicina (neomicina) e a

tetraciclina [Davies et al. (1978) Annu. Rev. Microbiol. 32:469]. Os marcadores de selecção 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 vectores de transformação. Os vectores de transformação incluem geralmente um marcador de selecção que poderá ser mantido num replicon ou desenvolvido para um vector integrativo, tal como acima descrito.

Têm sido desenvolvidos diversos vectores de expressão e de transformação, tanto sob a forma de replicons extracromossómicos como de vectores integrativos, para a transformação de um grande número de bactérias. Por exemplo, foram desenvolvidos vectores 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 dos EUA N° 4.745.056].

Os métodos para a introdução de DNA exógeno em hospedeiros bacterianos são bem conhecidos neste campo técnico e incluem geralmente a transformação de bactérias tratadas com CaCl₂ ou com outros agentes, tal como catiões divalentes e DMSO. O DNA poderá também ser introduzido nas células bacterianas por electroporação. Os procedimentos de transformação variam geralmente com a espécie bacteriana a transformar. Ver, p.ex., 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. 4th Evr. Congo 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 DNA que possua a capacidade de se ligar à RNA polimerase de levedura e de iniciar a transcrição a jusante (3') de uma sequência codificante (p.ex. um gene estrutural) para mRNA. 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 RNA polimerase (a "TATA box") e um local de iniciação da transcrição. O promotor de levedura poderá igualmente conter um segundo domínio, designado por sequência activadora 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 activa, 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 activação 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 activação 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 activação da transcrição de um gene de uma enzima

glicolítica como a GAP ou a PyK (Publ. EPO n° 164 556). Adicionalmente, um promotor de levedura poderá incluir promotores de ocorrência natural não derivados de leveduras que possuam a capacidade de se ligar à RNA 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); Mercer-Au-Puigalon et al. (1980) Gene 11:163; Panthier et al. (1980) Curr. Genet. 2:109].

Uma molécula de DNA poderá ser expressa intracelularmente em leveduras. Poderá existir uma sequência promotora directamente ligada à molécula de DNA, 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 DNA 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 DNA que se

encontra na junção das duas sequências de aminoácidos poderá ou não conter um local susceptível de corte. Ver, p.ex., 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 (p.ex., protease de processamento ubiquitina-específica) para cortar a ubiquitina a partir da proteína estranha. Usando este método, é possível isolar a proteína estranha nativa (ver, p.ex., WO 88/024066).

Alternativamente, as proteínas estranhas poderão igualmente ser secretadas pela célula para o meio de cultura, recorrendo à criação de moléculas quiméricas de DNA 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 DNA codificando para sequências de sinal adequadas poderá derivar de genes de proteínas secretadas 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 factor 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 factor alfa de levedura, o qual contém tanto uma sequência de sinal "pré" como uma região "pró". Os tipos de fragmentos de factor alfa que podem ser utilizados incluem o leader de factor alfa pré-pró de

extensão completa (com cerca de 83 resíduos de aminoácidos) assim como leaders de factor 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 n° 324 274). Outros leaders que utilizam um fragmento do leader de factor alfa para proporcionar a secreção incluem leaders híbridos de factor alfa construídos com uma pré-sequência de uma primeira levedura, mas uma pró-região proveniente de um segundo factor alfa de levedura (ver, p.ex., Publ. PCT N° 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 DNA. 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 geralmente reunidos em construções de expressão. As construções de expressão são frequentemente mantidas num replicon, tal como um elemento extracromossómico (p.ex., plasmídeos) com a capacidade de se manter de modo estável num hospedeiro, tal como uma levedura ou bactéria. O replicon 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 vectores de vai-vem de levedura-bactéria incluem o YEp24 [Botstein et al. (1979) Gene 8:17-24], pCl1/1 [Brake et al. (1984) Proc. Natl. Acad. Sci USA

81:4642-4646] e YRp17 [Stinchcomb et al. (1982) J. Mol. Biol. 158:157]. 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 seleccionar-se um vector com um elevado ou com um baixo número de cópias, dependendo do efeito que o vector 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 vector integrativo. Os vectores integrativos contêm geralmente pelo menos uma sequência homóloga a um cromossoma de levedura que permite a integração do vector, e de preferência contêm duas sequências homólogas flanqueando a construção de expressão. As integrações parecem resultar de recombinações entre o DNA homólogo no vector e o cromossoma da levedura [Orr-Weaver et al. (1983) Methods in Enzymol. 101:228-245]. Um vector integrativo poderá ser dirigido para um locus específico na levedura através da selecção da sequência homóloga apropriada para inclusão no vector. Ver Orr-Weaver et al., *supra*. Um ou mais construções de expressão poderão sofrer integração, possivelmente afectando os níveis de proteína recombinante produzidos [Rine et al. (1983) Proc. Natl. Acad. Sci. USA 80:6750]]. As sequências cromossómicas incluídas no vector poderão ocorrer sob a forma de um único segmento no vector, o que resulta na integração do vector completo, ou de dois segmentos homólogos a segmentos adjacentes no cromossoma e flanqueando a construção de expressão no vector, o que poderá resultar na integração estável de apenas a construção de expressão.

Usualmente, as construções de expressão extracromossómicas e integrativas poderão conter marcadores de selecção que permitam a selecção das estirpes de leveduras que foram

transformadas. Os marcadores de selecção 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 selecção 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 vectores de transformação. Os vectores de transformação incluem geralmente um marcador de selecção que poderá ser mantido num replicon ou desenvolvido para um vector integrativo, tal como acima descrito.

Têm sido desenvolvidos diversos vectores de expressão e de transformação, tanto sob a forma de replicons extracromossómicos como de vectores integrativos, para a transformação de um grande número de leveduras. Por exemplo, foram desenvolvidos vectores de expressão e métodos para introdução de DNA 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 guillerimondii* [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 DNA 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 cationes 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 *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.

Uma "origem de replicação" corresponde a uma sequência polinucleotídica que inicia e regula a replicação de polinucleótidos, tal como um vector 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 vector se replique numa determinada célula hospedeira. Com certas origens de replicação, um vector 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 antigénio T viral, o qual é eficaz nas células COS-7.

Uma sequência "mutante" é definida como uma sequência de DNA, RNA 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% (p.ex., 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 sequência codificante codifica tipicamente para uma proteína que possui uma actividade 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, p.ex., a Patente dos EUA N° 5.753.235).

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 forma de superfície interna e uma distribuição de cargas que é complementar às características de um epítopo de um antigénio, o que permite a ligação entre o anticorpo e o antigénio. 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 do invento encontram utilidade em procedimentos de cromatografia por afinidade, em imunoensaios e na destrinça/identificação de proteínas de *Neisseria menB*. Os anticorpos dirigidos contra as proteínas do presente invento ligam-se a polipéptidos antigénicos ou proteínas, ou fragmentos, de proteína que estão presentes em, e especificamente associados a, estirpes de *Neisseria meningitidis menB*. Em alguns casos, estes antigénios poderão estar associados a estirpes específicas, tal como os antigénios específicos das estirpes de menB. Os anticorpos do invento podem ser immobilizados numa matriz e utilizados em imunoensaios ou em procedimentos de cromatografia em coluna por afinidade, de maneira a permitir a detecção e/ou separação de polipéptidos, proteínas ou fragmentos de proteína ou células que incluem esses polipéptidos, proteínas ou fragmentos de proteína. Alternativamente, esses polipéptidos, proteínas ou fragmentos de proteína poderão ser immobilizados de forma a que possam ser detectados anticorpos especificamente capazes de se ligarem a esses polipéptidos, proteínas ou fragmentos de proteína.

Os anticorpos dirigidos contra as proteínas do invento, 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 invento, é considerado como equivalente à imunização in vivo. Os soros policlonais são obtidos 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 (p.ex. 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 seleccionados (após a remoção das células aderentes não específicas) através da aplicação de uma suspensão celular a uma placa ou poço

revestidos com o antigénio proteico. As células B que expressem uma imunoglobulina ligada à membrana específica para o antigénio 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 selectivo (p.ex., 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 liguem especificamente ao antigénio imunizante (e que não se liguem a antigénios não relacionados). Os hibridomas secretores de mAc seleccionados são então cultivados *in vitro* (p.ex., em frascos de cultura de tecidos ou em reactores 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 radioactivos (particularmente ^{32}P e ^{125}I), reagentes densos a electrões, enzimas e ligandos possuindo parceiros de ligação específica. As enzimas são tipicamente detectadas por meio da sua actividade. Por exemplo, a peroxidase de rábano é geralmente detectada 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}I poderá funcionar como um marcador radioactivo ou como um reagente denso a electrões. 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 ao se usarem no âmbito deste invento; deste modo, poder-se-á marcar um mAc com biotina e detectar a sua presença através de avidina marcada com ^{125}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 do presente invento.

Os antigénios, imunogénios, polipéptidos, proteínas ou fragmentos de proteína do presente invento estimulam a formação de anticorpos com parceiros específicos de ligação. Estes antigénios, imunogénios, polipéptidos, proteínas ou fragmentos de proteína do presente invento compreendem composições imunogénicas do presente invento. 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 invento. 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 do invento. As composições farmacêuticas compreenderão uma quantidade

terapeuticamente eficaz dos polipéptidos, anticorpos ou polinucleótidos do invento 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 detectável. O efeito poderá ser detectado através de, por exemplo, marcadores químicos 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 seleccionadas para administração. Assim, não será útil especificar antecipadamente uma quantidade eficaz exacta. 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 invento, uma dose eficaz situar-se-á entre cerca de 0,01 mg/kg e 50 mg/kg, ou 0,05 mg/kg a cerca de 10 mg/kg das construções de DNA para o indivíduo ao qual são administradas.

Uma composição farmacêutica poderá igualmente incluir um veículo farmacêuticamente aceitável. O termo "veículo farmacêuticamente 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 inactivas. Tais veículos são bem conhecidos dos técnicos de perícia média neste campo.

Poderão utilizar-se sais farmacologicamente 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 farmacologicamente 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 injectáveis, 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 do presente invento poderão ser administradas directamente ao indivíduo. Os indivíduos a tratar poderão ser animais e, em particular, seres humanos.

A administração directa 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 genes 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 o invento poderão ter fins profilácticos (i.e., de prevenção da infecção) ou terapêuticos (i.e., de tratamento da doença após a infecção).

Estas vacinas compreenderão antigénio(s) imunizante(s), imunogénio(s), polipéptido(s), proteína(s) ou fragmentos de proteína ou ácido nucleico (p.ex., ácido ribonucleico ou ácido desoxiribonucleico), geralmente em combinação com "veículos farmacologicamente 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 inactivas. 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 toxóide bacteriano, tal como um toxóide de ou dos agentes patogénicos de difteria, tétano, cólera, *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® (WO 90/14837); (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 submicrónica 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 (p.ex., interferão gama), factor estimulador de colónias de macrófagos (M-CSF), factor de necrose de tumor (TNF), etc.; e (6) outras substâncias que actuam 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 (p.ex., que incluam o antigénio, o veículo farmacêuticamente 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 farmacêuticamente 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 (p.ex., um primata não humano, um primata, etc.) a capacidade de produção de anticorpos evidenciada pelo sistema imunológico do indivíduo, o grau de protecção desejada, a formulação da vacina, a avaliação da situação clínica por parte do médico e outros factores 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 injectáveis, 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, p.ex. 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 DNA [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, utilizados para a administração de construções que incluem uma sequência codificante para um produto terapêutico do invento, poderão ser administrados localmente ou sistemicamente a um mamífero para que se dê a expressão no mamífero. Estas construções poderão utilizar abordagens com vectores 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.

O invento 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 vector viral e, mais preferencialmente, a um vector retroviral, adenoviral, viral adeno-associado (AAV), de herpesvírus ou de alfavírus. O vector 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 vectores 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 vector retroviral para terapia genética poderão derivar de diferentes retrovírus. Por exemplo, as LTRs do retrovector poderão derivar do Vírus do Sarcoma Murino, um local de ligação ao tRNA 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 vectores retrovirais recombinantes poderão ser usados para gerar partículas de vector retroviral competentes para transdução, recorrendo à sua introdução em linhas celulares de empacotamento apropriadas (ver a Patente dos EUA

Nº 5.591.624). Os vectores de retrovírus poderão ser construídos para integração específica de local no DNA da célula hospedeira através da incorporação de uma enzima integrase quimérica na partícula retroviral (ver WO96/37626). Prefere-se que o vector viral recombinante corresponda a um vírus recombinante de replicação defectiva.

As linhas celulares de empacotamento adequadas para uso em conjunção com os vectores de retrovírus acima descritos são bem conhecidos da técnica, são facilmente preparados (ver WO95/30763 e WO92/05266) e podem ser usados para criar linhas celulares produtoras (também designadas por linhas celulares de vector ou "LCVs") para a produção de partículas de vector recombinante. Preferencialmente, as linhas celulares de empacotamento são produzidas a partir de células humanas progenitoras (p.ex., células HT1080) ou linhas celulares progenitoras de arminho, o que elimina a inactivação no soro humano.

Os retrovírus preferidos para a construção de vectores 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 (Nº ATCC VR-999), Friend (Nº ATCC VR-245), Graffi, Gross (Nº ATCC VR-590), Kirsten, Vírus do Sarcoma de Harvey e Rauscher (Nº ATCC VR-998) e Vírus da Leucemia Murina de Moloney (Nº 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 invento 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 5,219,740, US 4,405,712, US 4,861,719, US 4,980,289, US 4,777,127, US 5,591,624. 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 neste invento. 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 neste invento 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 DNA 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 o invento incluem vetores de vírus associado a adenovírus (AAV). Os principais, e preferidos, exemplos de tais vetores para uso neste invento 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, e 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 vectores AAV utilizáveis encontram-se o pWP-19 e o pWN-1, ambos descritos em Nahreini (1993) Gene 124:257-262. O vector psub201 constitui outro exemplo de um vector AAV (ver Samulski (1987) J. Virol. 61:3096). Ainda outro exemplo de vector AAV é dado pelo vector ITR Double-D. A construção do vector ITR Double-D é exposta na Patente dos EUA nº 5.478.745. Outros vectores 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 vector AAV utilizável neste invento corresponde ao vector SSV9AFABTKneo, o qual contém o enhancer 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 vectores 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 vectores para terapia genética compreendendo sequências deste invento incluem igualmente vectores de herpesvírus. Os exemplos principais e preferidos correspondem a vectores de vírus herpes simplex 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 vectores de vírus herpes simplex incluem os vectores HFEM/ICP6-LacZ apresentadas na patente WO95/04139 (Wistar Institute), pHSVlac, descrito em Geller (1988) Science 241:1667-1669, em WO90/109441 e em WO92/07945; 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 vectores depositados na ATCC com os números de acesso ATCC VR-977 e ATCC VR-260.

Está igualmente contemplado neste invento o uso de vectores de alfavírus para terapia genética. Os vectores de alfavírus preferidos são os vectores 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 dos EUA nos 5.091.309, 5.217.879 e WO92/10578. Mais particularmente, são utilizáveis os vectores de alfavírus descritos na patente dos EUA com o nº de série 08/405.627, registada em 15 de Março de 1995, e nas patentes dos EUA nos 5.091.309 e 5.217.879 e WO94/21792, WO92/10578 e WO95/07994. Tais alfavírus 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. Serão preferencialmente usados vectores de alfavírus com reduzida citotoxicidade (ver USSN 08/679640).

Os sistemas vectores de DNA tais como os sistemas eucarióticos de expressão em camadas são igualmente úteis para a expressão dos ácidos nucleicos do invento. Ver a WO95/07994 para obter uma descrição detalhada dos sistemas eucarióticos de expressão em camadas. Preferencialmente, os sistemas eucarióticos de expressão em camadas de acordo com este invento derivam de vectores de alfavírus e, mais preferencialmente, de vectores de vírus Sindbis.

Outros vectores virais adequados para uso no âmbito do presente invento 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 dos EUA 4.603.112, EUA 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 Kyzylgach, 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 Trinita, 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 do invento a células não se encontra limitada aos vectores virais acima mencionados. Poderão utilizar-se outros métodos e meios de libertação, tal como, por exemplo, vectores de expressão de ácidos nucleicos, DNA 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), DNA 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 dos EUA n° 5.149.655), radiação ionizante (tal como descrito na Patente dos EUA 5.206.152 e em WO92/11033), 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 dos EUA com o n° de série 60/023.867. Abreviadamente, a sequência poderá ser inserida em vectores convencionais que contêm sequências de controlo 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 DNA, como poli-lisina, protamina e albumina, ligadas a ligandos que se direccionam 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 DNA nu para transformar uma célula hospedeira. Encontram-se exemplos de métodos para introdução de DNA nu nas patentes WO90/11092 e US 5.580.859. A eficiência de captação poderá ser melhorada usando esferas biodegradáveis de látex. As esferas de látex revestidas por DNA 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 ruptura do endossoma e a libertação do DNA para o citoplasma.

Encontram-se descrições de lipossomas que podem actuar como veículos para libertação de genes na patente dos EUA 5.422.120 e nas patentes WO95/13796, WO94/23697, WO91/14445 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 vectores convencionais que contêm sequências de controlo 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 DNA, 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 DNA compreendendo o gene sob o controlo de diversos promotores específicos de tecido ou ubiquitariamente activos. 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 5.149.655; o uso de radiação ionizante para a activação do gene transferido, tal como descrito na patente dos EUA 5.206.152 e em WO92/11033.

Como exemplos de veículos lipossómicos ou policatiónicos para transferência de genes encontram-se os descritos nas patentes dos EUA 5.422.120 e 4.762.915; WO95/13796; WO94/23697; WO91/14445; 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 terapeuticamente eficaz de um veículo para

transferência de genes, de acordo com a definição do termo acima apresentada. Para os fins do presente invento, 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 DNA para o indivíduo ao qual serão administradas.

Métodos de Administração

Uma vez formuladas, as composições polinucleotídicas do invento serão administradas (1) directamente 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 directa das composições será geralmente efectuada 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, p.ex., 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 microinjecção directa do DNA 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 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 circunsporozoítio de Plasmodium falciparum.

B. Hormonas, Vitaminas, etc.

Outros grupos que poderão ser incluídos correspondem, por exemplo, a: hormonas, esteróides, androgénios, estrogénios, hormona da tiróide, 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 aspecto, 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 DNA: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 no presente invento incluem preparações catiónicas (carregadas positivamente), aniónicas (carregadas negativamente), e neutras. Verificou-se que os lipossomas catiónicos medeiam a libertação intracelular de DNA plasmídico (Felgner (1987) *Proc. Natl. Acad. Sci. USA* 84:7413-7416); mRNA (1989) *Proc. Natl. Acad. Sci. USA* 86:6077-6081); e factores 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-diioleiloxi)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, igualmente, Szoka (1978) Proc. Natl. Acad. Sci. USA 75:4194-4198, e WO90/111092, para obter uma descrição da síntese de lipossomas de DOTAP (1,2-bis(oleiloxi)-3-(trimetilamonio)propano).

De forma similar, os lipossomas aniônicos 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 nucleico são preparados usando métodos conhecidos neste campo. Ver, p.ex., 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 direccionador 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 e adquirem 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 actividade de ligação ao receptor. A composição em lípidos poderá igualmente ser escolhida de modo a facilitar a interacção 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 em uma 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 eléctrica 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 DNA, proteínas cromossómicas não-histónicas, proteínas de revestimento de vírus DNA, tais como a proteína X174; os factores transcricionais contêm igualmente domínios que se ligam ao DNA e que, deste modo, poderão ser úteis como agentes condensadores de ácidos nucleicos. Abreviadamente, os factores 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 DNA.

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

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

Policatiónicos Sintéticos

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

Ensaio para Imunodiagnóstico

Os antigénios de *Neisseria* do invento poderão ser usados em imunoensaios para detectar as concentrações de anticorpo (ou, inversamente, os anticorpos anti-*Neisseria* podem ser usados para detectar as concentrações de antigénio). Poderão desenvolver-se imunoensaios baseados em antigénios recombinantes bem definidos para substituir os métodos invasivos de diagnóstico. Será possível detectar anticorpos contra proteínas de *Neisseria* em amostras biológicas, incluindo, por exemplo, amostras de sangue ou de soro. A concepção 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 reacção directa 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, radioactivos 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 invento, 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 factores que afectam esta ligação incluem: tipo e volume do solvente; temperatura de reacção; 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 reacção 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 DNA genómico immobilizadas 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, p.ex., um Southern blot são (1) a complexidade do DNA a ser submetido a blot e (2) a homologia entre a sonda e a sequência a detectar. A quantidade total do(s) fragmento(s) a estudar poderá variar numa magnitude de 10, desde 0,1 a 1 μg 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, com tempos de blot, hibridização e exposição substancialmente mais curtos, poderão usar-se uma quantidade menor de polinucleótidos de partida e uma menor actividade específica das sondas. Por exemplo, um gene de levedura de uma só cópia poderá ser detectado com um tempo de exposição de apenas 1 hora, partindo de 1 μg de DNA de levedura e efectuando 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 DNA, 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 afectar a temperatura de fusão (T_m) de um híbrido DNA-DNA entre a sonda e o fragmento de interesse e, conseqüentemente, 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 factores 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 concepção de um protocolo experimental de hibridização, será possível alterar de modo conveniente alguns dos factores que afectam a hibridização dos ácidos nucleicos. A temperatura de hibridização e das lavagens, e a concentração em sais durante as lavagens, são os factores 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 interespécies), a temperatura de hibridização terá de ser reduzida, o que resulta no aumento do efeito de fundo. A temperatura das lavagens afecta 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 de elevado efeito de fundo após autoradiografia, o filtro poderá ser lavado em condições de elevada restrição e re-exposto. Se o tempo requerido para a exposição tornar esta abordagem impraticável, deverão testar-se em paralelo diversas restrições de hibridização e/ou lavagem.

Ensaio com Sondas de Ácidos Nucleicos

Os métodos como a PCR, os ensaios de sonda de DNA ramificada ou as técnicas de blotting usando sondas de ácido nucleico de acordo com o invento têm a capacidade de determinar a presença de cDNA ou mRNA. Diz-se que uma sonda se "hibridiza" com uma sequência do invento 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* do invento (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 mRNA representa uma sequência codificante, pelo que uma sua sonda deverá ser complementar à sequência codificante; o cDNA de cadeia simples é complementar ao mRNA, pelo que uma sonda de cDNA 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, os quais poderão ser detectados. 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 detectado.

A extensão e sequência exactas 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 conterà tipicamente pelo menos 10-20 nucleótidos, preferencialmente 15-25, e mais preferencialmente pelo menos 30 nucleótidos, se bem que possa ser mais curta. Os primers 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 seleccionada segundo pretendido. Para algumas aplicações, o DNA ou o RNA são adequados. Para certas 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 RNA, 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]; 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 reacção em cadeia por polimerase (PCR) constitui um outro meio bem conhecido para a detecção de pequenas quantidades de ácido nucleico 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 "primer" hibridizam-se com os ácidos nucleicos alvo e são usados como primers para a reacção. Os primers poderão compreender uma sequência que não se hibridize com a sequência do alvo de amplificação (ou com o seu complemento) e 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 primers 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 detectados através de métodos mais

tradicionais, tais como 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 cDNA poderão ser detectados através de técnicas convencionais de blotting como as descritas em Sambrook et al. [*supra*]. O mRNA, ou o cDNA 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 detectados. Tipicamente, a sonda é marcada com uma porção molecular radioactiva.

Exemplos

Os exemplos que se seguem descrevem sequências de ácidos nucleicos que foram identificados em *N.meningitidis* e em *N.gonorrhoeae*, bem como os seus respectivos produtos de tradução deduzidos. Os exemplos incluem:

- Uma sequência de nucleótidos que foi identificada em *N.meningitidis*
- O produto de tradução deduzido para a dita sequência de *N.meningitidis*
- 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*.

As comparações de sequência foram realizadas no NCBI - National Center For Biotechnology Information em Bethesda, Maryland, E.U.A. (<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.

As letras minúsculas representam ambiguidades que surgiram durante o alinhamento das reacções de sequenciação independentes (algumas das sequências de nucleótidos presentes nos Exemplos derivam da combinação de resultados entre duas ou mais experiências).

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 ORFs, 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).

Prevê-se que as proteínas de *N.meningitidis* e de *N.gonorrhoeae* e os seus respectivos epitopos possam ser úteis como antigénios ou composições imunogénicas para aplicação em vacinas ou como meios de diagnóstico. Esta previsão baseia-se na presença de uma presumível sequência leader e/ou de vários presumíveis domínios transmembranares (a sublinhado simples) na proteína do gonococo.

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

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

Preparação do DNA 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%, Tris-HCl 50 mM e EDTA 50 mM, ajustados a um pH de 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 extracções, uma com fenol (equilibrado a pH 8) e uma com CHCl₃/álcool isoamílico (24:1). Por adição de acetato de sódio 0,3M e de 2 volumes de etanol, o DNA foi precipitado e recolhido por centrifugação. O pellet foi lavado uma vez com etanol a 70% e foi redissolvido em 4 ml de tampão (Tris-HCl 10 mM e EDTA 1mM a pH 8). Mediu-se a concentração de DNA através de leitura da DO a 260 nm.

Concepção de oligonucleótidos

Foram concebidos primers 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 preferencial do codão do meningococo. Deduzindo a extremidade 5' da sequência do primer de amplificação imediatamente a jusante da sequência leader prevista foi possível omitir todos os péptidos de sinal previstos.

Para a maior parte dos ORFs, os primers 5' incluíram dois locais de reconhecimento para enzimas de restrição (*Bam*HI-*Nde*I, *Bam*HI-*Nhe*I, ou *Eco*RI-*Nhe*I, dependendo do padrão de restrição do gene); os primers 3' incluíram um local de restrição para *Xho*I. 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 ou *Eco*RI-*Xho*I), e pET21 b+ (usando *Nde*I-*Xho*I ou *Nhe*I-*Xho*I)

cauda 5'-terminal	<u>CGCGGATCCCATATG</u>	(<i>Bam</i> HI- <i>Nde</i> I)	<ID NO:7>	SEQ
do primer:	<u>CGCGGATCCGCTAGC</u>	(<i>Bam</i> HI- <i>Nde</i> I)	<ID NO:8>	SEQ
	<u>CCGGAATTCTAGCTAGC</u>	(<i>Eco</i> RI- <i>Nhe</i> I)	<ID NO:9>	SEQ
cauda 3'-terminal	<u>CCCGCTCGAG</u>	(<i>Xho</i> I)	<ID NO:10>	SEQ
do primer:				

Para alguns ORFs, foram realizadas duas amplificações diferentes para possibilitar a clonagem de cada ORF nos dois sistemas de expressão. Foram usados dois primers 5' diferentes para cada ORF, tendo o primer 3' de *Xho*I sido utilizado como anteriormente:

cauda 5'-terminal GGAATTCATATGGCCATGG (*NdeI*) <ID SEQ
do primer: NO:11>
cauda 5'-terminal CGGGATCC (*BamHI*) <ID SEQ
do primer: NO:12>

Outros ORFs poderão ser clonados no vector de expressão pTRC e expressos como um produto de fusão amino-terminal com His. O péptido de sinal previsto poderá ser incluído no produto final. Os locais de restrição para *NheI-BamHI* foram incorporados usando primers:

cauda 5'-terminal GATCAGCTAGCCATATG (*NheI*) <ID SEQ
do primer: NO:13>
cauda 3'-terminal CGGGATCC (*BamHI*) <ID SEQ
do primer: NO:14>

Os primers incluíram não só as sequências de reconhecimento para enzimas de restrição como também nucleótidos que se hibridizaram com a sequência a amplificar. O número de nucleótidos hibridizantes dependeu do ponto de fusão do primer completo e foi determinado para cada primer usando a seguinte fórmula:

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

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

A temperatura média de fusão dos oligonucleótidos seleccionados foi de 65-70°C para o oligonucleótido completo e de 50-55°C para a região de hibridização apenas.

Os oligonucleótidos foram sintetizados recorrendo a um sintetizador automatizado de oligonucleótidos Perkin Elmer 394 DNA/RNA, eluídos das colunas usando 2 ml NH₄-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 DO₂₆₀ foi determinada recorrendo a um espectrofotómetro Perkin Elmer Lambda Bio e a concentração foi determinada e ajustada para 2-10 pmol/µl.

Quando as amplificações iniciais são realizadas, poderão não se conhecer as sequências 5' e/ou 3' completas para alguns ORFs de meningococo, apesar de as sequências correspondentes poderem ter já sido identificadas no gonococo. Para a amplificação, as sequências de gonococo poderão deste modo ser usadas como base para a concepção de primers, sendo alteradas de acordo com o codão de preferência. Em particular, destacam-se os seguintes codões como passíveis de modificação: ATA → ATT; TCG → TCT; CAG → CAA; AAG → AAA; GAG → GAA; CGA and CGG → CGC; GGG → GGC.

Amplificação

Foi usado o seguinte protocolo padronizado de PCR: utilizou-se 50-200 ng de DNA genómico como molde na presença de 20-40 µM de cada oligonucleótido, dNTPs em solução 400-800 µM, 1 x tampão PCR (incluindo MgCl₂ 1,5 mM), 2,5 unidades de DNA polimerase *TaqI* (usando AmpliTaq da Perkin-Elmer, Platinum da GIBCO, DNA 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 betaína 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 oligonucleótidos sem cauda para enzimas de

restrição, seguindo-se 30 ciclos efectuados à temperatura de hibridização que corresponde aos oligonucleótidos de extensão completa. Os ciclos foram seguidos por um passo final de extensão de 10 minutos a 72°C.

Os ciclos-padrão foram realizados da seguinte forma:

	Desnaturalização	Hibridização	Alongamento
Primeiros 5 ciclos	30 segundos, 95°C	30 segundos, 50-55°C	30-60 segundos, 72°C
Últimos 30 ciclos	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% e a dimensão de cada fragmento amplificado foi então comparada com um marcador de peso molecular de DNA.

O DNA 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%. O fragmento de DNA correspondente à banda de tamanho apropriado foi então eluído e purificado a partir do gel, usando o Gel Extraction Kit da Qiagen de acordo com as instruções do fabricante. O volume final do fragmento de DNA correspondeu a 30µl ou 50µl de água ou de Tris 10mM, pH 8,5.

Digestão dos fragmentos de PCR

O DNA purificado que corresponde ao fragmento amplificado foi dividido em 2 alíquotas e submetido a dupla digestão com:

(a) *NdeI/XhoI* ou *NheI/XhoI* para clonagem em pET-21 b+ e expressão subsequente da proteína como produto de fusão C-terminal com His; (b) *BamHI/XhoI* ou *EcoRI/XhoI* para clonagem em pGEX-KG e expressão subsequente da proteína como produto de fusão N-terminal com GST.

Para alguns ORFs, poderá utilizar-se *NheI/BamHI* para clonagem num vector pTRC-HisA e expressão subsequente da proteína como produto de fusão N-terminal com His.

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

Digestão dos vectores de clonagem (pET22B, pGEX-KG e pTRC-His A)

Submeteu-se 10 µg de plasmídeo a digestão dupla com 50 unidades de cada enzima de restrição num volume de reacção de 200 µl e na presença do tampão apropriado, num 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 vector digerido foi purificada a partir do gel utilizando o QIAquick Gel Extraction Kit da Qiagen e o DNA foi eluído em 50 µl de Tris-HCl 10 mM, pH 8,5. A concentração

de DNA foi avaliada através da medição da DO_{260} da amostra, sendo então ajustada a 50 g/ μ l. Utilizou-se 1 μ l de plasmídeo por cada procedimento de clonagem.

Clonagem

Os fragmentos correspondentes a cada ORF, previamente digeridos e purificados, foram ligados em pET22b bem como em pGEX-KG. Num volume final de 20 μ l, o procedimento de ligação efectuou-se com uma razão molar fragmento/vector de 3:1 e utilizando 0,5 μ l de DNA ligase NEB T4 (400 unidades/ μ l), na presença do tampão fornecido pelo fabricante. A reacção 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 μ l de células competentes de *E. coli* DH5 foram incubadas com a solução de reacção de ligase em gelo durante 40 minutos, seguindo-se uma incubação a 37°C durante 3 minutos, e seguidamente, após a adição de 800 μ l 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 μ l do sobrenadante. A suspensão foi então plaqueada em LB-ampicilina (100 mg/ml).

A detecção 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 μ g/ml. Foi então preparado um pellet das células e o DNA 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 μ l. 5 μ l 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 (fragmento de DNA de 1 Kb, GIBCO). A detecção dos clones positivos foi realizada com base no tamanho correcto do fragmento inserido.

Clonagem

Alguns ORFs podem ser clonados no vector pGEX-HIS utilizando os locais de clonagem para *EcoRI-PstI*, ou *EcoRI-SalI*, ou *SalI-PstI*. Após a clonagem, os plasmídeos recombinantes podem ser introduzidos no hospedeiro *E.coli* W3110.

Expressão

Cada ORF clonado no vector 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 (vector pGEX), *E.coli* TOP 10 (vector pTRC) ou *E.coli* BL21-DE3 (vector pET), de acordo com o acima descrito. No caso do vector 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 DO₆₀₀ desta diluição inicial se encontrava entre 0,1 e 0,15. Os frascos foram incubados a 30°C dentro de 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 vectores pET e pTRC; 0,8-1 de

DO para os vectores pGEX e pGEX-His). A expressão de proteínas nos vectores pET, pTRC e pGEX-His foi induzida por adição de IPTG 1mM, enquanto que 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 ressuspensão 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 ressuspensão em PBS para utilização subsequente.

Purificação em larga escala de proteínas de fusão com GST

Cultivou-se uma única colónia durante 12 horas a 37°C numa placa de agar de LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido de 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 óptima (20-37°C) até uma DO_{550} 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_{280} de 0,02-0,06. A proteína de fusão com GST foi eluída pela adição de 700 μ l de tampão de eluição de Glutathione frio (glutathione reduzida 10mM, Tris-HCl 50mM) e foram recolhidas fracções até que a DO_{280} atingisse 0,1. Carregou-se um volume de 21 μ l 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.

Purificação a larga escala de proteínas de fusão com His solúveis

Cultivou-se uma única colónia durante 12 horas a 37°C numa placa de agar de LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido de 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 óptima (20-37°C) até uma DO_{550} 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 150 μ l de resina de Ni²⁺ (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 DO₂₈₀ de 0,02-0,06. A resina foi lavada a 4°C 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 DO₂₈₀ 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 fracções até que a DO₂₈₀ atingisse 0,1. Carregou-se um volume de 21µl de cada fracção 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 de LB + Amp. As bactérias foram inoculadas em 20 ml de meio líquido de 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 óptima (20-37°C) até uma DO₅₅₀ 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 ressuspensa em 7,5 ml de tampão B (ureia 8 M, Tris-HCl 10 mM, tampão fosfato 100 mM, imidazol 10 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 Ni²⁺ (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 DO₂₈₀ 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 DO₂₈₀ 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 fracções até que a DO₂₈₀ atingisse 0,1. Carregou-se um volume de 21µl de cada fracção num gel de SDS a 12%.

Renaturação de proteínas de fusão com His

Adicionou-se glicerol a 10% às proteínas desnaturadas. As proteínas foram então diluídas até 20µg/ml recorrendo ao tampão de diálise I (glicerol 10%, arginina 0,5M, tampão fosfato 50mM, glutatona reduzida 5mM, glutatona 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%, arginina 0,5M, tampão fosfato 50mM, glutatona reduzida 5mM, glutatona oxidada 0,5mM, pH 8.8) durante 12-14 horas a 4°C. A concentração de proteína foi determinada através da seguinte fórmula:

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

Imunização em ratinhos

Utilizaram-se 20µg de cada proteína purificada para imunizar ratinhos por via intraperitoneal. No caso de alguns ORFs, os ratinhos Balb-C foram imunizados utilizando Al(OH)₃ 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. Para os ORFs 25 e 40, imunizaram-se ratinhos CD1 utilizando o adjuvante de Freund, tendo sido seguido o mesmo protocolo de imunização, com a exceção de a resposta imunitária ter sido medida ao dia 42 em vez de ao dia 56. De modo similar, para ainda outros ORFs, imunizaram-se ratinhos CD1 utilizando o adjuvante de Freund, mas a resposta imunitária foi medida ao dia 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 DO₆₂₀. As bactérias foram deixadas em desenvolvimento 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%, NaN₃ 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 of H₂O) a cada poço e as placas foram deixadas à temperatura ambiente durante 20 minutos. Adicionaram-se 100 µl de H₂SO₄ a cada poço e mediu-se a DO₄₉₀. O teste ELISA foi considerado positivo quando se encontraram DO₄₉₀ 2,5 vezes superiores às determinadas com os respectivos soros de pré-imunização.

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 colônias 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 DO₆₂₀. As bactérias foram deixadas em cultura 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 ressuspenso num tampão de

bloqueio (1% BSA, 0.4% NaN₃) e centrifugado durante 5 minutos a 4000 rpm. As células foram ressuspensas num tampão de bloqueio até se verificar uma DO₆₂₀ de 0,07. Adicionaram-se 100µl de células bacterianas a cada poço de uma placa Costar de 96 poços. Adicionaram-se 100µl de soro diluído (1:200 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 200µl de tampão de bloqueio a cada poço. Adicionaram-se a cada poço 100 µl de 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 200µl de tampão de bloqueio a cada poço. O sobrenadante foi aspirado e as células foram ressuspensas em 200µl de PBS com 0,25% de formaldeído. As amostras foram transferidas para tubos FACScan e lidas. As configurações do FACScan foram : FL1 on, FL2 e FL3 off; Limiar FSC-H: 92; Voltagem FSC PMT: E 02; SSC PMT: 474; Ganhos de Amp. 7.1; FL-2 PMT: 539. Valores de Compensação: 0.

Preparações de OMV

As bactérias foram deixadas em cultura durante um período de 12 horas em 5 placas GC e foram recolhidas com uma ansa e ressuspensas em 10 ml de Tris-HCl 20mM. A inactivação por calor foi realizada 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 fracção 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 das membranas externas, a fracção foi ressuspensa na

totalidade em Sarkosyl (Sigma) a 2% e incubadas 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 Extractos Integrais

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

Western Blot

As proteínas purificadas e transferidas para uma membrana de nitrocelulose (500ng/aplicação), as vesículas de membranas externas (5 µg) e os extractos totais de células (25µg) derivados da estirpe 2996 de MenB foram carregados em SDS-PAGE a 15%.

A transferência foi realizada durante 2 horas a 150mA a 4°C, num tampão de transferência (0.3 % de base Tris, 1.44 % glicina, 20% metanol). A membrana foi saturada por incubação durante 12 horas a 4°C num 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 reacção foi interrompida através da adição de água.

Ensaio Bactericida

A estirpe MC58 foi cultivada 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 desenvolvimento até que o valor de DO₆₂₀ atingisse valores de 0,5-0,8. A cultura foi dividida em alíquotas em 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 DO₆₂₀ de 0,5, sendo então diluída 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 25µl de soro de ratinho diluído (1:100; tampão de diluição: tampão de Gey BSA a 0,2%) foram adicionados a cada poço, tendo a placa sido incubada de seguida a 4°C. Adicionaram-se 25µl da suspensão de bactérias acima a cada poço. Adicionaram-se a cada poço 25µl de complemento de coelho inactivado pelo calor (banho de água a 56°C durante 30 minutos) ou complemento normal não inactivados. 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 semeadas 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 submetidas a um processo de contagem.

Sequências

Utilizando os procedimentos acima descritos, identificou-se a seguinte sequência parcial de DNA em *N.gonorrhoeae*:

< ID SEQ NO:1>:

```
g741.seq
1   GTGAACCGAA CTACCTTCTG CTGCCTTTCT TTGACCGCCG GCCCTGATTC
51  TGACCGCCTG CAGCAGCGGA GGGGCGGAGG CGGTGGTGTG GCCGCCGACA
101 TCGGCACGGG GCTTGCCGAT GCATTAACCG CGCCGCTCGA CCATAAAGAC
151 AAAGTTTTGA AATCCCTAAC ATTGGAAGCC TCCATTCCCC AAAACGGAAC
201 ACTGACCCTG TCGGCACAAG GTGCGGAAAA AACTTTCAA AATCAGCCGC
251 AAGACAACAG CCTCAACAGG GCCAAAGTGA AGAACGACA AATCAGCCGC
301 TTCGACTTCG TGCAAAAAAT CGAAGTGGAC GGACAAACCA TCACACTGGC
351 AAGCGGCGAA TTTCAAATAT ACAAACAGGA TCACTCCGCC gtcgtTgcc
401 TacgGATTGA AAAAATCAAC AACCCCGACA AAATCGACAG CCTGATAAAC
451 CAACGCTCCT TCCTTGTGAG CGATTTGGGC GGAGAACATA CCGCCTTCAA
501 CCAACTGCCT GACGGCAAAG CCGAGTATCA CGGCAAAGCA TTCAGTCCG
551 ACGATGCCGA CGGAAAATCG ACCTATACCA TAGATTTTCGC GCCCAAACAG
601 GGACACGGCA AAATCGAACA CCTGAAAACA CCCGAGCAGA ATGTTGAGCT
651 TGCCTCCGCC GAACTCAAAG CAGATGAAAA ATCACACGCC GTCATTTTGG
701 GCGACACGCG CTACGGCGGC GAAGAGAAAG GCACCTACCG CCTCGCCCTT
751 TTCGGCGACC GCGCCCAAGA AATCGCTGGC TCGGCAACCG TGAAGATAGG
801 GGAAAAGGTT CACGAAATCG GCATCGCCGA CAAACAGTAG
```

A sequência seguinte corresponde à sequência de aminoácidos < ID SEQ NO:2; ORF 741.ng>:

```
g741.pep
1   VNRTTFCCLS LTAGPDSRDL QRRRGGGGGV AADIGTGLAD ALTAPLDHKD
51  KGLKSLTLEA SIPQNGTLTL SAQGAEKTFK AGGKDNSLNT GKLNKDKISR
101 FDFVQKIEVD GQITLASEG FQIYKQDHS A VVALRIEKIN NPKIDSLIN
151 QRSFLVSDLG GEHTAFNQLP DGKAEYHGKA FSSDDADGKL TYTIDFAAKQ
201 GHGKIEHLKT PEQNVELASA ELKADEKSHA VILGDTRYGG EEKGTYRLAL
251 FGDRAQEIAG SATVKIGEKV HEIGIADKQ*
```

A seguinte sequência parcial de DNA foi identificada em *N.meningitidis* < ID SEQ NO:3>:

```
m741.seq
1   GTGAATCGAA CTGCCTTCTG CTGCCTTTCT CTGACCACTG CCCTGATTCT
51  GACCGCCTGC AGCAGCGGAG GGGGTGGTGT CGCCGCCGAC ATCGGTGCGG
101 GGCTTGCCGA TGCACTAACC GCACCGCTCG ACCATAAAGA CAAAGGTTTG
151 CAGTCTTTGA CGCTGGATCA GTCCGTCAGG AAAAACGAGA AACTGAAGCT
201 GCGGCACAAA GGTGCGGAAA AAACCTTATGG AAACGGTGAC AGCCTCAATA
251 CGGGCAAATT GAAGAACGAC AAGGTCAGCC GTTTCGACTT TATCCGCCAA
301 ATCGAAGTGG ACGGGCAGCT CATTACCTTG GAGAGTGGAG AGTYCCAAGT
351 ATACAAACAA AGCCATTCGG CCTTAACCGC CTTTCAGACC GAGCAAATAC
401 AAGATTCGGA GCATTCCGGG AAGATGGTTG CGAAACGCCA GTTCAGAATC
451 GCGGACATAG CGGGCGAACA TACATCTTTT GACAAGCTTC CCGAAGGCGG
501 CAGGGCGACA TATCGCGGGA CGGCGTTCGG TTCAGACGAT GCCGGCGGAA
551 AACTGACCTA CACCATAGAT TCGCCGCCA AGCAGGGAAA CCGCAAATC
601 GAACATTTGA AATCGCCAGA ACTCAATGTC GACCTGGCCG CCGCCGATAT
651 CAGGCCGGAT GGAAGACGCG ATGCGGTLAT CAGCGGIFLL GELLIITACA
701 ACCAAGCCGA GAAAGGCAGT TACTCCCTCG GTATCTTTGG CGGAAAAGCC
751 CAGGAAGTTG CCGGCAGCGC GGAAGTGAAA ACCGTAAACG GCATACGCCA
801 TATCGGCCCT GCCGCCAAGC AATAA
```

A sequência seguinte corresponde à sequência de aminoácidos ID SEQ NO:4; ORF 741>:

m741.pep

```

1  VNRTAFCCLS LTTALILTAC SSGGGGVAAD IGAGLADALT APLDHKDKGL
51 QSLTLDQSVR KNEKLLAAQ GAEKTYGNGD SLNTGKLNKND KVSRLFDFIRQ
101 IEVDGQLITL ESGEFQVYKQ SHSALTAFQT EQIQDSEHSG KMYAKRQFRI
151 GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD AGGKLTYTID FAAKQGNQKI
201 EHLKSPELNV DLAAADIKPD GKRHAVISGS VLYNQAEKGS YSLGIFGGKA
251 QEVAGSAEVK TVNGIRHIGL AAKQ*
```

m741/g741 61,4% de identidade numa sobreposição de 280 aa

```

m741.pep      10      20      30      40      50
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g741           10      20      30      40      50      60
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m741.pep      60      70      80      90      100     110
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g741           70      80      90      100     110     120
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m741.pep     120     130     140     150     160     170
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g741          130     140     150     160     170
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m741.pep     180     190     200     210     220     230
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g741          180     190     200     210     220     230
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
m741.pep     240     250     260     270
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
g741          240     250     260     270     280
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
                | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

A seguinte sequência parcial de DNA foi identificada na N.meningitidis <ID SEQ NO:5>:

a741.seq

```

1  GTGAACCGAA CTGCCTTCTG CTGCCFTTCT TTGACCGCCG CCCTGATTCT
51 GACCGCCTGC AGCAGCGGAG GCGGCGGTGT CGCCGCCGAC ATCGGCGCGG
101 TGCTTGCCGA TGCCTAACC GCACCGCTCG ACCATAAAGA CAAAAGTTTG
151 CAGTCTTTGA CGCTGGATCA GTCCGTCAGG AAAAAACGAGA AACTGAAGCT
201 GCGGCGACAA GGTGCGGAAA AAACCTTATGG AAACGCGGAC AGCCTCAATA
251 CGGGCAAATT GAAGAACGAC AAGGTCAGCC GCTTCGACTT TATCCGTCAA
301 ATCGAAGTGG ACGGGCAGCT CATTACCTTG GAGAGCGGAG AGTTCCAAAGT
351 GTACAAACAA AGCCATTCCG CCTTAACCGC CCTTCAGACC GAGCAAGTAC
401 AAGATTCGGA GCATTCAGGG AAGATGGTTG CGAAACGCCA GTTCAGAATC
451 GCGGATATAG CGGGTGAACA TACATCTTTT GACAAGCTTC CCGAAGGCGG
501 CAGGGCGACA TATCGCGGGA CGGCATTCCG TTCAGACGAT GCCAGTGGAA
551 AACTGACCTA CACCATAGAT TTCGCCGCCA AGCAGGGACA CGGCAAAATC
601 GAACATTTGA AATCGCCAGA ACTCAATGTT GACCTGGCCG CCTCCGATAT
651 CAAGCCGGAT AAAAAACGCC ATGCCGTCAT CAGCGGTTCC GTCCTTTACA
701 ACCAAGCCGA GAAAGGCAGT TACTCTCTAG GCATCTTTGG CCGGCAAGCC
751 CAGGAAGTTG CCGGCAGCGC AGAAGTGGAA ACCGCAAACG GCATACGCCA
801 TATCGGTCTT GCCGCCAAGC AGTAA
```

A sequência seguinte corresponde à sequência de aminoácidos <ID SEQ NO:6; ORF 741.a>:

```

a741.pep
1  VNRTAFCCLS LTAALILTAC SSGGGGVAAD IGAVLADALT APLDHKDKSL
51  QSLTLDQSVR KNEKLKLAQ GAEKTYGNGD SLNTGKLNK KVS RFD FIRQ
101 IEVDGQLITL ESGEFQVYKQ SHSALTALQT EQVQDSEHSG KMVAKRQFRI
151 GDIAGEHTSF DKLPEGGRAT YRGTAFGSDD ASGKLYTID FAAKQGHGKI
201 EHLKSPELNV DLAASDIKPD KKRHAVISGS VLYNQAEGKS YSLGIFGGQA
251 QEVAGSAEVE IANGIRHIGL AAKQ*

```

a741/m741 95,6% de identidade numa sobreposição de 274 aa

```

a741.pep      10      20      30      40      50      60
VNRTAFCCLSLTAALILTACSSGGGGVAADIGAVLADALTAPLDHKDKSLQSLTLDQSVR
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m741          10      20      30      40      50      60
VNRTAFCCLSLTTALILTACSSGGGGVAADIGAGLADALTAPLDHKDKGLQSLTLDQSVR
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

a741.pep      70      80      90      100     110     120
KNEKLKLAQGAEKTYGNGDSLNTGKLNKDKVSRFD FIRQIEVDGQLITLESGEFQVYKQ
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m741          70      80      90      100     110     120
KNEKLKLAQGAEKTYGNGDSLNTGKLNKDKVSRFD FIRQIEVDGQLITLESGEFQVYKQ
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

a741.pep     130     140     150     160     170     180
SHSALTALQTEQVQDSEHSGKMOVAKRQFRIGDIAGEHTSFDKLPEGGRATYRGTAFGSDD
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m741         130     140     150     160     170     180
SHSALTAFQTEQIQDSEHSGKMOVAKRQFRIGDIAGEHTSFDKLPEGGRATYRGTAFGSDD
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

a741.pep     190     200     210     220     230     240
ASGKLYTIDFAAKQGHGKIEHLKSPELNVDLAASDIKPKKRHAVISGSVLYNQAEGKS
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m741         190     200     210     220     230     240
AGGKLYTIDFAAKQGNKIEHLKSPELNVDLAAADIKPDGKRHAVISGSVLYNQAEGKS
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

a741.pep     250     260     270
YSLGIFGGQAQEVAGSAEVEIANGIRHIGLAAKQX
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
m741         250     260     270
YSLGIFGGKAQEVAGSAEVKTVNGIRHIGLAAKQX
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

LISTAGEM DE SEQUÊNCIAS

<110> CHIRON CORPORATION

THE INSTITUTE FOR GENOMIC RESEARCH

<120> ANTIGÉNIOS E COMPOSIÇÕES DE NEISSERIA

<130> P042614EP

<150> US19980083758P

<151> 1998-05-01

<150> US19980094869P

<151> 1998-07-31

<150> US19980098994P

<151> 1998-09-02

<150> US19980099062P

<151> 1998-09-02

<150> US19980103749P

<151> 1998-10-09

<150> US19980103794P

<151> 1998-10-09

<150> US19980103796P

<151> 1998-10-09

<150> US19990121528P

<151> 1999-02-25

<150> EP99922752.2

<151> 1999-04-30

<160> 14

<170> SeqWin99, versão 1.02

<210> 1

<211> 840

<212> DNA

<213> *Neisseria gonorrhoeae*

<400> 1

```
gtgaaccgaa ctaccttctg ctgcctttct ttgaccgccc gccctgatte tgaccgcctg 60
cagcagcggg ggggcggagg cgggtggtgt gccqccgaca tcggcacggg gcttgccgat 120
gcattaaccg cgccgctcga ccataaagac aaaggtttga aatccctaac attggaagcc 180
tccattcccc aaaacggaac actgaccctg tcggcacaaag gtgeggaaaa aactttcaaa 240
gccggcggca aagacaacag cctcaacacg ggcaaaactga agaacgacaa aatcagccgc 300
ttcgacttcg tgcaaaaaat cgaagtggac ggacaaacca tcacactggc aagcggcgaa 360
tttcaaatat acaaacagga tcaactccgc gtcggtgccc tacggattga aaaaatcaac 420
aaccccgaca aaatcgacag cctgataaac caacgctcct tcttgtcag cgatttgggc 480
ggagaacata ccgccttcaa ccaactgcct gacggcaaag ccgagtatca cggcaaagca 540
ttcagctcgg acgatgccga cggaaaactg acctatacca tagatttcgc cgccaaacag 600
ggacacggca aaatcgaaca cctgaaaaca cccgagcaga atgttgagct tgccctccgc 660
gcactcaaaag cagatgaaaa atcacaagcc gtcattttgg gcgacaagcg ctacggcggc 720
gaagagaaaag gcacttaccg cctcgccctt ttcggcgacc gcgccaaga aatcgtggc 780
tcggcaaccg tgaagatagg ggaaaagggt cacgaaatcg gcatcgccga caaacagtag 840
```

<210> 2

<211> 279

<212> PRT

<213> *Neisseria gonorrhoeae*

<400> 2

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

<210> 3

<211> 825

<212> DNA

<213> Neisseria meningitidis

<400> 3

```
gtgaatcgaa ctgccttctg ctgcctttct ctgaccactg ccctgattct gaccgcctgc 60
agcagcggag ggggtggtgt cgccgccgac atcggtgccg ggcttgccga tgcaactaacc 120
gcaccgctcg accataaaga caaaggtttg cagtctttga cgctggatca gtccgtcagg 180
aaaaacgaga aactgaagct ggcggcaciaa ggtgcgaaa aaacttatgg aaacggtgac 240
agcctcaata cgggcaaatt gaagaacgac aaggtcagcc gtttcgactt tatccgcca 300
atcgaagtgg acgggcagct cattaccttg gagagtggag agttccaagt atacaaacaa 360
agccattccg ccttaaccgc ctttcagacc gagcaaatc aagattcggg gcattccggg 420
aagatggttg cgaaacgccs gttcagaatc ggcgacatag cgggcgaaca tacatctttt 480
gacaagcttc ccgaaggcgg cagggcgaca tatcgccgga cggcgttcgg ttcagacgat 540
gccggcggaa aactgacctc caccatagat ttcgccgcca agcagggaaa cggcaaaatc 600
gaacatttga aatcgccaga actcaatgtc gacctggccg ccgccgatat caagccgat 660
ggaaaacgcc atgccgtcat cagcggttcc gtcctttaca accaagccga gaaaggcagt 720
tactccctcg gtatctttgg cggaaaagcc caggaagtgg ccggcagcgc ggaagtgaaa 780
accgtaaacc gcatacgcca tatcggcctt gccgccaaagc aataa 825
```

<210> 4

<211> 274

<212> PRT

<213> Neisseria meningitidis

<400> 4

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

<210> 5

<211> 825

<212> DNA

<213> Neisseria meningitidis

<400> 5

```
gtgaaccgaa ctgccttctg ctgcctttct ttgaccgccg ccttgattct gaccgcctgc 60
agcagcggag gggcggtgt cggcgccgac atcggcgcgg tgcttgccga tgcactaacc 120
gcaccgctcg accataaaga caaaagtttg cagtctttga cgctggatca gtccgtcagg 180
aaaaacgaga aactgaagct ggcggcacia ggtgcggaia aaacttatgg aaacggcgac 240
agcctcaata cgggcaaatt gaagaacgac aaggtcagcc gcttcgactt tatccgtcaa 300
atcgaagtgg acgggcagct cattaccttg gagagcggag agttccaagt gtacaaacia 360
agccattccg cottaaccgc ccttcagacc gagcaagtac aagattcggg gcaftcaggg 420
aagatggttg cgaaaeccca gttcagaatc ggcgatatag cgggtgaaca tacatctttt 480
gacaagcttc ccgaaggcgg cgggcgaca tatcgcgga cggcattcgg ttcagacgat 540
gccagtgga aactgaccta caccatagat ttcgccgcca agcagggaca cggcaaaatc 600
gaacatttga aatcgccaga actcaatggt gacctggccg cctccgatat caagccggat 660
aaaaaacgcc atgcccgtcat cagcggttcc gtcctttaca accaagccga gaaaggcagt 720
tactctctag gcatcttttg cgggcaagcc caggaagttg ccggcagcgc agaagtggaa 780
accgcaaacg gcatacgcga tatcgggtctt gccgccaagc agtaa 825
```

<210> 6

<211> 274

<212> PRT

<213> Neisseria meningitidis

<400> 6

```
Val Asn Arg Thr Ala Phe Cys Cys Leu Ser Leu Thr Ala Ala Leu Ile
1          5          10          15
Leu Thr Ala Cys Ser Ser Gly Gly Gly Gly Val Ala Ala Asp Ile Gly
20          25          30
Ala Val Leu Ala Asp Ala Leu Thr Ala Pro Leu Asp His Lys Asp Lys
35          40          45
Ser Leu Gln Ser Leu Thr Leu Asp Gln Ser Val Arg Lys Asn Glu Lys
50          55          60
Leu Lys Leu Ala Ala Gln Gly Ala Glu Lys Thr Tyr Gly Asn Gly Asp
65          70          75          80
```

Ser Leu Asn Thr Gly Lys Leu Lys Asn Asp Lys Val Ser Arg Phe Asp
 85 90 95
 Phe Ile Arg Gln Ile Glu Val Asp Gly Gln Leu Ile Thr Leu Glu Ser
 100 105 110
 Gly Glu Phe Gln Val Tyr Lys Gln Ser His Ser Ala Leu Thr Ala Leu
 115 120 125
 Gln Thr Glu Gln Val Gln Asp Ser Glu His Ser Gly Lys Met Val Ala
 130 135 140
 Lys Arg Gln Phe Arg Ile Gly Asp Ile Ala Gly Glu His Thr Ser Phe
 145 150 155 160
 Asp Lys Leu Pro Glu Gly Gly Arg Ala Thr Tyr Arg Gly Thr Ala Phe
 165 170 175
 Gly Ser Asp Asp Ala Ser Gly Lys Leu Thr Tyr Thr Ile Asp Phe Ala
 180 185 190
 Ala Lys Gln Gly His Gly Lys Ile Glu His Leu Lys Ser Pro Glu Leu
 195 200 205
 Asn Val Asp Leu Ala Ala Ser Asp Ile Lys Pro Asp Lys Lys Arg His
 210 215 220
 Ala Val Ile Ser Gly Ser Val Leu Tyr Asn Gln Ala Glu Lys Gly Ser
 225 230 235 240
 Tyr Ser Leu Gly Ile Phe Gly Gly Gln Ala Gln Glu Val Ala Gly Ser
 245 250 255
 Ala Glu Val Glu Thr Ala Asn Gly Ile Arg His Ile Gly Leu Ala Ala
 260 265 270
 Lys Gln

<210> 7

<211> 15

<212> DNA

<213> Sequência Artificial

<220>

<223> primer

<400> 7

cgcggatccc atatg 15

<210> 8
<211> 15
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 8

cgcggatccg ctagc 15

<210> 9
<211> 17
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 9
ccggaattct agctagc 17

<210> 10
<211> 10
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 10
cccgctcgag 10

<210> 11
<211> 20
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 11
ggaattccat atggccatgg 20

<210> 12
<211> 8
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 12
cgggatcc 8

<210> 13
<211> 17
<212> DNA
<213> Sequência Artificial

<220>
<223> primer

<400> 13
gatcagctag ccatatg 17

<210> 14

<211> 8

<212> DNA

<213> Sequência Artificial

<220>

<223> primer

<400> 14

cgggatcc 8

Lisboa, 23 de Janeiro de 2008

REIVINDICAÇÕES

1. Uma proteína, caracterizada por compreender:

- a sequência de aminoácidos ID SEQ NO:2;
- uma sequência de aminoácidos possuindo uma identidade de sequência de 50% ou mais relativamente à ID SEQ NO:2 e sendo eficaz na prevenção da doença devida à bactéria *Neisseria*;
- uma sequência de aminoácidos compreendendo um fragmento de 8 ou mais aminoácidos consecutivos da ID SEQ NO: 2, em que o dito fragmento compreende um epítipo da ID SEQ NO:2;
- a sequência de aminoácidos ID SEQ NO:4;
- uma sequência de aminoácidos possuindo uma identidade de sequência de 50% ou mais relativamente à ID SEQ NO:4 e sendo eficaz na prevenção da doença devida à bactéria *Neisseria*;
- uma sequência de aminoácidos compreendendo um fragmento de 10 ou mais aminoácidos consecutivos da ID SEQ NO:4 em que o dito fragmento compreende um epítipo da ID SEQ NO:4;
- a sequência de aminoácidos ID SEQ NO:6;
- uma sequência de aminoácidos possuindo uma identidade de sequência de 50% ou mais relativamente à ID SEQ NO:6 e sendo eficaz na prevenção da doença devida à bactéria *Neisseria*;
- uma sequência de aminoácidos compreendendo um fragmento de 10 ou mais aminoácidos consecutivos da ID SEQ NO:6, em que os referidos fragmentos compreendem um epítipo de SEQ ID NO:6.

2. A proteína, de acordo com a Reivindicação N^o.1, caracterizada por possuir mais de 60% de identidade de sequência relativamente à SEQ ID NO:4.

3. A proteína, de acordo com a Reivindicação N°.2, caracterizada por possuir mais de 70% de identidade de sequência relativamente à SEQ ID NO:4.

4. A proteína, de acordo com a Reivindicação N°.3, caracterizada por possuir mais de 80% de identidade de sequência relativamente à SEQ ID NO:4.

5. A proteína, de acordo com a Reivindicação N°.4, caracterizada por possuir mais de 90% de identidade de sequência relativamente à SEQ ID NO:4.

6. A proteína, de acordo com a Reivindicação N°.5, caracterizada por possuir mais de 95% de identidade de sequência relativamente à SEQ ID NO:4.

7. A proteína, de acordo com a Reivindicação 6, caracterizada por possuir mais de 99% de identidade de sequência relativamente à SEQ ID NO:4.

8. A proteína, de acordo com a Reivindicação N°.1, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 12 ou mais aminoácidos consecutivos da ID SEQ NO:4.

9. A proteína, de acordo com a Reivindicação N°.8, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 14 ou mais aminoácidos consecutivos da ID SEQ NO:4.

10. A proteína, de acordo com a Reivindicação N°.9, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 16 ou mais aminoácidos consecutivos da ID SEQ NO:4.

11. A proteína, de acordo com a Reivindicação N°.10, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 18 ou mais aminoácidos consecutivos da ID SEQ NO:4.

12. A proteína, de acordo com a Reivindicação N°.11, caracterizada por compreender uma sequência de aminoácidos

compreendendo um fragmento de 20 ou mais aminoácidos consecutivos da ID SEQ NO:4.

13. A proteína, de acordo com a Reivindicação N°.1, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 10 ou mais aminoácidos consecutivos da ID SEQ NO:2.

14. A proteína, de acordo com a Reivindicação N°.13, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 12 ou mais aminoácidos consecutivos da ID SEQ NO:2.

15. A proteína, de acordo com a Reivindicação N°.14, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 14 ou mais aminoácidos consecutivos da ID SEQ NO:2.

16. A proteína, de acordo com a Reivindicação N°.15, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 16 ou mais aminoácidos consecutivos da ID SEQ NO:2.

17. A proteína, de acordo com a Reivindicação N°.16, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 18 ou mais aminoácidos consecutivos da ID SEQ NO:2.

18. A proteína, de acordo com a Reivindicação N°.17, caracterizada por compreender uma sequência de aminoácidos compreendendo um fragmento de 20 ou mais aminoácidos consecutivos da ID SEQ NO:2.

19. A proteína, de acordo com qualquer uma das Reivindicações anteriores, caracterizada por a proteína ser um antigénio de Neisseria.

20. Um anticorpo, caracterizado por se ligar especificamente a uma sequência de aminoácidos seleccionada a partir do grupo composto por: ID SEQ NO:2, ID SEQ NO:4 e ID SEQ NO:6.

21. O anticorpo, de acordo com a Reivindicação N°.20, caracterizado por corresponder a um anticorpo monoclonal.

22. Uma molécula de ácido nucleico, caracterizada por codificar para uma proteína, de acordo com qualquer uma das Reivindicações N°.1 a N°.19.

23. Uma molécula de ácido nucleico, de acordo com a Reivindicação N°.22, caracterizada por compreender uma sequência nucleotídica, seleccionada a partir do grupo composto por: ID SEQ NO:1, ID SEQ NO:3 e ID SEQ NO:5.

24. Uma molécula de ácido nucleico que codifica para uma proteína eficaz na prevenção da doença devida à bactéria *Neisseria*, caracterizada por compreender (i) um fragmento de 12 ou mais nucleótidos consecutivos da ID SEQ NO:1 codificando para um epítipo da ID SEQ NO:2; (ii) um fragmento de 20 ou mais nucleótidos consecutivos da ID SEQ ID NO:3 codificando para um epítipo da ID SEQ NO:4; ou (iii) um fragmento de 20 ou mais nucleótidos consecutivos da ID SEQ NO:5 codificando para um epítipo da ID SEQ NO:6.

25. A molécula de ácido nucleico, de acordo com a Reivindicação N°.24, caracterizada por compreender um fragmento de 25 ou mais nucleótidos consecutivos da ID SEQ NO:3.

26. A molécula de ácido nucleico, de acordo com a Reivindicação N°.25, caracterizada por compreender um fragmento de 30 ou mais nucleótidos consecutivos da ID SEQ NO:3.

27. A molécula de ácido nucleico, de acordo com a Reivindicação N°.26, caracterizada por compreender um fragmento de 35 ou mais nucleótidos consecutivos da ID SEQ NO:3.

28. A molécula de ácido nucleico, de acordo com a Reivindicação N°.27, caracterizada por compreender um

fragmento de 40 ou mais nucleótidos consecutivos da ID SEQ NO:3.

29. Uma molécula de ácido nucleico, caracterizada por compreender uma sequência nucleotídica complementar a uma molécula de ácido nucleico, de acordo com qualquer uma das Reivindicações N°.22 a N°.28.

30. Uma composição, caracterizada por compreender uma proteína, uma molécula de ácido nucleico ou um anticorpo, de acordo com qualquer uma das Reivindicações anteriores.

31. Uma composição, de acordo com a Reivindicação N°.30, caracterizada por corresponder a uma composição imunogénica.

32. Uma composição, de acordo com a Reivindicação N°.30, caracterizada por corresponder a uma composição de vacina ou uma composição para diagnóstico.

33. Uma composição, de acordo com a Reivindicação N°.30, caracterizada por corresponder a uma composição farmacêutica.

34. O uso de uma proteína, de acordo com qualquer uma das Reivindicações N°.1 a N°.19, de uma molécula de ácido nucleico, de acordo com qualquer uma das Reivindicações N°.22 a N°.29, ou de um anticorpo, de acordo com as Reivindicações N°.20 ou N°.21, no fabrico de um medicamento para a prevenção de uma infecção pela bactéria *Neisseria*.

35. O uso, de acordo com a Reivindicação N°.34, caracterizado por o medicamento se destinar à prevenção de uma infecção por *N.meningitidis*.

36. O uso, de acordo com a Reivindicação N°.35, caracterizado por o medicamento se destinar à prevenção de uma infecção por *N.meningitidis* B.

37. A composição, de acordo com qualquer uma das Reivindicações N°.31 a N°.33, ou o uso de qualquer uma das Reivindicações N°.34 a N°.36, caracterizada por a composição ou medicamento incluir um adjuvante.

38. Um vector caracterizado por compreender a sequência nucleotídica de uma molécula de ácido nucleico, de acordo com qualquer uma das Reivindicações N°.22 a N°.29.

39. Uma célula hospedeira, caracterizada por ser transformada com o vector da Reivindicação N°.38.

40. Um processo para produção da proteína, de acordo com qualquer uma das Reivindicações N°.1 a N°.19, caracterizado por compreender o passo de cultivo de uma célula hospedeira, de acordo com a Reivindicação N°.39, sob condições que induzem a expressão de proteínas.

Lisboa, 23 de Janeiro de 2008