

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

(22) Data de pedido: 2011.05.04	(73) Titular(es): GLAXOSMITHKLINE BIOLOGICALS SA	
(30) Prioridade(s): 2010.05.06 US 332170 P	RUE DE L'INSTITUT, 89 1330 RIXENSART	BE
(43) Data de publicação do pedido: 2013.03.13	(72) Inventor(es):	
(45) Data e BPI da concessão: 2017.11.22	MICHAEL KOWARIK	CH
026/2018	MICHAEL WACKER	CH
	MICHAEL WETTER	CH
	(74) Mandatário:	
	ANTÓNIO INFANTE DA CÂMARA TRIGUEIROS DE ARAGÃO	PT
	RUA DO PATROCÍNIO, Nº 94 1399-019 LISBOA	

(54) Epígrafe: **VACINAS DE BIOCONJUGADO DE BACTÉRIAS GRAM-POSITIVAS CAPSULARES**

(57) Resumo:

UMA FORMA DE REALIZAÇÃO DA PRESENTE INVENÇÃO É DIRIGIDA A UMA NOVA VACINA DE BIOCONJUGADO DE S. AUREUS. MAIS GENERICAMENTE, A INVENÇÃO É DIRIGIDA A VACINAS DE BIOCONJUGADOS GRAM-POSITIVOS E OUTROS BIOCONJUGADOS, COMPREENDENDO: UM VEÍCULO PROTEICO COMPREENDENDO UMA SEQUÊNCIA CONSENSO DE ÁCIDOS NUCLEICOS INSERIDA, PELO MENOS, UM POLISSACÁRIDO, TAL COMO UM POLISSACÁRIDO GRAM-POSITIVO CAPSULAR LIGADO À SEQUÊNCIA CONSENSO; E, OPCIONALMENTE, UM ADJUVANTE OU VEÍCULO FARMACEUTICAMENTE ACEITÁVEL. NUM ASPETO ADICIONAL, A PRESENTE INVENÇÃO É DIRIGIDA A MÉTODOS DE PRODUÇÃO DE VACINAS DE BIOCONJUGADOS GRAM-POSITIVOS E OUTROS BIOCONJUGADOS. NOUTRO ASPETO, É PROPORCIONADA UMA PROTEÍNA N-GLICOSILADA QUE COMPREENDE UM OU MAIS POLISSACÁRIDOS, TAL COMO POLISSACÁRIDOS GRAM-POSITIVOS. A PRESENTE INVENÇÃO É ADICIONALMENTE DIRIGIDA A ORGANISMOS PROCARIÓTICOS MANIPULADOS, COMPREENDENDO SEQUÊNCIAS NUCLEOTÍDICAS QUE CODIFICA UMA GLICOSILTRANSFERASE DE UM PRIMEIRO ORGANISMO PROCARIÓTICO E UMA GLICOSILTRANSFERASE DE UM SEGUNDO ORGANISMO PROCARIÓTICO. A INVENÇÃO INCLUI AINDA PLASMÍDEOS E CÉLULAS PROCARIÓTICAS TRANSFORMADOS COM PLASMÍDEOS QUE CODIFICAM POLISSACÁRIDOS E ENZIMAS QUE PRODUZEM UMA PROTEÍNA N-GLICOSILADA E/OU VACINA DE BIOCONJUGADO. ALÉM DISSO, A INVENÇÃO É DIRIGIDA A MÉTODOS DE INDUÇÃO DE UMA RESPOSTA IMUNITÁRIA NUM MAMÍFERO, COMPREENDENDO A ADMINISTRAÇÃO DAS REFERIDAS VACINAS DE BIOCONJUGADO.

RESUMO

"VACINAS DE BIOCONJUGADO DE BACTÉRIAS GRAM-POSITIVAS CAPSULARES"

Uma forma de realização da presente invenção é dirigida a uma nova vacina de bioconjugado de *S. aureus*. Mais genericamente, a invenção é dirigida a vacinas de bioconjugados Gram-positivos e outros bioconjugados, compreendendo: um veículo proteico compreendendo uma sequência consenso de ácidos nucleicos inserida, pelo menos, um polissacárido, tal como um polissacárido Gram-positivo capsular ligado à sequência consenso; e, opcionalmente, um adjuvante ou veículo farmacêuticamente aceitável. Num aspeto adicional, a presente invenção é dirigida a métodos de produção de vacinas de bioconjugados Gram-positivos e outros bioconjugados. Noutro aspeto, é proporcionada uma proteína N-glicosilada que compreende um ou mais polissacáridos, tal como polissacáridos Gram-positivos. A presente invenção é adicionalmente dirigida a organismos procarióticos manipulados, compreendendo sequências nucleotídicas que codifica uma glicosiltransferase de um primeiro organismo procariótico e uma glicosiltransferase de um segundo organismo procariótico. A invenção inclui ainda plasmídeos e células procarióticas transformados com plasmídeos que codificam polissacáridos e enzimas que produzem uma proteína N-glicosilada e/ou vacina de bioconjugado. Além disso, a invenção é dirigida a métodos de indução de uma resposta imunitária num mamífero, compreendendo a administração das referidas vacinas de bioconjugado.

DESCRIÇÃO

"VACINAS DE BIOCONJUGADO DE BACTÉRIAS GRAM-POSITIVAS CAPSULARES"

REFERÊNCIA CRUZADA A PEDIDOS RELACIONADOS

Este pedido reivindica o benefício ao abrigo do artigo 35 U.S.C. § 119(e) do Pedido de Patente Provisório U.S. N° 61/332170, apresentado a 6 de maio de 2010.

DECLARAÇÃO A RESPEITO DE INVESTIGAÇÃO PATROCINADA FEDERALMENTE

Esta invenção foi realizada com apoio governamental ao abrigo da subvenção 1R01AI088754-2 concedida pelos Institutos Nacionais de Saúde. O governo possui determinados direitos na invenção.

LISTAGEM DE SEQUÊNCIAS

O presente pedido contém uma Listagem de Sequências que foi submetida no formato ASCII via EFS-Web e é aqui incorporada por referência na sua totalidade. A referida cópia ASCII, criada em 2 de maio de 2011, é denominada 031229US.txt e tem 206590 bytes de tamanho.

ANTECEDENTES DA INVENÇÃO

As vacinas foram uma das grandes invenções de saúde pública da medicina moderna e salvaram milhões de vidas. As imunizações deram provas de serem um meio ideal para prevenir e controlar as infecções. Todos os anos, as vacinas previnem até 3 milhões de mortes e 750000 crianças são protegidas da incapacidade. (Global Alliance for Vaccines and Immunization - Comunicados de Imprensa (11 de março de 2006) em www.gavialliance.org/media_centre/press_releases/2006_03_09_en_pr_queenrania_delhi.php). Em 1999, o CDC declarou as imunizações a conquista número um em saúde pública do século XX (Ten great public health achievements-United States, 1900-1999. *MMWR Morb Mortal Wkly Rep* 48:241-3 (2 de abril de 1999)). Algumas bactérias, como aquelas que causam tétano ou difteria, produzem uma toxina que é em grande parte responsável pela doença. Esta toxina pode ser utilizada numa forma destoxificada como vacina. Contudo, para a maior parte das bactérias não existe uma toxina única que possa ser utilizada para desenvolver uma vacina.

Entre as vacinas mais bem-sucedidas estão os polissacáridos de superfície de patógenos bacterianos, como *Haemophilus influenzae*, *Neisseria meningitidis* e *Streptococcus pneumoniae*, conjugados a proteínas veículo. Estas bactérias estão rodeadas por uma cápsula, a qual promove a virulência microbiana e a resistência à neutralização fagocítica, assim como evita a sua desidratação.

Os polissacáridos bacterianos podem deduzir uma resposta imunitária duradoura em humanos se forem ligados a um veículo proteico que contenha epitopos de célula T. Este conceito foi elaborado há 80 anos (Avery, O. T. e W. F. Goebel. 1929. Estudos

químioimunológicos sobre proteínas de hidratos de carbono conjugadas. II Especificidade imunológica de proteínas de açúcar sintéticas. *J. Exp. Med.* 50:521-533), e provado mais tarde para o polissacárido do *Haemophilus influenza* tipo B (HIB) ligado ao veículo proteico toxina da difteria (Anderson, P. 1983. Respostas de anticorpo à toxina da difteria de *Haemophilus influenzae* tipo b e induzidas por conjugados de oligossacáridos de cápsula do tipo b com a proteína não tóxica CRM197. *Infect Immun* 39:233-8; Schneerson, R., O. Barrera, A. Sutton e J. B. Robbins. 1980. Preparação, caracterização e imunogenicidade de conjugados de polissacárido de *Haemophilus influenzae* tipo b com proteína. *J Exp Med* 152:361-76). Este glicoconjugado também foi a primeira vacina conjugada a ser autorizada nos EUA em 1987 e introduzida no programa de imunização infantil dos EUA pouco depois. Além do HIB, as vacinas conjugadas foram utilizadas com êxito contra os patógenos humanos encapsulados *N. meningitidis* e *S. pneumoniae*. A utilização de rotina destas vacinas resultou em colonização nasofaríngea, assim como infecção, diminuídas. Atualmente, aproximadamente ~25% de todo o mercado de vacinas global compreende vacinas conjugadas.

As bactérias gram-positivas possuem uma membrana celular que está rodeada por polissacáridos capsulares. O *Staphylococcus* é uma dessas bactérias Gram-positivas.

O *Staphylococcus aureus* causa infecção. O *S. aureus* é um patógeno bacteriano oportunista, responsável por um espectro diverso de doenças humanas. Embora o *S. aureus* possa colonizar superfícies de mucosa de humanos normais, também é uma causa muito importante de infecções de feridas e tem o potencial invasivo de induzir infecções graves, incluindo osteomielite, endocardite e bacteriemia, com complicações metastáticas (Lowy,

F. D. 1998. Infecções por *Staphylococcus aureus*. *New Engl J Med* 339:520-32). O *S. aureus* é um dos agentes mais comuns implicados na pneumonia associada a ventilador e é uma causa importante e emergente de pneumonia adquirida em comunidade, afetando adultos anteriormente saudáveis e crianças desprovidas de fatores de predisposição de risco (Kollef, M. H., A. Shorr, Y. P. Tabak, V. Gupta, L. Z. Liu e R. S. Johannes. 2005. Epidemiologia e resultados de pneumonia associada aos cuidados de saúde: resultados de uma base de dados de grandes dimensões dos EUA de pneumonia positiva para cultura. *Chest* 128:3854-62; Shorr, A. F. 2007. Epidemiologia e impacto económico de *Staphylococcus aureus* resistente à meticilina: revisão e análise da literatura. *Pharmacoeconomics* 25:751-68).

O *S. aureus* é a segunda causa mais comum de bacteriemia nosocomial e as estirpes de *S. aureus* resistentes à meticilina (MRSA) contribuem para mais de 50% de todas as infeções nas unidades de cuidados intensivos nos EUA. As infeções por *S. aureus* no hospital e na comunidade estão a aumentar. As estirpes de MRSA foram isoladas de 2% das infeções estafilocócicas em 1974 e de 63% das infeções estafilocócicas em 2004. Muitas das estirpes de MRSA nosocomiais são multirresistentes e até as estirpes sensíveis à meticilina podem ser mortais. Um relato recente utilizando verificação de caso ativo baseado em população revelou que ocorreram 94360 infeções por MRSA invasivo nos EUA em 2005 e que a maioria destas (58%) ocorreu fora do hospital (Klevens, R. M., M. A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L. H. Harrison, R. Lynfield, G. Dumyati, J. M. Townes, A. S. Craig, E. R. Zell, G. E. Fosheim, L. K. McDougal, R. B. Carey e S. K. Fridkin. 2007. Infecções por *Staphylococcus aureus* invasivo resistente à meticilina nos Estados Unidos. *JAMA* 298:1763-71). Nesta análise,

morreram mais Americanos de MRSA (>18000 mortes) em 2005 do que de SIDA.

O *S. aureus* USA100, também conhecido pelo clone de Nova Iorque/Japão, é uma estirpe de MRSA que representa a estirpe de MRSA adquirida em hospital predominante nos EUA (McDougal, L. K., C. D. Steward, G. E. Killgore, J. M. Chaitram, S. K. McAllister e F. C. Tenover. 2003. Tipagem de eletroforese em gel de campo pulsado de isolados de *Staphylococcus aureus* resistentes à oxacilina dos Estados Unidos: estabelecimento de uma base de dados nacional. *J Clin Microbiol* 41:5113-20).

As análises epidemiológicas indicam que, só nos EUA, o *S. aureus* causa, aproximadamente, 2 milhões de infecções clínicas todos os anos (Fridkin, S. K., J. C. Hageman, M. Morrison, L. T. Sanza, K. Como-Sabetti, J. A. Jernigan, K. Harriman, L. H. Harrison, R. Lynfield e M. M. Farley. 2005. Doença de *Staphylococcus aureus* resistente à meticilina em três comunidades. *N Engl J Med* 352:1436-44; King, M. D., B. J. Humphrey, Y. F. Wang, E. V. Kourbatova, S. M. Ray e H. M. Blumberg. 2006. Surgimento do clone de *Staphylococcus aureus* USA 300 resistente à meticilina adquirido em comunidade como a causa predominante de infecções de pele e de tecidos moles. *Ann Intern Med* 144:309-17; Klevens, R. M., M. A. Morrison, J. Nadle, S. Petit, K. Gershman, S. Ray, L. H. Harrison, R. Lynfield, G. Dumyati, J. M. Townes, A. S. Craig, E. R. Zell, G. E. Fosheim, L. K. McDougal, R. B. Carey, S. K. Fridkin e M. I. para a vigilância do Núcleo Bacteriano Ativo. 2007. Infecções por *Staphylococcus aureus* resistente à meticilina invasivas nos Estados Unidos. *JAMA* 298:1763-1771). As infecções por *S. aureus* estão não apenas a aumentar em número, mas a resistência do *S. aureus* a antibióticos também está a aumentar. O MRSA é

responsável por 40%-60% das infecções por *S. aureus* nosocomiais nos EUA e muitas destas estirpes são multirresistentes. Conhecido como uma fonte muito importante de infecções nosocomiais, o *S. aureus* adotou recentemente um novo papel na medida em que causa um número crescente de infecções adquiridas em comunidade, em pessoas não hospitalizadas, sem fatores de predisposição de risco. As estirpes virulentas de MRSA associadas à comunidade (CA-MRSA) estão a tornar-se mais predominantes por todos os EUA e Europa e a sua disseminação foi observada globalmente (Baggett, H. C., T. W. Hennessy, K. Rudolph, D. Bruden, A. Reasonover, A. Parkinson, R. Sparks, R. M. Donlan, P. Martinez, K. Mongkolrattanothai e J. C. Butler. 2004. *Staphylococcus aureus* resistente à meticilina de começo na comunidade associado com a utilização antibiótica e a citotoxina leucocidina de Panton-Valentine durante um surto de furunculose no Alasca rural. *J Infect Dis* 189:1565-73; Gilbert, M., J. MacDonald, D. Gregson, J. Siushansian, K. Zhang, S. Elsayed, K. Laupland, T. Louie, K. Hope, M. Mulvey, J. Gillespie, D. Nielsen, V. Wheeler, M. Louie, A. Honish, G. Keays e J. Conly. 2006. Surto em Alberta de *Staphylococcus aureus* resistente à meticilina adquirido em comunidade (USA300) em pessoas com um historial de uso de drogas, privação de habitação ou reclusão. *Canad Med Assoc J* 175:149-54; Kazakova, S. V., J. C. Hageman, M. Matava, A. Srinivasan, L. Phelan, B. Garfinkel, T. Boo, S. McAllister, J. Anderson, B. Jensen, D. Dodson, D. Lonsway, L. K. McDougal, M. Arduino, V. J. Fraser, G. Killgore, F. C. Tenover, S. Cody e D. B. Jernigan. 2005. Um clone de *Staphylococcus aureus* resistente à meticilina entre jogadores de futebol profissionais. *N Engl J Med* 352:468-75).

A resistência do *S. aureus* à meticilina tornou-se não apenas mais comum, mas foram referidos numerosos isolados com

reduzida suscetibilidade à vancomicina. Nos EUA, foram isolados sete isolados clínicos de *S. aureus* que transportam *vanA* e são totalmente resistentes à vancomicina. Estes isolados também são resistentes à meticilina (Chang, S., D. M. Sievert, J. C. Hageman, M. L. Boulton, F. C. Tenover, F. P. Downes, S. Shah, J. T. Rudrik, G. R. Pupp, W. J. Brown, D. Cardo e S. K. Fridkin. 2003. Infecção com *Staphylococcus aureus* resistente à vancomicina contendo o gene de resistência *vanA*. *New Engl J Med* 348:1342-7). Em virtude de o *S. aureus* nem sempre poder ser controlado pelos antibióticos e os isolados de MRSA estarem a tornar-se crescentemente predominantes na comunidade, estratégias adicionais de controle, tal como uma vacina, são extremamente necessárias.

Os polissacáridos capsulares de *S. aureus* estão envolvidos na infecção. Muitos fatores de virulência contribuem para a patogénese das infeções estafilocócicas, incluindo adesões associadas à superfície, exoproteínas e toxinas segregadas e fatores de evasão imunitária (Foster, T. J. 2005. Evasão imunitária por estafilococos. *Nature Reviews Microbiology* 3:948-58). Como muitos patogénios bacterianos invasivos, o *S. aureus* produz um polissacárido capsular (CP) (FIG. 4) que melhora a sua resistência à eliminação pelas defesas imunitárias inatas do hospedeiro. Muitos isolados clínicos de *S. aureus* são encapsulados e as estirpes de serotipo 5 e 8 predominam (Arbeit, R. D., W. W. Karakawa, W. F. Vann e J. B. Robbins. 1984. Predominância de dois tipos de polissacáridos capsulares recentemente descritos, entre isolados clínicos de *Staphylococcus aureus*. *Diagn Microbiol Infect Dis* 2:85-91). Os polissacáridos capsulares de tipo 5 (CP5) e tipo 8 (CP8) têm unidades de repetição de trissacáridos semelhantes constituídas por ácido N-acetilmanosaminurónico (ManNAcA), N-acetil-L-

fucosamina (L-FucNAc) e N-acetil-D-fucosamina (D-FucNAc) (Jones, C. 2005. Estruturas revistas para os polissacáridos capsulares de *Staphylococcus aureus* tipos 5 e 8, componentes de novas vacinas de glicoconjugados. *Carbohydr Res* 340:1097-106). Os CP5 e CP8 são serologicamente distintos e tal pode ser atribuído a diferenças nas ligações entre os açúcares e nos sítios de O-acetilação (FIG. 4).

Estudos anteriores correlacionaram a produção de cápsula do *S. aureus* com a resistência à absorção e neutralização fagocítica *in vitro* (Fattom, A., R. Schneerson, S. C. Szu, W. F. Vann, J. Shiloach, W. W. Karakawa e J. B. Robbins. 1990. Síntese e propriedades imunológicas em murganhos de vacinas compostas por polissacáridos capsulares tipo 5 e tipo 8 de *Staphylococcus aureus* conjugados a exotoxina A de *Pseudomonas aeruginosa*. *Infect Immun* 58:2367-74; Thakker, M., J.-S. Park, V. Carey e J. C. Lee. 1998. O polissacárido capsular de serotipo 5 de *Staphylococcus aureus* é antifagocítico e melhora a virulência bacteriana num modelo de bacteriemia murina. *Infect Immun* 66:5183-5189; Watts, A., D. Ke, Q. Wang, A. Pillay, A. Nicholson-Weller e J. C. Lee. 2005. Estirpes de *Staphylococcus aureus* que expressam polissacáridos capsulares de serotipo 5 ou serotipo 8 diferem em virulência. *Infect Immun* 73:3502-11). Os neutrófilos humanos fagocitam mutantes negativos para cápsula na presença de soro não imune com atividade de complemento, enquanto os isolados encapsulados requerem anticorpos específicos de cápsula e complemento para neutralização opsonofagocítica ótima (Bhasin, N., A. Albus, F. Michon, P. J. Livolsi, J.-S. Park e J. C. Lee. 1998. Identificação de um gene essencial para a O-acetilação do polissacárido capsular tipo 5 de *Staphylococcus aureus*. *Mol Microbiol* 27:9-21; Thakker, M., J.-S. Park, V. Carey e J. C. Lee. 1998. O polissacárido capsular

de serotipo 5 de *Staphylococcus aureus* é antifagocítico e melhora a virulência bacteriana num modelo de bacteriemia murina. *Infect Immun* 66:5183-5189; Watts, A., D. Ke, Q. Wang, A. Pillay, A. Nicholson-Weller e J. C. Lee. 2005. Estirpes de *Staphylococcus aureus* que expressam polissacáridos capsulares de serotipo 5 ou serotipo 8 diferem em virulência. *Infect Immun* 73:3502-11). Nilsson et al. (Nilsson, I.-M., J. C. Lee, T. Bremell, C. Ryden e A. Tarkowski. 1997. O papel da expressão da microcápsula de polissacárido estafilocócica em septicemia e artrite séptica. *Infect Immun* 65:4216-4221) referiram que macrófagos peritoneais de murganhos fagocitaram números significativamente maiores de um mutante negativo para CP5, em comparação com a estirpe parental Reynolds. Uma vez fagocitada, a estirpe positiva para CP5 sobreviveu intracelularmente em maior extensão do que a estirpe mutante. Cunnion et al. (Cunnion, K. M., J. C. Lee e M. M. Frank. 2001. Produção de cápsula e fase de crescimento influenciam a ligação do complemento a *Staphylococcus aureus*. *Infect Immun* 69:6796-6803) compararam a opsonização de estirpes isogénicas de *S. aureus* e demonstraram que a estirpe positiva para CP5 ligou-se 42% menos ao complemento (C') sérico do que o mutante acapsular.

O desenvolvimento de vacinas de *S. aureus* tem, convencionalmente, envolvido a cápsula como um alvo. A concepção de vacinas para proteção contra doença estafilocócica é complicada pelas manifestações proteanas e complexidade clínica de infeções por *S. aureus* em humanos. Foram investigados muitos candidatos vacinais de *S. aureus* em modelos animais de infeção, mas foi referido que apenas dois regimes de imunização completaram ensaios clínicos de fase III (Schaffer, A. C. e J. C. Lee. 2008. Vacinação e imunização passiva contra *Staphylococcus aureus*. *Int J Antimicrob Agents* 32 Suppl

1:S71-8). A primeira vacina é baseada nos dois polissacáridos capsulares (CP) (FIG. 4) que são mais predominantes entre estirpes clínicas de *S. aureus*. Fattom et al. (Fattom, A.R. Schneerson, S. C. Szu, W. F. Vann, J. Shiloach, W. W. Karakawa e J. B. Robbins. 1990. Síntese e propriedades imunológicas em murganhos de vacinas compostas por polissacáridos capsulares tipo 5 e tipo 8 de *Staphylococcus aureus* conjugados a exotoxina de *Pseudomonas aeruginosa*. *Infect Immun* 58:2367-74) conjugaram os polissacáridos de serotipo 5 (CP5) e serotipo 8 (CP8) à exoproteína A não tóxica recombinante de *P. aeruginosa* (rEPA). As vacinas conjugadas foram imunogênicas em murganhos e humanos e induziram anticorpos opsonícos que mostraram eficácia a protegerem roedores de letalidade e de infecção estafilocócica não letal (Fattom, A.R. Schneerson, S. C. Szu, W. F. Vann, J. Shiloach, W. W. Karakawa e J. B. Robbins. 1990. Síntese e propriedades imunológicas em murganhos de vacinas compostas por polissacáridos capsulares tipo 5 e tipo 8 de *Staphylococcus aureus* conjugados a exotoxina de *Pseudomonas aeruginosa*. *Infect Immun* 58: 2367-74; Fattom, A., R. Schneerson, D. C. Watson, W. W. Karakawa, D. Fitzgerald, I. Pastan, X. Li, J. Shiloach, D. A. Bryla e J. B. Robbins. 1993. Avaliação laboratorial e clínica de vacinas conjugadas compostas por polissacáridos capsulares tipo 5 e tipo 8 de *S. aureus* ligados a exoproteína A recombinante de *Pseudomonas aeruginosa*. *Infect Immun* 61:1023-32; Fattom, A. I., J. Sarwar, A. Ortiz e R. Naso. 1996. Uma vacina de polissacárido capsular (CP) de *Staphylococcus aureus* e anticorpos específicos de CP protege murganhos contra estímulo bacteriano. *Infect Immun* 64:1659-65; Lee, J. C., J. S. Park, S. E. Shepherd, V. Carey e A. Fattom. 1997. Eficácia protetora de anticorpos para o polissacárido capsular tipo 5 de *Staphylococcus aureus* num modelo modificado de endocardite em ratos. *Infect Immun* 65:4146-51). Estudos de imunização passiva indicaram que

anticorpos específicos para CP5 e CP8 reduziram significativamente a infecção num modelo murino de mastite de *S. aureus* (Tuchscherr, L. P., F. R. Buzzola, L. P. Alvarez, J. C. Lee e D. O. Sordelli. 2008. Anticorpos para polissacárido capsular e fator de aglutinação A previnem mastite e o surgimento de variantes não encapsulados e de colônias pequenas de *Staphylococcus aureus* em murganhos. *Infect Immun* 76:5738-44). Mostrou-se que a vacina de conjugado de CP5 e CP8 combinada era segura em humanos e deduzia anticorpos que mostraram atividade opsonofagocítica.

O desenvolvimento de vacinas de *S. aureus* também envolveu proteínas de superfície como um alvo. O segundo ensaio clínico vacinal de *S. aureus* foi baseado na eficácia protetora de anticorpos para adesões estafilocócicas na prevenção de infecções estafilocócicas. O fator de aglutinação A de *S. aureus* é uma proteína ancorada à parede celular que é expressa à superfície, medeia a aderência estafilocócica ao fibrinogénio (Foster, T. J. e M. Hook. 1998. Adesões de proteínas de superfície de *Staphylococcus aureus*. *Trends Microbiol* 6:484-8) e promove a conjugação de *S. aureus* a superfícies de biomateriais (Vaudaux, P. E., P. Francois, R. A. Proctor, D. McDevitt, T. J. Foster, R. M. Albrecht, D. P. Lew, H. Wabers e S. L. Cooper. 1995. Utilização de mutantes de adesão defeituosa de *Staphylococcus aureus* para definir o papel de proteínas plasmáticas específicas na promoção da adesão bacteriana a derivações arteriovenosas caninas. *Infection & Immunity* 63:585-90), coágulos de sangue e superfícies endoteliais danificadas (Moreillon, P., J. M. Entenza, P. Francioli, D. McDevitt, T. J. Foster, P. Francois e P. Vaudaux. 1995. Papel da coagulase e fator de aglutinação de *Staphylococcus aureus* na patogénese de endocardite experimental. *Infection & Immunity* 63:4738-43). O domínio de ligação ao

fibrinogénio de ClfA está localizado na região A da proteína de comprimento completo (McDevitt, D., P. Francois, P. Vaudaux e T. J. Foster. 1995. Identificação do domínio de ligação de ligando do recetor de fibrinogénio localizado à superfície (fator de aglutinação) de *Staphylococcus aureus*. *Molecular Microbiology* 16:895-907). O ClfA desempenha um papel importante na ligação de *S. aureus* a plaquetas, uma interação que é crítica em modelos animais de endocardite estafilocócica induzida por cateter (Sullam, P. M., A. S. Bayer, W. M. Foss e A. L. Cheung. 1996. A ligação plaquetária diminuída *in vitro* por *Staphylococcus aureus* está associada com virulência reduzida num modelo de coelho de endocardite infecciosa. *Infection & Immunity* 64:4915-21).

Nanra *et al.* referiram que anticorpos para ClfA induziram a neutralização opsonofagocítica de *S. aureus in vitro* (Nanra, J. S., Y. Timofeyeva, S. M. Buitrago, B. R. Sellman, D. A. Dilts, P. Fink, L. Nunez, M. Hagen, Y. V. Matsuka, T. Mininni, D. Zhu, V. Pavliak, B. A. Green, K. U. Jansen e A. S. Anderson. 2009. Expressão heterogénea *in vivo* de fator de aglutinação A e polissacárido capsular pelo *Staphylococcus aureus*: Implicações para a conceção vacinal. *Vaccine* 27:3276-80). Além disso, murganhos imunizados com uma forma recombinante da região de ligação A de ClfA mostraram reduções em artrite e letalidade induzidas por *S. aureus* (Josefsson, E., O. Hartford, L. O'Brien, J. M. Patti e T. Foster. 2001. Proteção contra artrite experimental de *Staphylococcus aureus* por vacinação com fator de aglutinação A, um novo determinante de virulência. *Journal of Infectious Diseases* 184:1572-80). Foram realizadas experiências de imunização passiva em coelhos aos quais foi dada uma preparação de imunoglobulina policlonal humana que continha níveis elevados de anticorpos específicos para ClfA (Vernachio, J., A. S. Bayer, T. Le, Y. L. Chai, B. Prater, A. Schneider, B.

Ames, P. Syribeys, J. Robbins, J. M. Patti, J. Vernachio, A. S. Bayer, T. Le, Y.-L. Chai, B. Prater, A. Schneider, B. Ames, P. Syribeys, J. Robbins e J. M. Patti. 2003. Imunoglobulina de fator de antiaglutinação A reduz a duração de bacteriemia de *Staphylococcus aureus* resistente a meticilina num modelo experimental de endocardite infecciosa. *Antimicrobial Agents & Chemotherapy* 47:3400-6). A terapia de combinação resultou em melhor eliminação bacteriana do sangue de coelhos com endocardite de *S. aureus* induzida por cateter do que o tratamento de vancomicina isoladamente. Além disso, a transferência passiva de anticorpos específicos de ClfA reduziu significativamente a infecção num modelo murino de mastite de *S. aureus* (Tuchscherr, L. P., F. R. Buzzola, L. P. Alvarez, J. C. Lee e D. O. Sordelli. 2008. Anticorpos para polissacárido capsular e fator de aglutinação A previnem mastite e o surgimento de variantes não encapsulados e de colônias pequenas de *Staphylococcus aureus* em murganhos. *Infect Immun* 76: 5738-44).

um ensaio clínico de fase III foi alegadamente concebido para proteger contra sépsis de início tardio em 2000 recém-nascidos prematuros de baixo peso à nascença. Os bebés receberam até quatro administrações de Veronate, uma preparação de imunoglobulina humana agregada de doadores com elevados títulos de anticorpo contra ClfA e SdrG. Não obstante os resultados promissores de um ensaio clínico de fase II semelhante, esta terapia profilática não resultou na redução na frequência de infeções estafilocócicas nos recém-nascidos (DeJonge, M., D. Burchfield, B. Bloom, M. Duenas, W. Walker, M. Polak, E. Jung, D. Millard, R. Schelonka, F. Eyal, A. Morris, B. Kapik, D. Roberson, K. Kesler, J. Patti e S. Hetherington. 2007. Ensaio clínico de segurança e eficácia de INH-A21 para a

prevenção de infecção da corrente sanguínea estafilocócica nosocomial em bebês prematuros. *J Pediatr* 151:260-5).

Mostrou-se que a glicosilação proteica ocorre mas, em organismos procarióticos, raramente acontece naturalmente. Por outro lado, a glicosilação proteica ligada a N é um processo essencial e conservado ocorrendo no retículo endoplasmático dos organismos eucarióticos. É importante para o enrolamento proteico, oligomerização, estabilidade, controlo de qualidade, separação e transporte de proteínas secretórias e membranares (Helenius, A. e Aebi, M. (2004). Papéis dos glicanos ligados a N no retículo endoplasmático. *Annu. Rev. Biochem.* 73, 1019-1049). A glicosilação proteica tem uma influência profundamente favorável na antigenicidade, na estabilidade e na meia-vida de uma proteína. Além disso, a glicosilação pode facilitar a purificação de proteínas por cromatografia, e. g., cromatografia de afinidade com ligandos de lectina ligados a uma fase sólida que interage com unidades glicosiladas da proteína. É, deste modo, prática estabelecida produzir muitas proteínas glicosiladas de modo recombinante em células eucarióticas, para proporcionar padrões de glicosilação, biologicamente e farmacologicamente úteis.

As vacinas conjugadas foram utilizadas com êxito para proteger contra as infecções bacterianas. Para resposta de memória protetora, é requerida a conjugação de um polissacárido antigénico a um veículo proteico, dado que os polissacáridos são antigénios independentes de células T. Os polissacáridos foram conjugados a transportadores proteicos por diferentes métodos químicos, utilizando grupos de reativos de ativação no polissacárido, assim como no veículo proteico. (Qian, F., Y. Wu, O. Muratova, H. Zhou, G. Dobrescu, P. Duggan, L. Lynn, G. Song,

Y. Zhang, K. Reiter, N. MacDonald, D. L. Narum, C. A. Long, L. H. Miller, A. Saul e G. E. Mullen. 2007. Conjugação de proteínas recombinantes à ExoProteína A de *Pseudomonas aeruginosa*: uma estratégia para melhoramento da imunogenicidade de candidatos vacinais maláricos. *Vaccine* 25:3923-3933; Pawlowski, A., G. Kallenius e S. B. Svenson. 2000. Preparação de vacinas de conjugados de polissacárido capsular pneumocócico com proteína utilizando novas tecnologias de fragmentação e conjugação. *Vaccine* 18:1873-1885; Robbins, J. B., J. Kubler-Kielb, E. Vinogradov, C. Mocca, V. Pozsgay, J. Shiloach e R. Schneerson. 2009. Síntese, caracterização e imunogenicidade em murganhos de conjugados de oligossacárido específico de O de *Shigella sonnei* com proteína nuclear. *Proc Natl Acad Sci U S A* 106:7974-7978).

As vacinas conjugadas podem ser administradas a crianças para as proteger contra infecções bacterianas e podem proporcionar uma resposta imunitária duradoura a adultos. Verificou-se que as construções da invenção produzem uma resposta de IgG em animais. Considera-se que o polissacárido (*i. e.*, resíduo de açúcar) desencadeia uma resposta imunitária de curto prazo que é específica de açúcar. De facto, o sistema imunitário humano produz uma forte resposta a estruturas de superfície de polissacáridos específicas de bactérias, tais como antigénios O e polissacáridos capsulares. Contudo, dado que a resposta imunitária aos polissacáridos é dependente de IgM, o sistema imunitário não desenvolve memória. O veículo proteico que transporta o polissacárido, contudo, desencadeia uma resposta de IgG que é dependente de célula T e que proporciona proteção duradoura, dado que o sistema imunitário desenvolve memória. Por esta razão, no desenvolvimento de uma vacina, é vantajoso desenvolvê-la como um conjugado de veículo proteico com polissacárido.

Os organismos procarióticos raramente produzem proteínas glicosiladas. Contudo, foi demonstrado que uma bactéria, o patógeno transmitido por alimentos *Campylobacter jejuni*, pode glicosilar as suas proteínas (Szymanski, et al. (1999). Evidência para um sistema de glicosilação proteica geral em *Campylobacter jejuni*. *Mol. Microbiol.* 32, 1022-1030). A maquinaria requerida para glicosilação é codificada por 12 genes que estão aglomerados no locus *pgl*. A perturbação da glicosilação afeta a invasão e a patogênese de *C. jejuni*, mas não é letal como na maioria dos organismos eucarióticos (Burda P. e M. Aebi, (1999). A via dolicol da glicosilação ligada a N. *Biochim Biophys Acta* 1426(2):239-57). Foi mostrado que o locus *pgl* é responsável pela glicosilação proteica ligada a N em *Campylobacter* e que é possível reconstituir a N-glicosilação de proteínas de *C. jejuni* por expressão recombinante do locus *pgl* e glicoproteína aceitadora em *E. coli* ao mesmo tempo (Wacker, M., D. Linton, P. G. Hitchen, M. Nita-Lazar, S. M. Haslam, S. J. North, M. Panico, H. R. Morris, A. Dell, B. W. Wren e M. Aebi. 2002. Glicosilação ligada a N em *C. jejuni* e sua transferência funcional para *E. coli*. *Science* 298:1790-3).

A via biossintética da glicosilação proteica ligada a N de *Campylobacter* tem semelhanças significativas com a via da biossíntese de polissacáridos em bactérias (Bugg, T. D. e P. E. Brandish. 1994. Do peptidoglicano às glicoproteínas: características comuns da biossíntese de oligossacáridos ligada a lípido. *FEMS Microbiol Lett* 119:255-62). Com base no conhecimento de que os polissacáridos antigénicos de bactérias e os oligossacáridos de *Campylobacter* são ambos sintetizados no lípido transportador, undecaprenilpirofosfato (UndPP), as duas vias foram combinados em *E. coli* (Feldman, M. F., M. Wacker, M.

Hernandez, P. G. Hitchen, C. L. Marolda, M. Kowarik, H. R. Morris, A. Dell, M. A. Valvano e M. Aebi. 2005. Manipulação de glicosilação proteica ligada a *N* com diversas estruturas de lipopolissacáridos de antígeno O em *Escherichia coli*. *Proc Natl Acad Sci U S A* 102:3016-21). Foi demonstrado que o PglB não tem uma especificidade estrita para o substrato de açúcar ligado a lípido. Os polissacáridos antigénicos agrupados em UndPP são capturados por PglB no periplasma e transferidos para um veículo proteico (Feldman, M. F., M. Wacker, M. Hernandez, P. G. Hitchen, C. L. Marolda, M. Kowarik, H. R. Morris, A. Dell, M. A. Valvano e M. Aebi. 2005. Manipulação de glicosilação proteica ligada a *N* com diversas estruturas de lipopolissacáridos de antígeno O em *Escherichia coli*. *Proc Natl Acad Sci U S A* 102:3016-21; Wacker, M., M. F. Feldman, N. Callewaert, M. Kowarik, B. R. Clarke, N. L. Pohl, M. Hernandez, E. D. Vines, M. A. Valvano, C. Whitfield e M. Aebi. 2006. A especificidade de substrato de oligossacariltransferase bacteriana (OTase) sugere um mecanismo de transferência comum para os sistemas bacterianos e eucarióticos. *Proc Natl Acad Sci U S A* 103:7088-93). Foi mostrado que o PglB de *Campylobacter* transfere uma gama variada de oligossacáridos ligados a UndPP se contiverem uma hexosamina *N*-acetilada no terminal de redução (Wacker *et al.* (2006)), permitindo a conjugação de um polissacárido antigénico a uma proteína de escolha através de uma ligação *N*-glicosídica. Embora tal possa proporcionar uma base teórica para a produção de vacinas conjugadas *in vivo*, necessitam de ser superados muitos desafios difíceis de modo a compreender-se esta possibilidade teórica.

Com base nesta descoberta anterior, de que o *C. jejuni* contém um sistema geral de glicosilação proteica ligada a *N*, a *E. coli* foi modificada para incluir a maquinaria de glicosilação

proteica ligada a *N* de *C. jejuni*. Deste modo, foram produzidas formas glicosiladas de proteínas nativas a *C. jejuni* num hospedeiro de *E. coli*. Foi ainda mostrado que este processo poderia ser utilizado para produzir proteínas glicosiladas de origens diferentes em hospedeiro de *E. coli* modificado para utilização como produtos vacinais (o documento WO 06/119987 e "GlycoVaxyn e hospital afiliado da Universidade de Harvard receberam uma subvenção de 3,4 milhões de dólares para o desenvolvimento de vacinas de *Staphylococcus aureus*". URL:<http://www.glycovaxyn.com/content/news/releases/10%2005%2004.pdf>). A produção por *E. coli* é vantajosa, porque podem ser produzidas culturas de grandes dimensões desses hospedeiros de *E. coli* modificados, as quais produzem grandes quantidades de vacina útil.

A utilização deste processo para produzir uma proteína glicosilada num hospedeiro de *E. coli* modificado, para utilização como um produto vacinal para *S. aureus*, encontra problemas que foram considerados intransponíveis. Em primeiro lugar, a *E. coli* é uma bactéria Gram-negativa e as suas vias de biossíntese de sacáridos diferem grandemente daquelas de uma bactéria Gram-positiva, tal como *S. aureus*, após o passo de polimerização. Além disso, seria inexequível manipular geneticamente a *E. coli* para produzir o polissacárido capsular de *S. aureus* diretamente consistente com tecnologias anteriores. Por exemplo, o *S. aureus* é um organismo Gram-positivo e a síntese da sua cápsula está associada com a estrutura do envelope celular e a construção do invólucro celular. A maquinaria biossintética de produção da cápsula está especificamente concebida para dispor o polissacárido capsular (PS) no exterior da célula e sua parede celular. Seria extremamente difícil, pelo menos pela razão de que seria

altamente intensivo em termos de recursos, produzir esta cápsula num organismo de *E. coli* modificado, porque o envelope celular de *E. coli* é construído numa forma fundamentalmente diferente. A maquinaria biossintética para o agrupamento da cápsula a partir do precursor de PS seria não funcional, devido ao ambiente diferente. Enquanto a cápsula de *S. aureus* deve transitar uma única membrana, na *E. coli* existe uma membrana adicional que necessita de ser atravessada para atingir a localização final de uma cápsula autêntica. Além disso, como a cápsula de *S. aureus* é muito grande, considerou-se ser inexequível preparar uma cápsula de grandes dimensões como a cápsula de *S. aureus* entre as duas membranas de *E. coli*.

O princípio de que as enzimas de organismos diferentes podem funcionar em conjunto foi mostrado anteriormente (e. g., Rubires, X., F. Saigi, N. Pique, N. Climent, S. Merino, S. Alberti, J. M. Tomas e M. Regue. 1997. Um gene (wbbL) de *Serratia marcescens* N28b (04) complementa a mutação rfb-50 de derivados de *Escherichia coli* K-12. *J. Bacteriol* 179(23): 7581-6). Contudo, considera-se que não foi anteriormente produzido num organismo Gram-negativo qualquer polissacárido de LPS modificado de um organismo Gram-positivo.

BREVE SUMÁRIO DA INVENÇÃO

A requerente identificou, surpreendentemente, uma nova vacina de bioconjugado de *S. aureus*. Esta nova vacina de *S. aureus* é baseada na nova e inesperada verificação de que um oligossacárido ou um polissacárido de um procariota que tem uma estirpe Gram pode glicosilar uma proteína num procariota hospedeiro que tem uma estirpe Gram diferente. Consequentemente,

é proporcionado um organismo procariótico hospedeiro Gram-negativo, em que o referido organismo procariótico hospedeiro é *E. coli*, compreendendo:

(i) uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-positiva, em que a referida bactéria Gram-positiva é *Staphylococcus aureus*;

(ii) uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-negativa, em que a referida bactéria Gram-negativa é *Pseudomonas aeruginosa*;

(iii) uma sequência nucleotídica que codifica uma proteína transportadora, em que a referida proteína transportadora compreende a sequência consenso de aminoácidos D/E-X-N-Z-S/T, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e

(iv) uma sequência nucleotídica que codifica uma oligossacariltransferase.

Características novas e inesperadas adicionais da divulgação incluem, sem limitação, as formas de realização apresentadas abaixo.

Mais genericamente, a presente divulgação é dirigida a uma vacina de bioconjugado, tal como uma vacina Gram-positiva, compreendendo um veículo proteico compreendendo uma sequência consenso de ácidos nucleicos inserida; pelo menos um oligossacárido ou polissacárido de uma bactéria, tal como uma

bactéria Gram-positiva, ligado à sequência consenso e, opcionalmente, um adjuvante. Além disso, a divulgação é dirigida a uma vacina de bactéria Gram-positiva, tal como uma vacina de *S. aureus*, ou vacina de outras bactérias, preparada por um sistema de glicosilação utilizando uma via biossintética de LPS modificada, a qual compreende a produção de um polissacárido capsular ou LPS modificado.

A presente invenção é adicionalmente dirigida a uma vacina de *Staphylococcus aureus*, compreendendo: uma proteína compreendendo uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e, pelo menos, um polissacárido de *Staphylococcus aureus* ligado à referida sequência consenso por ligação N-glicosídica; e, opcionalmente, um veículo ou adjuvante farmacêuticamente aceitável.

A presente é, além disso, dirigida a uma combinação de um polissacárido capsular modificado de *S. aureus* com um antígeno proteico do mesmo organismo por ligação N-glicosídica.

A invenção é, ainda, dirigida a um organismo procariótico hospedeiro Gram-negativo, em que o referido organismo procariótico hospedeiro é *E. coli*, compreendendo uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-positiva, em que a referida bactéria Gram-positiva é *Staphylococcus aureus*; uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-negativa, em que a referida bactéria Gram-negativa é *Pseudomonas aeruginosa*; uma sequência nucleotídica que codifica uma proteína transportadora, em que a referida proteína transportadora compreende a sequência consenso

de aminoácidos D/E-X-N-Z-S/T, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e uma sequência nucleotídica que codifica uma oligossacariltransferase. A divulgação é, adicionalmente, dirigida a um organismo procariótico hospedeiro manipulado, compreendendo uma sequência nucleotídica introduzida que codifica glicosiltransferases nativas apenas a um organismo procariótico Gram-positivo; uma sequência nucleotídica que codifica uma proteína; e uma sequência nucleotídica que codifica uma OTase.

A divulgação é, além disso, dirigida a métodos de produção de uma vacina de bioconjugado num organismo procariótico hospedeiro, compreendendo ácidos nucleicos que codificam uma ou mais glicosiltransferases de uma primeira espécie procariótica, tal como uma espécie Gram-positiva, por exemplo, *S. aureus*; uma ou mais glicosiltransferases de uma segunda espécie procariótica, uma proteína; e uma OTase. Além disso, a presente invenção é dirigida à produção de vacinas de bioconjugado através da produção, em bactérias Gram-negativas, de polissacáridos capsulares modificados, que podem ser transferidos para o núcleo de lípido A por WaaL e/ou ligados a um veículo de escolha pela OTase.

A divulgação é, ainda, dirigida a métodos de produção de proteínas glicosiladas num organismo procariótico hospedeiro, compreendendo sequência nucleotídica que codifica glicosiltransferases nativas a um primeiro organismo procariótico e também que codifica glicosiltransferases nativas a um segundo organismo procariótico que é diferente do primeiro organismo procariótico. A presente divulgação é, adicionalmente, dirigida à produção de proteínas N-glicosiladas com polissacáridos capsulares de bactérias Gram-positivas, que são

sintetizados por uma combinação de diferentes glicosiltransferases de diferentes organismos. A divulgação é, além disso, dirigida à produção de proteínas glicosiladas num organismo procariótico hospedeiro, compreendendo uma sequência nucleotídica introduzida que codifica glicosiltransferases nativas apenas a um organismo procariótico Gram-positivo.

A presente divulgação é, além disso, dirigida a plasmídeos, tal como plasmídeos compreendendo uma ou mais das SEQ ID N°: 2, SEQ ID N°: 3 e SEQ ID N°: 4. A divulgação também inclui plasmídeos compreendendo uma ou mais das SEQ ID N°: 6; SEQ ID N°: 7; SEQ ID N°: 8 e SEQ ID N°: 16. A divulgação também se refere a plasmídeos compreendendo uma ou mais das SEQ ID N°: 10; SEQ ID N°: 11; e SEQ ID N°: 12. Além disso, a divulgação é dirigida a plasmídeos compreendendo uma ou mais das SEQ ID N°: 13; SEQ ID N°: 15; SEQ ID N°: 15; SEQ ID N°: 17; SEQ ID N°: 18; SEQ ID N°: 19; SEQ ID N°: 20; SEQ ID N°: 21 e SEQ ID N°: 27.

A divulgação é adicionalmente dirigida a células bacterianas transformadas, tal como, por exemplo, células bacterianas transformadas com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 2; SEQ ID N°: 3; SEQ ID N°: 4; SEQ ID N°: 17; SEQ ID N°: 18; SEQ ID N°: 19 e SEQ ID N°: 20; SEQ ID N°: 21; e SEQ ID N°: 27. A presente divulgação é, ainda, dirigida a uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 5; SEQ ID N°: 8; SEQ ID N°: 9; SEQ ID N°: 10; SEQ ID N°: 11; SEQ ID N°: 12; SEQ ID N°: 13; SEQ ID N°: 14; SEQ ID N°: 15 e SEQ ID N°: 16.

A presente divulgação é ainda dirigida a um método de indução de uma resposta imunitária contra uma infecção causada

por bactérias Gram-positivas e outras bactérias num mamífero. Numa forma de realização, o método compreende a administração ao referido mamífero de uma quantidade eficaz de uma composição farmacêutica compreendendo: proteína compreendendo, pelo menos, uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e um ou mais oligossacáridos ou polissacáridos, os um ou mais oligossacáridos ou polissacáridos que são iguais ou diferentes de outro dos um ou mais oligossacáridos ou polissacáridos, de uma bactéria Gram-positiva ligada à referida sequência consenso.

Noutro aspeto, a divulgação caracteriza um método de identificação de um polissacárido alvo para utilização na glicosilação de uma proteína com o referido polissacárido alvo, no todo ou em parte. A referida proteína glicosilada compreendendo o polissacárido alvo pode ser utilizada, por exemplo, em composições vacinais. Numa forma de realização, o método de identificação de um polissacárido alvo inclui: identificação de uma bactéria Gram-positiva, tal como *S. aureus*, como um alvo; identificação de uma primeira unidade de repetição de um polissacárido produzido pela referida bactéria Gram-positiva compreendendo, pelo menos, três monómeros; identificação de um polissacárido produzido por uma bactéria de uma espécie Gram-negativa, compreendendo uma segunda unidade de repetição compreendendo dois dos mesmos monómeros que a referida primeira unidade de repetição.

A presente divulgação também é dirigida a um método para a modificação de uma bactéria de uma primeira espécie bacteriana, tal como uma espécie Gram-negativa. Numa forma de realização, o método inclui: identificação de uma primeira unidade de repetição de um polissacárido de uma espécie Gram-positiva, tal

como *S. aureus*, compreendendo três monómeros; identificação de um polissacárido produzido por uma bactéria de uma segunda espécie Gram-negativa, compreendendo outra unidade de repetição compreendendo dois dos mesmos monómeros da primeira unidade de repetição; inserção na referida bactéria de uma primeira espécie Gram-negativa de uma ou mais sequências nucleotídicas que codifica glicosiltransferases que agrupam um trissacárido, compreendendo: a) referida segunda unidade de repetição; e b) um monómero da referida primeira unidade de repetição não presente na referida segunda unidade de repetição; inserção de uma sequência nucleotídica que codifica uma proteína; e inserção de uma sequência nucleotídica que codifica uma OTase.

BREVE DESCRIÇÃO DOS DESENHOS

A **FIG. 1** representa uma via para a biossíntese de antígeno O dependente de *wzx/wzy*, exemplificada pela biossíntese do antígeno O de *P. aeruginosa* O11. Os nomes das proteínas putativamente responsáveis pelas reações apresentadas estão indicados acima ou abaixo das setas, incluindo uridina difosfato (UDP) e uridina monofosfato (UMP).

A **FIG. 2** representa uma via proposta para a biossíntese do polissacárido capsular de serotipo 5 (CP5) de *S. aureus* manipulado em *E. coli*. As enzimas proporcionadas pelo aglomerado de antígeno O de *P. aeruginosa* O11 estão indicadas como na FIG. 1. As enzimas de CP5 de *S. aureus* estão indicadas como Cap5 (comparar com a FIG. 6). WecB e WecC são enzimas de *E. coli* requeridas para a produção de UDP-ManNAcA. Outras proteínas e enzimas representadas incluem uridina difosfato (UDP), uridina monofosfato (UMP) e coenzima A (CoA).

A **FIG. 3** representa uma via proposta para a biossíntese do polissacárido capsular de serotipo 8 (CP8) de *S. aureus* manipulado. Os nomes dos genes estão indicados por setas (comparar com as FIG. 1, 2 e 6). UDP, UMP: uridina difosfato, uridina monofosfato. CoA: coenzima A.

A **FIG. 4** representa a sobreposição estrutural de Estruturas de Unidade de Repetição (RU) de antígeno O capsular de *S. aureus* e *P. aeruginosa*.

A **FIG. 5A** representa a análise de SDS-PAGE do alongamento da RU (unidade de repetição) do antígeno O de O11 incompleto por enzimas de *S. aureus*.

A **FIG. 5B** representa a imunodeteção do alongamento da RU do antígeno O de O11 incompleto por enzimas de *S. aureus*.

A **FIG. 6** representa uma estratégia, numa forma de realização da invenção, para a construção dos aglomerados dos genes O11/CP55 e O11/CP8 quiméricos.

A **FIG. 7A** representa LPS de CP5 polimerizado de uma forma de realização da invenção detetado em extratos lipídicos de *E. coli*.

A **FIG. 7B** representa LPS de CP8 polimerizado de uma forma de realização da invenção detetado em extratos lipídicos de *E. coli*.

A **FIG. 8A** representa a produção de LPS de CP5 recombinante de uma forma de realização da invenção analisada por SDS-PAGE e

corada por prata, em dependência de gene de resistência a antibiótico no plasmídeo pLAFR contendo o aglomerado quimérico em células W3110 Δ wecA.

A **FIG. 8B** representa a produção de LPS de CP5 recombinante de uma forma de realização da invenção analisada por SDS PAGE, corada por prata e imunodeteção, em dependência de gene de resistência a antibiótico no plasmídeo pLAFR contendo o aglomerado quimérico em células W3110 Δ wecA.

A **FIG. 9** representa a produção de LPS de CP5 recombinante de uma forma de realização da invenção analisada por SDS PAGE e por imunodeteção, em dependência de promotor na frente do aglomerado quimérico em células W3110 Δ wecA.

A **FIG. 10A** mostra os resultados da análise de HPLC de uma forma de realização da RU recombinante de CP5 da presente invenção produzida utilizando o aglomerado de CP5 quimérico (SEQ ID N°: 2).

A **FIG. 10B** mostra os resultados da análise de HPLC de uma forma de realização da RU recombinante de CP8 da presente invenção produzida utilizando o aglomerado de CP8 quimérico desprovido da polimerase *cap8I*.

A **FIG. 11A** mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP5 quimérico da presente invenção em *E. coli* eluindo a 37 minutos observado na FIG. 10A.

A **FIG. 11B** mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de

realização do aglomerado de CP5 quimérico da presente invenção em *E. coli* eluindo a 40 minutos observado na FIG. 10A.

A **FIG. 11C** mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP8 quimérico da presente invenção em *E. coli* eluindo a 32 minutos observado na FIG. 10B.

A **FIG. 11D** mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP8 quimérico da presente invenção em *E. coli* eluindo a 38 minutos observado na FIG. 10B.

A **FIG. 11E** mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP8 quimérico da presente invenção em *E. coli* eluindo a 45 minutos observado na FIG. 10B.

A **FIG. 11F** mostra os resultados da análise de HPLC de uma forma de realização da otimização da estrutura de glicano.

A **FIG. 11G** apresenta os resultados da análise de HPLC do repertório de glicano de CP5 completo presente em UndPP em células *E. coli* numa forma de realização da presente invenção.

A **FIG. 11H** apresenta os resultados da análise de HPLC de glicanos de CP5 desacetilados e homogeneidade de RU numa forma de realização da invenção.

A **FIG. 11I** proporciona os resultados da análise de HPLC do repertório de glicano de CP8 presente em UndPP em células *E. coli* numa forma de realização da presente invenção.

A **FIG. 11J** mostra resultados de HPLC, numa forma de realização da presente invenção, da desacetilação de glicanos de CP8 e homogeneidade de RU.

A **FIG. 11K** apresenta resultados de HPLC mostrando redução na polimerização de RU e aumento em LLO induzido por co-expressão de wzz07 com o aglomerado quimérico de CP8 numa forma de realização da presente invenção.

A **FIG. 12** mostra os resultados da análise de SDS-PAGE do bioconjugado EPA-CP5 purificado de cromatografia de afinidade de Ni^{2+} de células, em formas de realização da presente invenção, sem e com o gene *cap5K* de flipase de *S. aureus* (SEQ ID N°: 2 e 3).

A **FIG. 13A** apresenta a análise de bioconjugado de CP5-EPA, de acordo com uma forma de realização da presente invenção, purificado por cromatografia de afinidade de Ni^{2+} e cromatografia de troca aniônica.

A **FIG. 13B** representa as massas M/Z verificadas para o sítio de glicosilação no péptido DNNNSTPTVISHR tripsinizado ligado N-glicosidicamente à massa de RU O-acetilada ($m/z=2088$ ($[\text{M}+\text{H}]^+$)), de acordo com uma forma de realização da presente invenção. A inserção ilustra a estrutura de RU conjugada ao péptido.

A **FIG. 13C** representa as massas M/Z verificadas para o sítio de glicosilação no péptido DQNR tripsinizado ligado N-glicosidicamente à massa de RU O-acetilada ($m/z=1165$ ($[\text{M}+\text{H}]^+$)),

de acordo com uma forma de realização da presente invenção. A inserção ilustra a estrutura de RU conjugada ao péptido.

A **FIG. 13D** representa uma análise de bioconjugado de CP8-EPA purificado por cromatografia de afinidade de Ni^{2+} e cromatografia de troca aniônica, de acordo com uma forma de realização da presente invenção.

A **FIG. 13E** representa bioconjugado de CP5-EPA purificado de células contendo 3 (esquerda) ou 2 plasmídeos (pista direita) para a produção de glicoconjugado, de acordo com uma forma de realização da presente invenção.

A **FIG. 13F** representa a análise de bioconjugado de CP8-EPA purificado por cromatografia de afinidade de Ni^{2+} , de acordo com uma forma de realização da presente invenção.

A **FIG. 14A** apresenta a análise de MALDI de Elevada Massa de um bioconjugado de CP5-EPA purificado de uma forma de realização da invenção produzido utilizando o sistema de 3 plasmídeos da FIG. 13A.

A **FIG. 14B** mostra a caracterização por cromatografia de exclusão por tamanho de bioconjugado de CP5-EPA de uma forma de realização da invenção produzido utilizando o sistema de 3 plasmídeos da FIG. 13A.

A **FIG. 14C** mostra a análise de SDS PAGE e imunodeteção de bioconjugado de CP5-Hla purificado, de acordo com uma forma de realização da presente invenção.

A **FIG. 14D** apresenta os resultados de bioconjugado de CP5-AcrA purificado, de acordo com uma forma de realização da presente invenção.

A **FIG. 14E** apresenta os resultados de bioconjugado de CP5-ClfA purificado, de acordo com uma forma de realização da presente invenção.

A **FIG. 15A** representa os anticorpos anti-CP5 específicos deduzidos em murganhos por bioconjugado de CP5-EPA, de acordo com uma forma de realização da presente invenção.

A **FIG. 15B** representa os anticorpos anti-CP5 específicos deduzidos em coelhos por bioconjugado de CP5-EPA, de acordo com uma forma de realização da presente invenção.

A **FIG. 16A** ilustra a atividade opsonofagocítica *in vitro* (em *S. aureus* Reynolds) de anticorpos específicos de CPS deduzidos por imunização de coelhos com CP5-EPA, de acordo com uma forma de realização da presente invenção.

A **FIG. 16B** ilustra a atividade opsonofagocítica *in vitro* (em *S. aureus* USA 100) de anticorpos específicos de CPS deduzidos por imunização de coelhos com CP5-EPA, de acordo com uma forma de realização da presente invenção.

A **FIG. 17A** representa os resultados de imunização passiva utilizando anticorpos anti-CP5-EPA, de acordo com uma forma de realização da presente invenção, em murganhos estimulados i.p. com $\sim 3,6 \cdot 10^7$ CFU da estirpe Reynolds de *S. aureus*.

A **FIG. 17B** representa os resultados de imunização passiva utilizando anticorpos anti-CP5-EPA, de acordo com uma forma de realização da invenção, em murganhos injetados com 2 mg de IgG de CP5-EPA.

A **FIG. 17C** representa os resultados de imunização passiva utilizando anticorpos anti-CP5-EPA, de acordo com uma forma de realização da invenção, em murganhos injetados com 300 µg de IgG de CP5-EPA

A **FIG. 18** representa os resultados de um ensaio de imunização ativa utilizando diferentes doses de CP5-EPA como vacina, de acordo com uma forma de realização da presente invenção e do modelo de bacteriemia de murganho para estímulo.

DESCRIÇÃO PORMENORIZADA DA INVENÇÃO

De acordo com uma forma de realização da presente divulgação, foi agora mostrado que um polissacárido de LPS de um organismo Gram-positivo é produzido num organismo Gram-negativo. A requerente considera que se trata de um novo resultado que representa um afastamento importante e significativo da técnica anterior.

Os ácidos nucleicos no âmbito da divulgação são exemplificados pelos ácidos nucleicos da divulgação contidos na Listagem de Sequências. Qualquer ácido nucleico que codifica um componente imunogénico, ou sua parte, o qual seja capaz de expressão numa célula hospedeira, pode ser utilizado na presente divulgação. As seguintes descrições de sequências são proporcionadas para facilitar a compreensão de determinados

termos utilizados por todo o pedido e não podem ser interpretadas como limitando as formas de realização da invenção.

A SEQ ID N°: 1 representa pLAFR1 (Acesso Gene Bank AY532632.1) contendo a sequência de antígeno O de O11 de *P. aeruginosa* PA0103 no sítio EcoRI, cadeia complementar (parcialmente do Acesso Gene Bank AF236052).

A SEQ ID N°: 2 representa pLAFR1 contendo o aglomerado quimérico de CP5, correspondendo ao pLAFR1-O11 com os genes *cap5HIJ* substituindo *wbjA-wzy* por recombinação homóloga. A sequência inserida também contém uma cassete de *cat* para seleção de clones recombinados homólogos.

A SEQ ID N°: 3 representa pLAFR1 contendo o aglomerado quimérico de CP5 com o gene de flipase *cap5K*, correspondendo ao pLAFR1-O11 com os genes *cap5HIJ* substituindo *wbjA-wzy* por recombinação homóloga e o *cap5K* clonado entre *cap5J* e a cassete de *cat*.

A SEQ ID N°: 4 representa pLAFR1 contendo o aglomerado quimérico de CP8 incluindo um gene de flipase, correspondendo ao pLAFR1-O11 com os genes *cap8KHIIJ* substituindo *wbjA-wzy*. A sequência inserida também contém uma cassete de *cat* para seleção de clones recombinados homólogos.

A SEQ ID N°: 5 representa um plasmídeo de expressão para a produção de Hla H35L. A ORF que codifica Hla H35L está clonada em *NdeIISacI* em pEC415.

A SEQ ID N°. 6 representa o plasmídeo de expressão para a produção do sítio 202 de Hla-H35L. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 202 de aminoácido e uma cauda de HIS C-terminal. Esta construção está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 7 representa o plasmídeo de expressão para a produção do sítio 238 de Hla-H35L. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 238 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 8 representa o plasmídeo de expressão para a produção do sítio 272 de Hla-H35L. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 272 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 9 representa um plasmídeo de expressão para a produção de ClfA. O gene foi sintetizado quimicamente e clonado no *NdeI* *SacI* no vetor de expressão pEC415.

A SEQ ID N°: 10 representa o plasmídeo de expressão para a produção do sítio 290 de ClfA. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 290 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 11 representa o plasmídeo de expressão para a produção do sítio 327 de ClfA. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição

327 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 12 representa o plasmídeo de expressão para a produção do sítio 532 de ClfA. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 532 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 13 representa a sequência de aminoácidos de EPA recombinante geneticamente destoxificado, com uma sequência sinal e dois sítios de glicosilação nas posições 260 e 402.

A SEQ ID N°: 14 representa a sequência de aminoácidos de EPA recombinante, geneticamente destoxificado, sem a sequência sinal e dois sítios de glicosilação nas posições 241 e 383.

A SEQ ID N°: 15 representa a ORF que codifica AcrA clonado por meio de NheI/SalI em pEC415.

A A SEQ ID N°: 6 representa o plasmídeo de expressão para a produção do sítio 130 de Hla-H35L. A ORF codifica um péptido sinal de DsbA N-terminal de *E. coli*, um glicosítio próximo da posição 130 de aminoácido e uma cauda de HIS C-terminal. A construção acima mencionada está clonada em NheI/SalI em pEC415.

A SEQ ID N°: 17 representa CP5 que produz o aglomerado génico com *cap5K* de flipase, seguido de uma cassete de expressão de *pglB* consistindo da sequência de ADN intergénica entre *galF* e *wbqA* do serotipo 0121 de *E. coli* e a ORF de *pglB*. Este inserto está clonado no sítio EcoRI de pLAFR1.

A SEQ ID N°: 18 representa CP8 que produz o aglomerado gênico com *cap8K* de flipase, seguido de uma cassette de expressão de *pglB* consistindo da sequência de ADN intergênica entre *galF* e *wbqA* do serotipo 0121 de *E. coli* e a ORF de *pglB*. Este inserto está clonado no sítio EcoRI de pLAFR1.

A SEQ ID N°: 19 representa CP8 que produz o aglomerado gênico com *cap8K* de flipase, seguido de uma cassette de expressão de *pglB* consistindo da sequência de ADN intergênica entre *galF* e *wbqA* do serotipo 0121 de *E. coli* e a ORF de *pglB*, além disso, esta sequência tem o gene para *wzz* do serovar 07 de *E. coli* clonado em SfaAI/BspTI, *i. e.*, entre *wzx* de *Pseudomonas aeruginosa* 011 e *cap8H*. O inserto está clonado no sítio EcoRI de pLAFR1.

A SEQ ID N°: 20 representa um plasmídeo de expressão para EPA e *wzz*. A estrutura é pACT3, na qual a cassette de resistência foi substituída (canamicina por cloranfenicol)

A SEQ ID N°: 21 representa *wzz* de serotipo 07 de *E. coli* clonado em pext21 Eco/Sal.

A SEQ ID N°: 22 representa uma sequência peptídica apresentada nos Exemplos.

A SEQ ID N°: 23 representa uma sequência peptídica apresentada nos Exemplos.

A SEQ ID N°: 24 representa uma sequência consenso proteica, D/E-X-N-Z-S/T, em que X e Z podem ser qualquer aminoácido natural, exceto prolina.

A SEQ ID N°: 25 representa um sítio de glicosilação.

A SEQ ID N°: 26 representa um sítio de glicosilação.

A SEQ ID N°: 27 representa um plasmídeo de expressão contendo a ORF de *pglB* clonada nos sítios EcoRI/BamHI.

As descrições de termos e abreviaturas surgem abaixo como utilizadas na descrição e consistentes com as utilizações conhecidas de alguém com conhecimentos gerais na técnica. As descrições são proporcionadas para facilitar a compreensão desses termos e abreviaturas e não podem ser interpretadas como formas de realização limitativas da invenção.

AcrA refere-se a uma glicoproteína de *C. jejuni*.

Imunização ativa refere-se à indução de imunidade (anticorpos) após exposição a um antigénio.

APC refere-se a células de apresentação de antigénios.

Amp refere-se a ampicilina.

Bacteriemia refere-se à presença de bactérias viáveis no sangue em circulação.

C' refere-se ao complemento.

CapA é uma enzima que se propõe ser um determinante de comprimento de cadeia em CP5 de *S. aureus*

CapB é uma enzima que se propõe ser um regulador do comprimento da cadeia de polissacárido em CP5 de *S. aureus*

CapC é uma enzima que se propõe codificar uma proteína veículo em CP5 de *S. aureus*

CapD uma enzima tendo uma atividade de 4,6 desidratase, convertendo o precursor UDPGlcNAc em UDP-2-acetamido-2,6 dideoxi-D-xilo-4-hexulose em CP5 de *S. aureus*.

CapE é uma 4,6-desidratase 3,5-epimerase catalisando a epimerização de UDP-D-GlcNAc em UDP-2-acetamido-2,6-dideoxi-D-xilo-4-hexulose em CP5 de *S. aureus*.

CapF é uma redutase, catalisa a redução de UDP-2-acetamido-2,6-dideoxi-D-xilo-4-hexulose em UDP-L-6dTalNAc em CP5 de *S. aureus*.

CapG é uma 2-Epimerase, catalisa a epimerização de UDP-L-6dTalNAc em UDP-LFucNAc em CP5 de *S. aureus*.

CapH em CP5 de *S. aureus* é uma O-acetiltransferase.

CapH em CP8 é uma transferase semelhante a capI de CP5 de *S. aureus*.

CapI em CP5 de *S. aureus* é uma glicosiltransferase que catalisa a transferência de UDP-ManNAcA em lípido transportador-D-FucNAc-L-FucNAc produzindo o lípido transportador-D-FucNAc-L-FucNAc-ManNAcA.

CapI em CP8 é uma polimerase a qual é semelhante a CapJ em CP5 de *S. aureus*.

CapJ em CP5 de *S. aureus* é uma polimerase.

CapJ em CP8 é uma O-acetiltransferase semelhante a CapH em CP5 de *S. aureus*.

CapK em CP5 de *S. aureus* é uma flipase em.

CapK em CP8 de *S. aureus* é uma flipase semelhante à CapK em CP5.

CapL é uma transferase que catalisa a transferência de UDP L-FucNAc para D-FucNAc-lípido transportador produzindo o lípido transportador-D-FucNAc-L-FucNAc em CP5 de *S. aureus*.

CapM é uma transferase que catalisa a transferência de UDP D-FucNAc para lípido transportador produzindo o lípido transportador-D-FucNAc em CP5 de *S. aureus*.

CapN é uma 4-redutase que catalisa a redução de UDP 2 acetamido-2, 6-didesoxi-D-xilo-4-hexulose em UDP-D-FucNAc em CP5 de *S. aureus*.

CapO é uma desidrogenase que catalisa a conversão de UDP D ManNAc em UDP-ManNAcA em CP5 de *S. aureus*.

CapP é uma 2-epimerase que catalisa a epimerização de UDP D-GlcNAc em UDP-D-ManNAc em CP5 de *S. aureus*.

CFU refere-se a unidade de formação de Colônia.

ClfA refere-se ao fator de aglutinação A de *S. aureus*, uma proteína ancorada à parede celular.

Vacina conjugada refere-se a uma vacina criada pela conjugação covalente de um antígeno polissacárido a uma proteína transportadora. A vacina conjugada deduz respostas imunitárias antibacterianas e memória imunológica. Em bebês e pessoas idosas, pode ser induzida uma resposta imunitária protetora contra antígenos polissacáridos se estes antígenos estiverem conjugados com proteínas que induzem uma resposta dependente de célula T.

Sequência consenso refere-se a uma sequência de aminoácidos, -D/E - X - N - Z - S/T-, em que X e Z podem ser qualquer aminoácido natural, exceto Prolina, dentro da qual se encontra o sítio de conjugação de hidrato de carbono a glicoproteínas ligadas a N.

Polissacárido capsular refere-se a uma camada espessa, de tipo mucoso, de polissacárido. Polissacáridos capsulares são solúveis em água; geralmente ácidos que consistem de unidades regularmente repetidas de um a vários monossacáridos/monómeros.

CP5 refere-se a polissacárido capsular tipo 5 ou polissacárido capsular de serotipo 5 de *Staphylococcus aureus*.

CP8 refere-se a polissacárido capsular tipo 8 ou polissacárido capsular de serotipo 8 de *Staphylococcus aureus*.

D-FucNAc refere-se a N-acetil D-fucosamina.

ECA refere-se a antígeno enterobacteriano comum.

ELISA refere-se a ensaio de imunoabsorção enzimática, uma técnica bioquímica utilizada, principalmente, em imunologia para detetar a presença de um anticorpo ou um antígeno numa amostra.

EPA ou EP Ar refere-se a exoproteína A de *P. aeruginosa* recombinante não tóxica

Vacina de glicoconjugado refere-se a uma vacina consistindo de um veículo proteico ligado a um oligossacárido antigénico.

Glicosiltransferase refere-se a enzimas que atuam como um catalisador para a transferência de uma unidade de monossacárido de um açúcar de nucleótido ativado para uma molécula aceitadora de glicosilo.

Estirpe Gram-positiva refere-se a uma estirpe bacteriana que cora de roxo com coloração Gram (uma valiosa ferramenta de diagnóstico). As bactérias Gram-positivas têm uma espessa parede celular semelhante a malha constituída por peptidoglicano (50-90% da parede celular).

Estirpe Gram-negativa refere-se a uma estirpe bacteriana que tem uma camada mais fina (10% da parede celular) que cora de rosa. As bactérias Gram-negativas também têm uma membrana exterior adicional que contém lípidos e está separada da parede celular pelo espaço periplasmático.

Hla (toxina alfa) refere-se a hemolisina alfa, que é uma toxina segregada formadora de poros e um antígeno fator de virulência essencial de *S. aureus*.

Hla H35L refere-se a uma forma mutante de toxina alfa não tóxica mutante de Hla de *S. aureus*.

Cauda de histidina ou cauda de poli-histidina, é um motivo de aminoácido em proteínas que consiste de, pelo menos, cinco resíduos de histidina (His), frequentemente no terminal N ou C, da proteína utilizada para purificar de um modo muito simples e rápido, através de ligação específica a uma coluna de afinidade de níquel.

IV refere-se a intravenosamente.

kDa refere-se a quiloDaltons, é uma unidade de massa atômica.

L-FucNAc refere-se a N-acetil L-fucosamina.

LPS refere-se a lipopolissacárido. Os lipopolissacáridos (LPS), também conhecidos por lipoglicanos, são moléculas de grandes dimensões consistindo de um lípido e um polissacárido unidos por uma ligação covalente; são encontrados na membrana exterior de bactérias Gram-negativas, atuam como endotoxinas e deduzem fortes respostas imunitárias em animais.

ManNAcA refere-se a ácido N-acetilmanosaminurónico.

Estirpes de *S. aureus* resistentes a meticilina (MRSA) refere-se a estirpe de *S. aureus* resistente a meticilina associada com permanência hospitalar mais longa e mais infecções em unidades de cuidados intensivos, o que conduz à administração de mais antibióticos.

N-glicanos ou oligossacáridos ligados a *N* refere-se a monossacáridos, oligossacáridos ou polissacáridos de composições variáveis que estão ligados a um azoto de ϵ -amida de um resíduo de asparagina numa proteína por meio de uma ligação *N* glicosídica.

Glicosilação proteica ligada a *N* refere-se a um processo ou via para ligar covalentemente "glicanos" (monossacáridos, oligossacáridos ou polissacáridos) a um azoto de cadeia lateral de asparagina (N) numa proteína alvo.

Antigénios O refere-se a um polímero de glicano repetitivo contido num LPS, também denominado polissacárido O. O antigénio O está conjugado ao oligossacárido nuclear e compreende o domínio mais exterior da molécula de LPS.

Oligossacáridos ou Polissacáridos refere-se a homopolímero ou heteropolímero formado por hidratos de carbono ligados covalentemente (monossacáridos), consistindo de unidades de repetição (monossacáridos, dissacáridos, trissacáridos, etc.) ligadas umas às outras por ligações glicosídicas.

Atividade opsonofagocítica refere-se a fagocitose de um patógeno na presença de complemento e anticorpos específicos. Considera-se que as atividades opsonofagocíticas (OPA) *in vitro* de anticorpos séricos representam as atividades funcionais dos anticorpos *in vivo* e, assim, correlacionam-se com imunidade protetora

OTase ou OST refere-se a oligossacariltransferase que catalisa uma transferência mecanisticamente única e seletiva de

um oligossacárido ou polissacárido (glicosilação) para o resíduo de asparagina (N) na sequência consenso de proteínas nascentes ou enroladas.

Imunização passiva é a transferência de imunidade humoral ativa na forma de anticorpos já preparados, de um indivíduo para outro.

Espaço periplasmático refere-se ao espaço entre a membrana citoplasmática interna e a membrana exterior externa de bactérias Gram-negativas.

PMN refere-se a neutrófilos polimorfonucleares, os quais são os glóbulos brancos mais abundantes no sangue periférico de humanos e muitos (embora não todos) mamíferos.

Veículo proteico refere-se a uma proteína que compreende a sequência consenso dentro da qual o oligo-"poli"-sacárido está conjugado.

RU refere-se a unidade de repetição, a qual é constituída por heteropolissacáridos específicos sintetizados através do agrupamento de monossacáridos individuais num oligossacárido num transportador de undecaprenilfosfato (Und-P), seguido de polimerização num oligossacárido.

Sequência sinal refere-se a um curto (e. g., aproximadamente, 3-60 aminoácidos de comprimento) péptido na extremidade N-terminal da proteína que dirige a proteína para diferentes localizações.

UDP-D-ManNAc é UDP-N-acetil-D-manosamina

UDP-D-ManNAcA é ácido UDP-N-acetil-D-manosaminurónico

UDP-D-QuiNAc é UDP-N-acetil-D-quinovosamina

UDP-L-FucNAc é UDP-N-acetil-L-fucosamina

UDP-L-6dTalNAc é UDPN-acetil-L-pneumosamina.

Und refere-se a lípido undecaprenil ou undecaprenol, composto por onze unidades de prenil.

UndP refere-se a undecaprenilfosfato, o qual é um transportador lipídico universal (derivado de Und) de intermediários biossintéticos de glicano para polímeros de hidrato de carbono que são exportados para o envelope da célula bacteriana.

UndPP refere-se a undecaprenilpirofosfato, o qual é uma versão fosforilada de UndP.

wbjA é uma glucosiltransferase em *P. aeruginosa* O11

wbjB é uma putativa epimerase semelhante às enzimas requeridas para a biossíntese da cápsula de CPS e CP8 em *S. aureus*.

wbj C é uma putativa epimerase em *P. aeruginosa* O11.

wbjD é uma putativa epimerase em *P. aeruginosa* O11.

wbjE é uma putativa epimerase em *P. aeruginosa* O11.

wbjF é uma Glicosiltransferase em *P. aeruginosa* O11.

wbpL glicosiltransferase, participa na biossíntese de LPS em *P. aeruginosa* O11.

wbpM glicosiltransferase, participa na biossíntese de LPS em *P. aeruginosa* O11

Formas de realização da invenção são, pelo menos parcialmente, baseadas na verificação de que *C. jejuni* contém um sistema geral de glicosilação de proteína ligada a N, uma característica pouco habitual para organismos procarióticos. Foi mostrado que diversas proteínas de *C. jejuni* são modificadas por um heptassacárido. Este heptassacárido é agrupado em UndPP, o lípido transportador, no lado citoplasmático da membrana interna, pela adição, passo a passo, de monossacáridos ativados por nucleótido catalisados por glicosiltransferases específicas. O oligossacárido ligado a lípido é, depois, deslocado (*i. e.*, difunde-se transversalmente) para o espaço periplasmático por uma flipase, *e. g.*, PglK. No passo final da glicosilação de proteína ligada a N, a OTase (*e. g.*, PglB) catalisa a transferência do oligossacárido do lípido transportador para resíduos de Asn na sequência consenso Asp/Glu-Xaa-Asn-Zaa-Ser/Thr (*i. e.*, D/E - X - N - Z - S/T), onde os Xaa e Zaa podem ser qualquer aminoácido, exceto Pro. A requerente transferiu com êxito o aglomerado de glicosilação para o heptassacárido dentro de *E. coli* e foi capaz de produzir glicoproteínas ligadas a N de *Campylobacter*.

Foi desenvolvido um método novo e inventivo para modificar uma bactéria de hospedeiro Gram-negativa, tal como *E. coli*, para

produzir proteínas glicosiladas para utilização como produtos vacinais contra uma bactéria Gram-positiva, tal como *S. aureus*. O desenvolvimento deste método requereu a superação de problemas significativos e, em muitos aspetos, inesperados, e afastou-se substancialmente do conhecimento convencional e da técnica anterior.

Neste método novo e inventivo, foi identificada outra bactéria Gram-negativa que produz um polissacárido que tem semelhança estrutural com o polissacárido de interesse do organismo alvo, por exemplo, *S. aureus*. Para objetivo desta invenção, a semelhança estrutural manifesta-se ela própria como unidades de repetição no polissacárido do alvo (e. g., *S. aureus*) que são parcialmente idênticas a unidades de repetição no polissacárido da bactéria identificada, outra Gram-negativa. Devido a esta última bactéria ser Gram-negativa, como é o hospedeiro, por exemplo, o organismo *E. coli*, a requerente admitiu inicialmente a hipótese (e, mais tarde, verificada por experiência, como discutido abaixo) de que a utilização destas vias de biossíntese num organismo *E. coli* modificado permitiria a biossíntese do antigénio de RU construído e o seu deslocamento do citoplasma para o periplasma do organismo *E. coli* modificado. Além disso, a requerente admitiu a hipótese (e, mais tarde, verificada por experiência, como discutido abaixo) de que o tamanho do polissacárido produzido através desta via de biossíntese seria muito menor do que o polissacárido produzido pela via de biossíntese do *S. aureus* Gram-positivo.

Como um resultado, e como discutido abaixo, o método novo e inovador desenvolvido pela requerente resolveu os difíceis problemas mencionados acima.

Além disso, foi surpreendentemente verificado que aspetos da via de LPS num organismo Gram-negativo poderiam ser utilizados para produzir polissacáridos que contêm algumas das mesmas unidades de repetição que polissacáridos capsulares nativos a bactérias Gram-positivas, tal como, por exemplo, *S. aureus*, como detalhado abaixo.

Deste modo, na preparação da secção de polissacárido da vacina de proteína glicosilada para *S. aureus*, a construção da secção de polissacárido baseada, pelo menos parcialmente, num polissacárido nativo a uma bactéria Gram-negativa, como *E. coli*, é uma solução surpreendente. A requerente verificou ainda que, agindo assim, é aparentemente importante encontrar uma bactéria que produza um polissacárido que seja tão semelhante quanto possível ao polissacárido de interesse produzido por *S. aureus*. *P. aeruginosa* é essa bactéria.

A FIG. 1 proporciona uma representação passo a passo de uma forma de realização da preparação de monossacáridos ativados por nucleótido no citoplasma, por enzimas proporcionadas no aglomerado de antigénio O, ou por enzimas de gestão da célula hospedeira Gram-negativa, como seria evidente a um especialista na técnica à luz desta descrição. Os passos do processo prosseguem da esquerda para a direita na representação da FIG. 1. Na forma de realização representada na FIG. 1, uma glicosilfosfatotransferase (WbpL) acrescenta D-FucNAc fosfato a UndP, formando UndPP-FucNAc. Glicosiltransferases específicas alongam, depois, ainda mais a molécula de UndPP-D-FucNAc, pela adição de monossacáridos, formando o oligossacárido de unidade de repetição (RU) (WbjE, WbjA). A RU é, depois, deslocada para o espaço periplasmático pela proteína Wzx. A enzima Wzy polimeriza

RU periplasmáticas para formar o polissacárido de antígeno O. O comprimento do polímero é controlado pela proteína Wzz. Muitos oligossacáridos e polissacáridos bacterianos são agrupados em UndPP e, depois, transferidos para outras moléculas. Por outras palavras, a UndPP é uma plataforma de construção geral para açúcares nas bactérias. Em *E. coli* e, considera-se, muitas outras bactérias Gram-negativas, o antígeno O é transferido de UndPP para o núcleo de Lípido A pela enzima WaaL de *E. coli*, para formar o lipopolissacárido (LPS).

A FIG. 2 representa uma forma de realização da preparação de monossacáridos ativados por nucleótido no citoplasma por enzimas proporcionadas no aglomerado de antígeno O de *P. aeruginosa* O11, por enzimas de gestão da célula hospedeira Gram-negativa e por enzimas de *S. aureus* e/ou *E. coli* que se sabe serem requeridas para a biossíntese de UDP-ManNAcA (Cap50P e/ou WecBC), como seria evidente a um especialista na técnica à luz desta descrição. Na representação da FIG. 2, os passos do processo prosseguem da esquerda para a direita. Como na biossíntese de O11, WbpL e WbjE sintetizam o dissacárido nuclear. Depois, a glicosiltransferase Cap5I de *S. aureus* adiciona D-ManNAcA. Cap5H adiciona um grupo acetilo ao segundo resíduo de FucNAc. A acetilação pode ser o passo final da síntese de RU, como mostrado na FIG. 2. O deslocamento é possível de uma ou da totalidade das proteínas Wzx no sistema, as quais são Wzx de *P. aeruginosa* ou Cap5K expressas recombinantemente, ou enzimas semelhantes a Wzx expressas endogenamente, e. g., do aglomerado de ECA codificado no cromossoma de *E. coli*. A polimerização é uma atividade exclusiva da polimerase Cap5J que forma o polissacárido CP5 em UndPP. Como outros polissacáridos ligados a UndPP, o açúcar de CP5 é

transferido para o núcleo de Lípido A pela enzima WaaL de *E. coli*, para formar LPS recombinante (cápsula de LPS).

A FIG. 3 representa a preparação de monossacáridos ativados por nucleótido no citoplasma, por enzimas proporcionadas no aglomerado de antigénio O de *P. aeruginosa* O11, por enzimas de gestão da célula hospedeira Gram-negativa e por enzimas de *S. aureus* e/ou *E. coli* que se sabe serem requeridas para a biossíntese de UDP-ManNAcA (Cap8OP e/ou WecBC), como seria evidente a um especialista na técnica à luz desta descrição. Na representação da FIG. 3, os passos do processo prosseguem da esquerda para a direita. Como na biossíntese de O11, WbpL e WbjE sintetizam o dissacárido nuclear. Depois, a glicosiltransferase Cap8H de *S. aureus* adiciona D-ManNAcA. Cap8J adiciona um grupo acetilo ao segundo resíduo de FucNAc. Não se sabe se ocorre acetilação no açúcar ativado ou na RU ligada ao lípido. O deslocamento é possível por uma ou a totalidade das proteínas Wzx no sistema, as quais são Wzx de *P. aeruginosa* ou Cap8K expressas recombinantemente, ou enzimas semelhantes a Wzx expressas endogenamente, e. g., do aglomerado de ECA codificado no cromossoma de *E. coli*. A polimerização é uma atividade exclusiva da polimerase Cap8I que forma o polissacárido CP8 em UndPP. O açúcar de CP8 é, depois, transferido para o núcleo de Lípido A em *E. coli* pela enzima WaaL.

A FIG. 4 ilustra as diferentes estruturas dos polissacáridos O11, CP5 e CP8. Mostra-se na FIG. 4 que as RU partilham a estrutura tronco idêntica, consistindo do UndPP e do dissacárido α -D-FucNAc-(1,3)-L-FucNAc. As RU de *S. aureus* são parcialmente decoradas com um único grupo O-acetilo, no L-FucNAc do meio ou no resíduo de ManNAcA, o que é característico das RU de *S. aureus*. A ligação do segundo e terceiro açúcar nas RU de

S. aureus difere entre si, assim como a ligação entre as RU polimerizadas. À direita, as estruturas de açúcar são mostradas numa representação diferente. O número próximo das setas para trás (CP5 e CP8) indica a posição do carbono modificado com um grupo O-acetilo. Uma representação alternativa das estruturas de RU está mostrada no canto inferior esquerdo. Como mostrado na FIG. 4, existe grande sobreposição entre a RU no antigénio O11 que é parte de um polissacárido nativo a *P. aeruginosa* e aquelas das cápsulas de CP5 e CP8 das respectivas estirpes de *Staphylococcus*. Em particular, como mostrado na FIG. 4, a porção L-FucNAc-->D-FucNAc na RU é idêntica em ambas.

Noutro aspeto, a divulgação caracteriza um método de identificação de um polissacárido alvo para utilização na glicosilação de uma proteína com o referido polissacárido alvo, no todo ou em parte. A referida proteína glicosilada compreendendo o polissacárido alvo pode ser utilizada, por exemplo, em composições vacinais. O método de identificação de um polissacárido alvo inclui: identificação de uma bactéria Gram-positiva, tal como *S. aureus*, como um alvo; identificação de uma primeira unidade de repetição de um polissacárido produzido pela referida bactéria Gram-positiva compreendendo, pelo menos, três monómeros; identificação de um polissacárido produzido por uma bactéria de uma espécie Gram-negativa, compreendendo uma segunda unidade de repetição compreendendo, pelo menos, dois dos mesmos monómeros que a referida primeira unidade monomérica de repetição.

Consequentemente, numa forma de realização da divulgação, um método de modificação de uma bactéria de uma primeira espécie Gram-negativa inclui: identificação de uma bactéria Gram-positiva, tal como *S. aureus*, como um alvo; identificação

de uma primeira unidade de repetição de um polissacárido produzido pela referida bactéria Gram-positiva compreendendo, pelo menos, três monómeros; identificação de um polissacárido produzido por uma bactéria de uma segunda espécie Gram-negativa, compreendendo uma segunda unidade de repetição compreendendo, pelo menos, dois dos mesmos monómeros que a referida primeira unidade de repetição; inserção na referida bactéria de uma primeira espécie Gram-negativa de uma ou mais sequências nucleotídicas que codifica glicosiltransferases que agrupam um trissacárido, contendo: a) referida segunda unidade de repetição; e b) um monómero da referida primeira unidade de repetição não presente na referida segunda unidade de repetição; inserção de uma sequência nucleotídica que codifica uma proteína, tal como uma proteína compreendendo, pelo menos, uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e inserção de uma sequência nucleotídica que codifica uma OTase.

Numa forma de realização da divulgação, o método compreende ainda a inserção numa bactéria Gram-negativa hospedeira de uma ou mais sequências nucleotídicas que codifica glicosiltransferases que agrupam um trissacárido contendo um monómero de uma primeira unidade de repetição não presente numa segunda unidade de repetição e que agrupam a segunda unidade de repetição. Uma forma de realização adicional da divulgação, envolve a inserção de uma ou mais glicosiltransferases de uma bactéria Gram-negativa que agrupam, pelo menos, uma unidade monomérica de uma primeira unidade de repetição e uma ou mais glicosiltransferases de uma bactéria Gram-positiva, tal como *S. aureus*, que agrupam, pelo menos, dois monómeros de uma segunda unidade de repetição. O método compreende, adicionalmente, a inserção numa bactéria hospedeira

Gram-negativa de uma sequência nucleotídica que codifica uma proteína e uma sequência nucleotídica que codifica uma OTase.

Em, pelo menos, uma forma de realização da divulgação, é produzida uma estirpe de *E. coli* hospedeira que transporta os correspondentes ácidos nucleicos que codificam as requeridas enzimas das estirpes CP5 e CP8 de *S. aureus*, as quais se acumulam, deslocam e polimerizam as unidades de repetição construídas. Numa forma de realização, as glicosiltransferases específicas necessárias correspondem àquelas que formam a RU L-FucNAc->D-FucNAc que são nativas a *P. aeruginosa*, e às glicosiltransferases correspondentes às que adicionam o monossacárido D-ManNAcA à RU completa que são nativas a cada das estirpes CP5 e CP8 de *S. aureus*. Uma tal forma de realização pode, ainda, incluir a utilização de um plasmídeo para injetar os ácidos nucleicos na célula hospedeira. Uma forma de realização adicional envolve a utilização, num plasmídeo, de ácidos nucleicos que codificam para as glicosiltransferases correspondentes a L-FucNAc->D-FucNAc e, num plasmídeo diferente, ácidos nucleicos que codificam para as glicosiltransferases correspondentes a D-ManNAcA. Um benefício dessas formas de realização, surpreendente, à luz da técnica anterior, é que a via de biossíntese de LPS modificada de *P. aeruginosa* que é agora responsável pela produção do polímero de RU construído da cápsula de *S. aureus* resulta numa estrutura que é muito menor do que a cápsula de *S. aureus*.

A presente divulgação é adicionalmente dirigida a uma proteína N-glicosilada recombinante compreendendo, pelo menos, uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e, pelo menos, um oligossacárido ou polissacárido de uma bactéria

Gram-positiva ligado à referida sequência consenso. Noutra forma de realização, a proteína N-glicosilada recombinante compreende duas ou mais das referidas sequências consenso inseridas. Ainda numa forma de realização adicional, a proteína N-glicosilada recombinante compreende dois ou mais dos referidos oligossacáridos ou polissacáridos de *S. aureus*. Ainda numa forma de realização adicional, a proteína N-glicosilada recombinante compreende duas ou mais das referidas sequências consenso inseridas e oligossacáridos ou polissacáridos de estirpes de *S. aureus* diferentes, por exemplo, da estirpe do polissacárido capsular 5 e estirpe do polissacárido capsular 8 de *S. aureus*.

A presente divulgação é, além disso, dirigida a uma combinação de um polissacárido capsular modificado de *S. aureus* com um antigénio proteico do mesmo organismo por ligação N-glicosídica.

Formas de realização da presente divulgação incluem uma proteína que, na natureza, está glicosilada. Essas proteínas naturalmente glicosiladas (e. g., proteínas de *C. jejuni*) contêm sequências consenso naturais mas não compreendem quaisquer sequências consenso otimizadas (i. e., introduzidas) adicionais. As proteínas naturalmente glicosiladas incluem proteínas procarióticas e eucarióticas. Formas de realização da presente divulgação incluem ainda uma proteína N-glicosilada recombinante, compreendendo uma ou mais da(s) seguinte(s) sequência(s) de aminoácidos parcial (parciais) N-glicosilada(s): D/E - X - N - Z - S/T, (sequência consenso otimizada), em que X e Z podem ser qualquer aminoácido natural, exceto Pro e em que é introduzida, pelo menos, uma da(s) referida(s) sequência(s) de aminoácidos parcial (parciais) N-glicosilada(s). A introdução da(s) sequência(s) de aminoácidos parcial (parciais)

específica(s) (sequência(s) consenso otimizada(s)) em proteínas conduz a proteínas que são *N*-glicosiladas eficientemente por uma OTase, tal como, por exemplo, uma OTase de *Campylobacter* spp., tal como, por exemplo, uma OTase de *C. jejuni*, nas posições de introdução.

A expressão "sequência(s) de aminoácidos parcial (parciais)" como é utilizada no contexto da presente invenção também será referida como "sequência(s) consenso otimizada(s)" ou "sequência(s) consenso". A sequência consenso otimizada é *N*-glicosilada por uma OTase, tal como, por exemplo, uma OTase de *Campylobacter* spp., tal como, por exemplo, uma OTase de *C. jejuni*.

De acordo com o código de uma letra internacionalmente aceite para aminoácidos, as abreviaturas D, E, N, S e T denotam ácido aspártico, ácido glutâmico, asparagina, serina e treonina, respetivamente.

A introdução da sequência consenso otimizada pode ser alcançada pela adição, eliminação e/ou substituição de um ou mais aminoácidos. A adição, eliminação e/ou substituição de um ou mais aminoácidos para introdução da sequência consenso otimizada pode ser alcançada por estratégias químicas sintéticas bem conhecidas dos especialistas na técnica, tal como síntese peptídica química assistida por fase sólida. Alternativamente, e preferido para polipéptidos maiores, as proteínas da presente invenção podem ser preparadas por técnicas recombinantes correntes, por adição de ácidos nucleicos que codificam para uma ou mais sequências consenso otimizadas à sequência de ácidos nucleicos de uma proteína de partida, a qual pode ser uma

proteína que seja naturalmente glicosilada ou pode ser uma proteína que não seja naturalmente glicosilada.

Numa forma de realização preferida, as proteínas da presente invenção podem compreender uma ou mais, de um modo preferido, pelo menos, duas ou, pelo menos, três e, de um modo mais preferido, pelo menos, cinco das referidas sequências de aminoácidos otimizadas *N*-glicosiladas introduzidas.

A presença de uma ou mais sequência(s) de aminoácidos otimizada(s) *N*-glicosilada(s) nas proteínas da presente invenção pode ter a vantagem para aumentar a sua antigenicidade, aumentando a sua estabilidade, afetando a sua atividade biológica, prolongando a sua meia-vida biológica e/ou simplificando a sua purificação.

A sequência consenso otimizada pode incluir qualquer aminoácido, exceto prolina, na posição ou posições X e Z. A expressão "quaisquer aminoácidos" pretende abranger aminoácidos naturais comuns e raros, assim como derivados e análogos de aminoácidos sintéticos, os quais vão permitir que a sequência consenso otimizada seja *N*-glicosilada pela OTase. Os aminoácidos comuns e raros de ocorrência natural são preferidos para X e Z. X e Z podem ser iguais ou diferentes.

É de notar que X e Z podem diferir para cada sequência consenso otimizada numa proteína de acordo com a presente invenção.

O *N*-glicano ligado à sequência consenso otimizada será determinado pelas glicosiltransferases específicas e sua interação quando se agrupa o oligossacárido num transportador lipídico para transferência pela OTase. Os especialistas na

técnica podem conceber o N-glicano por variação do(s) tipo(s) e quantidade das glicosiltransferases específicas presentes na célula hospedeira desejada. (Raetz & Whitfield, Endotoxinas de Lipopolissacáridos, *NIH-PA Manuscript de Autor* 1-57, 19-25 (publicado na forma editada final como: *Annual Rev. Biochem.*, 71: 635-700 (2002)); Reeves et al., Síntese de Polissacárido Bacteriano e Nomenclatura Génica, *Trends in Microbio.* 4(3): 495-503, 497-98 (Dec. 1996); e Whitfield, C. e I. S. Roberts. 1999. Estrutura, agrupamento e regulação da expressão de cápsulas em *Escherichia coli*. *Mol Microbiol* 31(5): 1307-19).

"Polissacáridos", como aqui utilizado, inclui sacáridos compreendendo, pelo menos, dois monossacáridos. Os polissacáridos incluem oligossacáridos, trissacáridos, unidades de repetição compreendendo um ou mais monossacáridos (ou monómeros) e outros sacáridos, reconhecidos como polissacáridos por alguém com conhecimentos gerais na técnica. Os N-glicanos são aqui referidos como monossacáridos, oligossacáridos ou polissacáridos de composições variáveis que estão ligados a um azoto de ϵ -amida de um resíduo de asparagina numa proteína por meio de uma ligação N-glicosídica.

Os polissacáridos das formas de realização da invenção incluem, sem limitação, polissacáridos de *S. aureus*, tais como CP5 e CP8. A forma de realização da invenção inclui, ainda, polissacáridos de *S. aureus* que visam uma bactéria, tal como um polissacárido que visa uma estirpe de *S. aureus* resistente a meticilina. Onde seja aqui mencionado que os polissacáridos visam uma estirpe bacteriana, esses polissacáridos incluem polissacáridos que são da bactéria contra a qual é desejada uma resposta imunitária e incluem polissacáridos que são iguais a,

baseados em derivados de, nativos de ou manipulados da bactéria contra a qual é desejada uma resposta imunitária.

Não existe limitação no que respeita à origem da proteína recombinante da invenção. Numa forma de realização, a referida proteína é derivada de proteínas de mamífero, bacterianas, virais, fúngicas ou vegetais. Numa forma de realização adicional, a proteína é derivada de proteínas de mamífero, de um modo muito preferido, humanas. Para a preparação de proteínas recombinantes antigénicas de acordo com a invenção, de um modo preferido, para utilização como componentes ativos em vacinas, prefere-se que a proteína recombinante seja derivada de uma proteína bacteriana, viral ou fúngica. A glicosilação de proteínas de diversas origens é conhecida de um especialista na técnica. Kowarik *et al.* "Definição da sequência consenso do sítio de N-glicosilação bacteriano" *EMBO J.* (2006) 1-10.

Num exemplo, numa forma de realização, a Exotoxina de *P. aeruginosa* (EPA) geneticamente destoxificada é um veículo proteico adequado. Para produção de uma versão de EPA que possa ser glicosilada, os ácidos nucleicos que codificam para EPA necessitam de ser modificados por inserção de sítios de glicosilação, como anteriormente discutido.

Os transportadores proteicos destinados à utilização em formas de realização da invenção devem, de um modo preferido, ter determinadas características imunológicas e farmacológicas. De uma perspetiva imunológica, de um modo preferido, um veículo proteico deve: (1) ter epitopos de célula T; (2) ser capaz de distribuir um antigénio a células apresentadoras de antigénios (APC) no sistema imunitário; (3) ser potente e durável; e (4) ser capaz de produzir uma resposta de IgG sistémica específica

de antigénio. De uma perspetiva farmacológica, um veículo proteico deve, de um modo preferido: (1) ser não tóxico; e (2) ser capaz de distribuir antigénios eficientemente através de barreiras epiteliais intactas. De um modo mais preferido, além destas características imunológicas e farmacológicas, um veículo proteico considerado para utilização na produção de um bioconjugado bacteriano deve: (1) ser facilmente segregado para o espaço periplasmático; e (2) ser capaz de ter epítopos antigénicos facilmente introduzidos como ganchos ou sequências lineares dentro dele. Informado por esta divulgação e conhecimento de alguém com conhecimentos gerais na técnica, um praticante com conhecimentos gerais na técnica pode, rotineiramente, considerar e identificar transportadores proteicos adequados que podem ser utilizados em formas de realização particulares da invenção.

Numa forma de realização da invenção, a proteína AcrA de *Campylobacter* é um veículo proteico.

Numa forma de realização adicional da invenção, a Exotoxina de *P. aeruginosa* (EPA) geneticamente destoxificada é um veículo proteico, no qual o organismo alvo para o qual se deseja uma vacina é o *S. aureus*. Contrariamente a AcrA, que contém sítios de glicosilação naturais, o EPA não contém tais sítios de glicosilação naturais e necessita de ser modificado por inserção de sítios de glicosilação (e. g., inserção de ácidos nucleicos que codificam para a sequência consenso otimizada, como discutido anteriormente, na sequência de ácidos nucleicos que codificam para EPA). Numa forma de realização adicional, o EPA é modificado para se introduzir dois sítios de glicosilação que vão permitir a glicosilação com o antigénio de *S. aureus*. Ainda numa forma de realização adicional, são introduzidas duas

sequências consenso, como discutido no Exemplo 10 do documento WO 2009/104074.

A sequência de aminoácidos de EPA, como modificada numa forma de realização desta invenção para conter dois sítios de glicosilação, é proporcionada como SEQ ID N°: 13 (com sequência sinal) e SEQ ID N°.: 14 (sem sequência sinal). Os sítios de glicosilação na SEQ ID N°: 13 são DNNNS e DQNRT, nas posições 260DNNNS e 402DQNRT. Os sítios de glicosilação na SEQ ID N°: 14 são DNNNS e DQNRT, nas posições 241DNNNS e 383DQNRT.

Uma proteína transportadora, tal como EPA, é uma proteína na qual podem ser adicionados sítios de *N*-glicosilação na produção de um bioconjugado bacteriano. Os sítios de *N*-glicosilação requerem a introdução das sequências consenso discutidas anteriormente, nomeadamente, inserção dos sequões D/E - X - N - Z-S/T, em que X e Z podem ser qualquer aminoácido natural, exceto prolina. A requerente verificou que essas sequências consenso, de um modo preferido, são introduzidas em ganchos de superfície, por inserção, ao invés de mutação, e pela utilização de resíduos flanqueantes adicionalmente inseridos e por mutação de resíduos flanqueantes, para otimizar o funcionamento do sítio de *N*-glicosilação.

Alguns antigénios de subunidade de proteína bem caracterizados de *S. aureus* são a hemolisina alfa (toxina alfa, Hla), fator de aglutinação alfa (ClfA), IsdB e Leucocidina de Panton-Valentine (PVL).

Hla é uma toxina de formação de poros segregada e um fator de virulência essencial de MRSA num modelo de murganho de pneumonia de *S. aureus*. O nível de expressão de Hla por estirpes

de *S. aureus* independentes correlaciona-se diretamente com a sua virulência. Mostrou-se que a imunização ativa com uma forma mutante de Hla (Hla H351, SEQ ID N°: 5), que não consegue formar poros (Menzies, B. E. e D. S. Kernodle. 1996. A imunização passiva com antissoro contra um mutante de toxina alfa não tóxico de *Staphylococcus aureus* é protetora num modelo murino. *Infect Immun* 64:1839-41; Jursch, R., A. Hildebrand, G. Hobom, J. Trantum-Jensen, R. Ward, M. Kehoe e S. Bhakdi. 1994. Resíduos de histidina próximos da terminação N de toxina alfa estafilocócica como repórteres de regiões que são críticas para a oligomerização e a formação de poros. *Infect Immun* 62(6): 2249-56), produzia respostas de imunoglobulina G específicas de antígeno e proporcionava proteção contra a pneumonia estafilocócica. A transferência de anticorpos específicos de Hla protege animais não imunizados contra estímulo por *S. aureus* e previne a lesão de células epiteliais de pulmão humano durante a infecção (Bubeck Wardenburg, J., A. M. Palazzolo-Ballance, M. Otto, O. Schneewind e F. R. DeLeo. 2008. A leucocidina de Panton-Valentine não é um determinante de virulência em modelos murinos de doença de *Staphylococcus aureus* resistente a meticilina associada à comunidade. *J Infect Dis* 198:1166-70). Para ser utilizada como uma vacina, a mutação H35L em Hla é requerida para eliminar a toxicidade da proteína (Menzies, B. E. e D. S. Kernodle. 1994. Mutagênese dirigida a sítio do gene da toxina alfa de *Staphylococcus aureus*: papel das histidinas na atividade da toxina *in vitro* e num modelo murino. *Infect Immun* 62:1843-7). O ClfA contém um domínio resistente a protease que é utilizado para imunização. A imunização passiva de murganhos com anticorpos anti-ClfA e anti-CP5 esterilizou eficazmente as glândulas mamárias num modelo de infecção de glândula mamária (Tuchscher, L. P., F. R. Buzzola, L. P. Alvarez, J. C. Lee e D. O. Sordelli. 2008. Anticorpos para polissacárido capsular e

fator de aglutinação A, previnem mastite e o surgimento de variantes não encapsulados e de colônias pequenas de *Staphylococcus aureus* em murganhos. *Infect Immun* 76: 5738-44).

Uma forma de realização adicional da invenção inclui a glicosilação de proteínas nativas a *S. aureus*, por exemplo, Hla e ClfA. Em formas de realização exemplificativas adicionais da invenção, o veículo proteico utilizado pode ser selecionado como sendo a proteína Hla, por exemplo, Hla H351 (por exemplo, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8 ou SEQ ID N°: 16). Noutra forma de realização exemplificativa adicional da invenção, o veículo proteico é a proteína ClfA (por exemplo, SEQ ID N°: 10, SEQ ID N°: 11 ou SEQ ID N°: 12).

A divulgação é ainda dirigida a organismos procarióticos hospedeiros recombinantes, compreendendo: uma sequência nucleotídica que codifica uma ou mais glicosiltransferases de uma primeira espécie procariótica, tal como uma espécie Gram-positiva; uma ou mais glicosiltransferases de uma espécie procariótica diferente, tal como uma espécie Gram-negativa; uma sequência nucleotídica que codifica uma proteína; e uma sequência nucleotídica que codifica uma OTase. A divulgação é, adicionalmente, dirigida a um organismo procariótico hospedeiro recombinante, compreendendo uma sequência nucleotídica introduzida que codifica glicosiltransferases nativas apenas a um organismo procariótico Gram-positivo; uma sequência nucleotídica que codifica uma proteína; e uma sequência nucleotídica que codifica uma OTase. A divulgação também é dirigida a um organismo procariótico hospedeiro recombinante ou manipulado, compreendendo: uma sequência nucleotídica que codifica uma glicosiltransferase nativa a uma primeira espécie procariótica, a qual é, por exemplo, diferente do organismo

procariótico hospedeiro; uma sequência nucleotídica que codifica uma glicosiltransferase nativa a uma segunda espécie procariótica diferente da espécie do referido primeiro organismo procariótico e, por exemplo, diferente do referido hospedeiro. O organismo procariótico manipulado também pode, por exemplo, compreender uma primeira espécie procariótica que seja uma espécie Gram-positiva. O organismo procariótico manipulado também pode, por exemplo, compreender uma segunda espécie procariótica que seja uma espécie Gram-negativa. A divulgação inclui, ainda, um organismo procariótico hospedeiro Gram-negativo recombinante ou manipulado, compreendendo: uma sequência nucleotídica que codifica uma glicosiltransferase nativa a uma espécie procariótica Gram-negativa que é, por exemplo, diferente do referido organismo procariótico hospedeiro; uma sequência nucleotídica que codifica uma glicosiltransferase nativa a *S. aureus*; uma sequência nucleotídica que codifica uma proteína; e uma sequência nucleotídica que codifica uma OTase. A invenção inclui, ainda, um hospedeiro de *E. coli* recombinante ou manipulado, compreendendo: uma sequência nucleotídica que codifica uma glicosiltransferase nativa a *P. aeruginosa*; uma sequência nucleotídica que codifica uma ou mais glicosiltransferases nativas a estirpe CP5 de *S. aureus* e/ou a estirpe CP8 de *S. aureus*; uma sequência nucleotídica que codifica um EPA de *P. aeruginosa*, hemolisina alfa de *S. aureus* ou veículo proteico do fator de aglutinação A de *S. aureus*; e uma sequência nucleotídica que codifica uma OTase, por exemplo, e OTase nativa a *C. jejuni*.

Além da utilização da via de biossíntese do outro organismo Gram-negativo no organismo de *E. coli* hospedeiro modificado, numa forma de realização adicional, também estão incluídos no

organismo de *E. coli* hospedeiro ácidos nucleicos que codificam para (i) glicosiltransferases para construção da estrutura das unidades de repetição do polissacárido do outro organismo Gram-negativo (que são idênticas às unidades de repetição do polissacárido de interesse do organismo de *S. aureus* Gram-positivo alvo) e (ii) glicosiltransferases para construção das unidades do polissacárido de interesse do organismo de *S. aureus* Gram-positivo alvo que não são encontradas no polissacárido relevante do outro organismo Gram-negativo e (iii) enzimas para deslocamento e polimerização da RU de interesse construída do organismo de *S. aureus* Gram-positivo alvo para formar um polissacárido semelhante à cápsula de *S. aureus*. Em particular, nesta forma de realização, os ácidos nucleicos que codificam para (i) tiveram origem numa outra bactéria Gram-negativa, enquanto que os ácidos nucleicos que codificam para (ii) e (iii) tiveram origem no organismo de *S. aureus* Gram-positivo alvo.

Outro aspeto da divulgação é dirigido a: um organismo procariótico hospedeiro manipulado, compreendendo: i) uma sequência nucleotídica que codifica glicosiltransferases nativas a uma espécie procariótica Gram-positiva; ii) uma sequência nucleotídica que codifica uma proteína; e iii) uma sequência nucleotídica que codifica uma OTase, em que as sequências que codifica genes transportadores da referida espécie procariótica Gram-positiva estão eliminadas. Essa forma de realização envolve uma construção de ácido nucleico introduzida que codifica apenas glicosiltransferases Gram-positivas.

Com respeito aos outros ácidos nucleicos que seriam inseridos no hospedeiro em uma ou mais outras formas de realização, ácidos nucleicos que codificam uma proteína, tais

como AcrA, Hla, ClfA ou EPA (SEQ ID N°: 15, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 16; SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12; SEQ ID N°: 13, SEQ ID N°: 14) e a oligossacariltransferase de *C. jejuni* (SEQ ID N°: 27), que são parte da maquinaria de glicosilação desse organismo, são injetados no hospedeiro além dos ácidos nucleicos que codificam para glicosiltransferases de cada de *P. aeruginosa* e *S. aureus*. Como um resultado, o organismo de *E. coli* modificado pode glicosilar a proteína AcrA com o polissacárido produzido nesse organismo por ação das glicosiltransferases de *S. aureus* e da outra bactéria Gram-negativa.

Uma forma de realização da divulgação envolve um organismo procariótico hospedeiro manipulado, compreendendo: i) uma sequência nucleotídica que codifica uma glicosiltransferase nativa a uma primeira espécie procariótica diferente do organismo procariótico hospedeiro; ii) uma sequência nucleotídica que codifica uma glicosiltransferase nativa a uma segunda espécie procariótica, por exemplo, uma espécie procariótica Gram-positiva, diferente do organismo procariótico hospedeiro; iii) uma sequência nucleotídica que codifica uma proteína; e iv) uma sequência nucleotídica que codifica uma OTase. Em formas de realização da invenção, a primeira espécie procariótica é uma espécie Gram-negativa, por exemplo, *P. aeruginosa*.

No contexto da presente divulgação, células hospedeiras referem-se a qualquer célula hospedeira, e. g., uma célula hospedeira eucariótica ou procariótica. Noutras formas de realização, a célula hospedeira é uma célula hospedeira procariótica, e. g., *Escherichia* ssp., *Campylobacter* ssp., *Salmonella* ssp., *Shigella* ssp., *Helicobacter* ssp., *Pseudomonas*

ssp. ou *Bacillus* ssp. Ainda em formas de realização adicionais, a célula hospedeira é *Escherichia coli*, *Campylobacter jejuni*, *Salmonella typhimurium*, etc.

A divulgação é, além disso, dirigida a métodos de produção de uma vacina de bioconjugado, compreendendo a introdução num organismo procariótico hospedeiro de ácidos nucleicos que codificam uma ou mais glicosiltransferases de *S. aureus*; uma ou mais glicosiltransferases de uma segunda espécie procariótica, uma proteína; e uma OTase. Além disso, a presente divulgação é dirigida à produção de vacinas de bioconjugado por produção, em bactérias Gram-negativas, de polissacáridos capsulares modificados em undecaprenol (Und) e ligação destes antigénios polissacáridos a um veículo proteico de escolha.

A divulgação é ainda dirigida a métodos de produção de proteínas glicosiladas num organismo procariótico hospedeiro, compreendendo sequência nucleotídica que codifica glicosiltransferases nativas a um primeiro organismo procariótico e que também codificam glicosiltransferases nativas a um segundo organismo procariótico que é diferente do primeiro organismo procariótico. A presente divulgação é adicionalmente dirigida à produção de proteínas N-glicosiladas com polissacáridos capsulares de bactérias Gram-positivas, os quais são sintetizados por uma combinação de diferentes glicosiltransferases de diferentes organismos. A divulgação é, além disso, dirigida à produção de proteínas glicosiladas num organismo procariótico hospedeiro, compreendendo uma sequência nucleotídica introduzida que codifica glicosiltransferases nativas apenas a um organismo procariótico Gram-positivo.

Como é conhecido na técnica, a biossíntese de diferentes polissacáridos em células bacterianas é conservada. Os polissacáridos são agrupados em lípidos transportadores de precursores comuns (nucleótidos de açúcar ativados) na membrana citoplasmática por diferentes glicosiltransferases com especificidade definida. (Whitfield, C. e I. S. Roberts. 1999. Estrutura, agrupamento e regulação da expressão de cápsulas em *Escherichia coli*. *Mol Microbiol* 31: 1307-19). A via biossintética para a produção de polissacáridos de antígeno O em Gram-negativos e para polissacárido capsular Tipo I em Gram-positivos é conservada. O processo utiliza o mesmo transportador lipídico, i. e., UndP, para agrupamento do polissacárido. Começa com a adição de um monossacárido-1-fosfato ao lípido transportador UndP no lado citoplasmático da membrana. O antígeno é acumulado por adição sequencial de monossacáridos de nucleótidos de açúcar, ativados por diferentes glicosiltransferases. O oligossacárido ligado a lípidos ou RU é, depois, deslocado através da membrana pela flipase. As RU são polimerizadas pela enzima Wzy no espaço periplasmático, que formam o denominado antígeno O em bactérias Gram-negativas ou polissacárido capsular em bactérias Gram-positivas. As bactérias Gram-negativas utilizam a enzima Wzz para regular o comprimento do polímero, o qual é, depois, transferido para o núcleo do lípido A, formando LPS. O LPS é, ainda, translocado para a membrana exterior expondo o antígeno O ao exterior (como representado, por exemplo, na FIG. 1). As bactérias Gram-positivas, em contraste, formam a cápsula deste precursor ligado ao lípido por transporte adicional, utilizando uma maquinaria enzimática diferente e especializada. As vias biossintéticas destes polissacáridos possibilitam a produção de bioconjugados *in vivo*, pela captura dos polissacáridos no periplasma num veículo proteico.

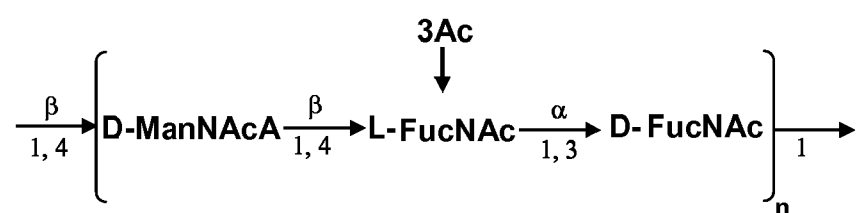
O processo da construção de polissacárido difere para polissacáridos capsulares, na medida em que o polissacárido capsular é libertado do lípido transportador após polimerização e exportado para a superfície. Em bactérias Gram-positivas, como *S. aureus*, os quais não contêm um compartimento periplasmático, a polimerização do antígeno ocorre no lado exterior da membrana. Além disso, a regulação do comprimento em *S. aureus* está incluída na maquinaria de três enzimas responsáveis pelo agrupamento da cápsula. Neste agrupamento, o polissacárido é libertado do lípido transportador e exportado para a superfície por um processo enzimático.

Os elementos genéticos encontrados no aglomerado génico requeridos para expressão da cápsula funcional em *S. aureus* assemelham-se à maquinaria genética encontrada em aglomerados de síntese de antígeno O dependente de *wzy*. (Dean, C. R., C. V. Franklund, J. D. Retief, M. J. Coyne, Jr., K. Hatano, D. J. Evans, G. B. Pier e J. B. Goldberg. 1999. Caracterização do locus do antígeno O do serogrupo O11 de *Pseudomonas aeruginosa* PA103. *J Bacteriol* 181:4275-4284).

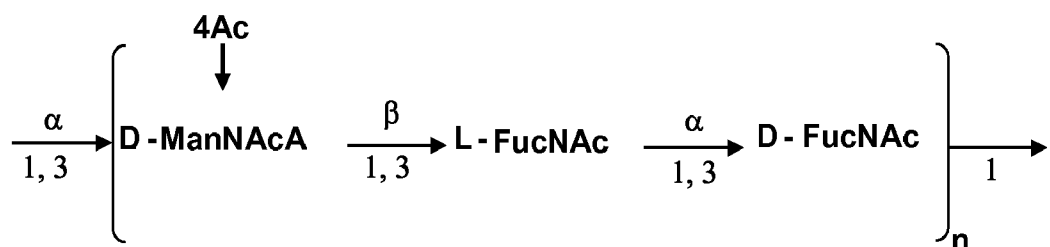
Não obstante estas diferenças entre a construção de polissacáridos em bactérias Gram-positivas e Gram-negativas, foi surpreendentemente identificado e verificado que os aspetos da via de LPS num organismo Gram-negativo poderiam ser utilizados para produzir polissacáridos que contêm algumas das mesmas unidades de repetição que polissacáridos capsulares nativos a bactérias Gram-positivas, tal como, por exemplo, *S. aureus*. Como esses polissacáridos são produzidos pelos mecanismos da via de LPS no hospedeiro Gram-negativo, a estrutura desses polissacáridos é a mesma que em precursores de polissacáridos

LPS. Tais polissacáridos produzidos em sistemas Gram-negativos da presente invenção podem, para os fins deste pedido, ser caracterizados, deste modo, como "polissacáridos capsulares modificados" ou "cápsulas de LPS". Além disso, este sistema de expressão e via biossintética sintetizados de novo, os quais combinam as vias biossintéticas de LPS e capsular, podem, para os objetivos deste pedido, ser caracterizados como sendo uma "via biossintética de LPS modificada".

Numa forma de realização da presente invenção, um polissacárido modificado produzido por uma via biossintética de LPS modificada compreende:



Numa forma de realização adicional da presente invenção, um polissacárido modificado produzido por uma via biossintética de LPS modificada compreende:



Utilizando a tecnologia da invenção, podem ser produzidos bioconjugados bacterianos que são imunogênicos. Podem ser realizadas modificações genéticas permitindo a conjugação *in*

vivo de polissacáridos bacterianos em proteínas desejadas e em posições desejadas.

Outro aspecto da divulgação envolve a produção de cápsulas de LPS ou LPS modificados conjugados a um veículo proteico utilizando a via biossintética de LPS modificada, como discutido acima.

Uma forma de realização adicional da divulgação inclui uma construção de sequência nucleotídica que codifica o aglomerado de biossíntese do polissacárido completo de Cap5 e Cap8, em que os genes transportadores eliminados são *capA*, *capB* e *capC* de *S. aureus* (ver a FIG. 6).

Uma forma de realização adicional da invenção inclui a integração do aglomerado quimérico CP5/O11 (SEQ ID N°. 2, SEQ ID N°. 3 ou SEQ ID N°. 17) ou do aglomerado quimérico CP8/O11 (SEQ ID N°. 4, SEQ ID N°. 18 ou SEQ ID N°. 19) no genoma de uma célula hospedeira. Uma forma de realização adicional da invenção envolve a integração no genoma de uma célula hospedeira: (a) o aglomerado quimérico CP5/O11 (SEQ ID N°. 2, SEQ ID N°. 3 ou SEQ ID N°. 17) ou o aglomerado quimérico CP8/O11 (SEQ ID N°. 4 SEQ ID N°. 18 ou SEQ ID N°. 19); (b) ácidos nucleicos que codificam a OTase; e (c) ácidos nucleicos que codificam uma proteína com ou sem uma sequência consenso introduzida.

Outra forma de realização da presente divulgação é dirigida a plasmídeos, tal como, por exemplo, plasmídeos compreendendo uma ou mais das SEQ ID N°: 2; SEQ ID N°: 3; SEQ ID N°: 4; SEQ ID N°: 17; SEQ ID N°: 18 e SEQ ID N°: 19. A invenção também inclui plasmídeos compreendendo uma ou mais das SEQ ID N°: 13;

SEQ ID N°: 14 e SEQ ID N°: 15. A invenção também se refere a plasmídeos compreendendo uma ou mais das SEQ ID N°: 16; SEQ ID N°: 6; SEQ ID N°: 7 e SEQ ID N°: 8. A invenção também se refere a plasmídeos compreendendo uma ou mais das SEQ ID N°: 10; SEQ ID N°: 11 e SEQ ID N°: 12. Além disso, a invenção é dirigida a plasmídeos compreendendo uma ou mais das SEQ ID N°: 20; SEQ ID N°: 21 e SEQ ID N°: 27.

Formas de realização da presente divulgação são, além disso, dirigidas a células bacterianas transformadas, tal como, por exemplo, incluindo uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°. 2; SEQ ID N°. 3; SEQ ID N°: 4; SEQ ID N°: 17; SEQ ID N°: 18; SEQ ID N°: 19; SEQ ID N°: 20; SEQ ID N°: 21 e SEQ ID N°: 27. Está ainda incluída na invenção uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 19 e SEQ ID N°: 20. Está adicionalmente incluída uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 13, SEQ ID N°: 19 e SEQ ID N°: 21. A presente invenção é, ainda, dirigida a uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 16, SEQ ID N°: 6; SEQ ID N°: 7; SEQ ID N°: 8; SEQ ID N°: 10; SEQ ID N°: 11 e SEQ ID N°: 12. A divulgação é adicionalmente dirigida a células bacterianas transformadas, tal como, por exemplo, incluindo uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°. 3; SEQ ID N°: 4; SEQ ID N°: 17; SEQ ID N°: 18; e SEQ ID N°: 19 e em que a referida célula bacteriana expressa uma glicosiltransferase nativa a *P. aeruginosa* e uma glicosiltransferase nativa a CP5 e/ou CP8 de *S. aureus*. Está ainda incluída na divulgação uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°: 17;

SEQ ID N°: 18 e SEQ ID N°: 19, em que a referida célula bacteriana expressa uma glicosiltransferase nativa a *P. aeruginosa*, uma glicosiltransferase nativa a CP5 e/ou CP8 de *S. aureus* e PglB. Está ainda incluída adicionalmente na presente divulgação (a) uma célula bacteriana transformada com um plasmídeo compreendendo a SEQ ID N°. 19, em que a referida célula bacteriana expressa uma glicosiltransferase nativa a *P. aeruginosa*, uma glicosiltransferase nativa a CP8 de *S. aureus*, Wzz de *E. coli* serovar 07 e PglB; (b) uma célula bacteriana transformada com um plasmídeo compreendendo uma ou mais das SEQ ID N°. 19 e SEQ ID N°. 20, em que a referida célula bacteriana expressa uma glicosiltransferase nativa a *P. aeruginosa*, uma glicosiltransferase nativa a CP8 de *S. aureus*, Wzz (regulador de comprimento), EPA e PglB; e (c) uma célula bacteriana compreendendo uma ou mais das SEQ ID N°. 16; SEQ ID N°. 6; SEQ ID N°. 7; SEQ ID N°. 8; SEQ ID N°. 13; SEQ ID N°. 14; SEQ ID N°. 15; SEQ ID N°. 10; SEQ ID N°. 11 e SEQ ID N°. 12.

As formas de realização da presente divulgação são, adicionalmente, dirigidas a um método de indução de uma resposta imunitária contra uma infecção causada por bactérias Gram-positivas e outras bactérias num mamífero, tal como, por exemplo, num humano. Numa forma de realização, o método compreende a administração ao referido mamífero de uma quantidade eficaz de uma composição farmacêutica compreendendo: proteína compreendendo, pelo menos, uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e um ou mais oligossacáridos ou polissacáridos, os um ou mais oligossacáridos ou polissacáridos que são iguais ou diferentes a outros dos um ou mais oligossacáridos ou polissacáridos, de uma bactéria

Gram-positiva ligada à referida sequência consenso. Uma forma de realização adicional da presente divulgação inclui um método de indução de uma resposta imunitária contra uma infecção causada por *S. aureus* num mamífero, compreendendo a administração ao referido mamífero de uma quantidade eficaz de uma composição farmacêutica compreendendo: uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; pelo menos, um oligossacárido ou polissacárido de *S. aureus*, tal como polissacárido CP5; e um adjuvante farmacêuticamente aceitável. Outra forma de realização da divulgação é dirigida à indução de uma resposta imunitária contra uma infecção causada por *S. aureus* num mamífero, compreendendo a administração ao referido mamífero de uma quantidade eficaz de uma composição farmacêutica compreendendo: uma proteína compreendendo uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; pelo menos, um polissacárido CP8 de *S. aureus*; e um adjuvante farmacêuticamente aceitável. Ainda uma forma de realização adicional é dirigida à indução de uma resposta imunitária contra uma infecção causada por *S. aureus* num mamífero, compreendendo a administração de uma quantidade eficaz de uma composição farmacêutica compreendendo uma proteína com duas ou mais sequências consenso e oligossacáridos ou polissacáridos de diferentes estirpes bacterianas Gram-positivas. Ainda uma forma de realização adicional é dirigida à indução de uma resposta imunitária contra uma infecção causada por *S. aureus* num mamífero, compreendendo a administração de uma quantidade eficaz de uma composição farmacêutica compreendendo uma proteína com duas ou mais sequências consenso e polissacáridos compreendendo CP5 de *S. aureus* e CP8 de *S. aureus*.

Nos casos, nesta descrição, onde sejam anotadas sequências de nucleótidos ou de aminoácidos específicas, será entendido que a presente invenção abrange sequências homólogas que ainda incorporam a mesma funcionalidade que as sequências anotadas. Numa forma de realização da invenção, essas sequências são, pelo menos, 85% homólogas. Noutra forma de realização, essas sequências são, pelo menos, 90% homólogas. Ainda em formas de realização adicionais, essas sequências são, pelo menos, 95% homólogas. A determinação da percentagem de identidade entre duas sequências de nucleótidos ou aminoácidos é conhecida de um especialista na técnica.

As sequências de ácidos nucleicos aqui descritas, tal como aquelas descritas nas listagens de sequências que acompanham esta descrição, são apenas exemplos e será evidente a um especialista na técnica que estas sequências podem ser combinadas de diferentes modos. Formas de realização adicionais da invenção incluem variantes de ácidos nucleicos. Um variante de um ácido nucleico (e. g., um ácido nucleico codão-otimizado) pode ser substancialmente idêntico, ou seja, pelo menos, 70% idêntico, por exemplo, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% ou 99,5% idêntico às SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22, SEQ ID N°: 23, SEQ ID N°: 24, SEQ ID N°: 25, SEQ ID N°: 26 e/ou SEQ ID N°: 27. Variantes de ácido nucleico de uma sequência que contém SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10,

SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22, SEQ ID N°: 23, SEQ ID N°: 24, SEQ ID N°: 25, SEQ ID N°: 26 e/ou SEQ ID N°: 27. Inclui ácidos nucleicos com uma substituição, variação, modificação, substituição, eliminação e/ou adição de um ou mais nucleótidos (e. g., 2, 3, 4, 5, 6, 8, 10, 12, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500 ou mais nucleótidos) de uma sequência que contém as SEQ ID N°: 1, SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 5, SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 9, SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12, SEQ ID N°: 13, SEQ ID N°: 14, SEQ ID N°: 15, SEQ ID N°: 16, SEQ ID N°: 17, SEQ ID N°: 18, SEQ ID N°: 19, SEQ ID N°: 20, SEQ ID N°: 21, SEQ ID N°: 22, SEQ ID N°: 23, SEQ ID N°: 24, SEQ ID N°: 25, SEQ ID N°: 26 e/ou SEQ ID N°: 27 ou suas partes.

Tais variantes incluem ácidos nucleicos que codificam glicosiltransferases procarióticas e que i) são expressos numa célula hospedeira, tal como *E. coli* e ii) são substancialmente idênticos às SEQ ID N°: 2, SEQ ID N°: 3, SEQ ID N°: 4, SEQ ID N°: 17, SEQ ID N°: 18 e/ou SEQ ID N°: 19 e/ou suas partes.

Os ácidos nucleicos aqui descritos incluem ADN recombinante e ADN sintético (e. g., sintetizado quimicamente). Os ácidos nucleicos podem ser de cadeia dupla ou cadeia simples. No caso de ácidos nucleicos de cadeia simples, o ácido nucleico pode ser uma cadeia sentido ou cadeia antissentido. Os ácidos nucleicos podem ser sintetizados utilizando análogos ou derivados

oligonucleotídicos, como conhecido de um especialista na técnica à luz desta descrição.

Os plasmídeos que incluem um ácido nucleico aqui descrito podem ser transformados em células hospedeiras para expressão. As técnicas para transformação são conhecidas dos especialistas na técnica à luz desta descrição.

Uma forma de realização adicional da divulgação envolve a produção de vacinas de bioconjugado Gram-positivo contendo cápsulas de LPS ou LPS modificados conjugados a um veículo proteico.

Uma forma de realização adicional da divulgação envolve uma nova vacina de bioconjugado. Uma forma de realização adicional da divulgação envolve uma nova abordagem para a produção dessas vacinas de bioconjugado que utilizam células bacterianas recombinantes que produzem diretamente bioconjugados imunogénicos ou antigénicos. Numa forma de realização, as vacinas de bioconjugado podem ser utilizadas para tratar ou prevenir doenças bacterianas, tais como diarreia, infeções nosocomiais e meningite. Em formas de realização adicionais, as vacinas de bioconjugado podem ter potencial terapêutico e/ou profilático para o cancro ou outras doenças.

Noutra forma de realização da presente divulgação, complexos sintetizados de polissacáridos (*i. e.*, resíduos de açúcar) e proteínas (*i. e.*, transportadores proteicos) podem ser utilizados como vacinas conjugadas para proteger contra infeções, tal como infeções por *S. aureus*. Numa forma de realização, uma vacina de bioconjugado, tal como uma vacina Gram-positiva, compreende um veículo proteico compreendendo uma

sequência consenso de ácidos nucleicos inserida; pelo menos, um oligossacárido ou polissacárido de uma bactéria Gram-positiva ligado à sequência consenso e, opcionalmente, um adjuvante. Noutra forma de realização, a presente divulgação é, ainda, dirigida a uma vacina de bioconjugado Gram-positiva, tal como uma vacina de *S. aureus*, compreendendo um veículo proteico compreendendo uma sequência consenso de ácidos nucleicos inserida; pelo menos, um oligossacárido ou polissacárido de uma bactéria Gram-positiva, tais como polissacárido capsular ou cápsula de LPS, ligado à sequência consenso e, opcionalmente, um adjuvante. Noutra forma de realização da invenção, a vacina de bioconjugado de *S. aureus* compreende duas ou mais destas sequências consenso inseridas. Numa forma de realização adicional, a vacina de bioconjugado de *S. aureus* compreende dois ou mais de oligossacáridos ou polissacáridos de *S. aureus*. Ainda uma forma de realização adicional compreende duas ou mais das referidas sequências consenso inseridas e oligossacáridos ou polissacáridos de estirpes de *S. aureus* diferentes, por exemplo, da estirpe de polissacárido capsular 5 (CP5) e da estirpe de polissacárido capsular 8 (CP8) de *S. aureus*.

Uma forma de realização adicional da presente invenção envolve uma vacina de *S. aureus* realizada por um sistema de glicosilação utilizando uma via de LPS modificada, que compreende a produção de um polissacárido capsular ou cápsula de LPS modificados. Ainda uma forma de realização adicional envolve uma vacina de *S. aureus* realizada por um sistema de glicosilação que utiliza uma via de LPS modificada, o qual compreende a produção de um polissacárido capsular modificado a partir de ácidos nucleicos introduzidos que não codificam glicosiltransferases de uma espécie procariótica Gram-negativa.

Uma forma de realização adicional envolve uma vacina de *S. aureus* produzida por um sistema de glicosilação, compreendendo ácidos nucleicos que codificam: i) uma ou mais glicosiltransferases responsáveis pela produção do L-FucNAc-->D-FucNAc da RU do antígeno O11 nativo a *P. aeruginosa*; ii) uma ou mais glicosiltransferases responsáveis pela produção do D-ManNAcA contendo RU nativa a qualquer das estirpes CP5 ou CP8 de *S. aureus*; iii) uma ou mais enzimas responsáveis pelo deslocamento e polimerização das RU de CP5 ou CP8 construídas, iv) uma proteína recombinante contendo sequências consenso introduzidas; e v) oligossacariltransferase de *C. jejuni*. Nesta forma de realização, o organismo hospedeiro pode ser uma bactéria Gram-negativa, por exemplo, *E. coli*.

Uma forma de realização adicional da invenção envolve uma vacina de *S. aureus* produzida por um sistema de glicosilação compreendendo ácidos nucleicos que codificam: i) glicosiltransferases responsáveis pela produção do L-FucNAc-->D-FucNAc da RU do antígeno O11 nativo a *P. aeruginosa*; ii) uma glicosiltransferase responsável pela produção do D-ManNAcA contendo RU nativa a qualquer das estirpes CP5 ou CP8 de *S. aureus*; iii) proteína AcrA de *C. jejuni*; e iv) oligossacariltransferase de *C. jejuni*. Nesta forma de realização, o organismo hospedeiro pode ser a bactéria Gram-negativa, *E. coli*.

As vacinas da presente invenção têm utilidades terapêuticas e profiláticas. Será entendido que a vacina da invenção pode ser útil nas áreas de medicina humana e medicina veterinária. Assim, o sujeito a ser imunizado pode ser um humano ou outro animal, por exemplo, animais de criação, incluindo vacas, ovelhas, porcos, cavalos, cabras e aves domésticas (e.g., galinhas,

perus, patos e gansos) e animais de companhia, tais como cães e gatos.

Noutro aspeto, a divulgação é dirigida a um método de produção de vacinas para imunização de um mamífero contra uma bactéria, tal como uma bactéria Gram-positiva. O método inclui: imunização de um sujeito com um bioconjugado, tal como um bioconjugado compreendendo um polissacárido Gram-positivo, e. g., um polissacárido de *S. aureus* e um veículo farmacêuticamente aceitável.

Esta divulgação também caracteriza composições vacinais para proteção contra infeção por uma bactéria gram-positiva, tal como *S. aureus*, ou para tratamento de uma infeção gram-positiva, tal como infeção por *S. aureus*. Numa forma de realização, as composições vacinais compreendem um ou mais componentes imunogénicos, tal como a polissacárido, ou um seu fragmento ou parte, de *S. aureus*. Numa forma de realização adicional, as composições vacinais compreendem um ou mais componentes imunogénicos, tal como uma proteína, ou um fragmento ou porção desta, de uma bactéria Gram-negativa ou Gram-positiva.

Um aspeto da divulgação proporciona uma composição vacinal para proteção contra infeção por *S. aureus* que contém, pelo menos, um componente imunogénico ou fragmento de um polissacárido de *S. aureus* e um veículo farmacêuticamente aceitável. Esses componentes ou fragmentos imunogénicos podem, incluir, por exemplo, um polissacárido de *S. aureus* de, pelo menos, cerca de dois monómeros em comprimento ou, pelo menos, cerca de três monómeros em comprimento. Num aspeto adicional da invenção, uma RU de *S. aureus* compreende os referidos monómeros.

Essas unidades de repetição podem incluir, por exemplo, uma RU de *S. aureus* de, pelo menos, 1 (um) em comprimento.

Os componentes ou fragmentos imunogénicos da divulgação podem ser obtidos, por exemplo, por rastreio de polissacáridos ou polipéptidos produzidos recombinantemente ou através de síntese química ou, por exemplo, por rastreio do bioconjugado compreendendo um polissacárido e uma proteína. O rastreio de componentes ou fragmentos imunogénicos da divulgação pode ser realizado, utilizando um ou mais dos vários diferentes ensaios. Por exemplo, os ensaios de rastreio incluem ELISA e outros ensaios conhecidos de alguém com conhecimentos gerais na técnica.

Numa forma de realização, os componentes ou fragmentos imunogénicos são identificados pela capacidade do polissacárido e/ou proteína estimular anticorpos de IgG contra bactérias Gram-positivas, tais como polissacáridos CP5 ou CP8 de *S. aureus*, como determinado, por exemplo, pela resposta imunitária obtida em murganhos (FIG. 15A) e no coelho (FIG. 15B), medindo anticorpos anti-CP5 específicos (quantificados por ELISA) contra o candidato vacinal de glicoconjugado CP5-EPA e outros meios conhecidos de uma pessoa com conhecimentos gerais na técnica.

Numa forma de realização, os componentes ou fragmentos imunogénicos são identificados pela capacidade do polissacárido e/ou proteína estimular a atividade opsonica, tal como a neutralização opsonofagocítica, como determinado, por exemplo, pela neutralização de *S. aureus* (atividade "in vitro") com anticorpos anti-CP5-EPA de coelho (obtidos no Exemplo 7 abaixo, ver a FIG. 15B) e outros meios conhecidos de uma pessoa com conhecimentos gerais na técnica.

Ainda numa forma de realização adicional, os componentes ou fragmentos imunogénicos são identificados pela capacidade de o polissacárido e/ou proteína estimular a imunidade humoral e/ou de mediação celular contra bactérias Gram-positivas, tal como *S. aureus*, como determinado, por exemplo, pela proteção contra a infeção bacteriana ("estímulo") utilizando imunização ativa em murganhos (FIG. 18) com CP5-EPA e outros meios conhecidos de uma pessoa com conhecimentos gerais na técnica.

Numa forma de realização da presente invenção, uma composição vacinal da invenção pode ser baseada numa glicoproteína compreendendo um componente ou fragmento imunogénico de um polissacárido de *S. aureus* da invenção e, opcionalmente, compreendendo ainda um veículo ou adjuvante farmacêuticamente aceitável. Em formas de realização adicionais da presente invenção, uma composição vacinal pode ser baseada numa glicoproteína compreendendo um componente ou fragmento imunogénico de uma proteína de *S. aureus* da invenção e, opcionalmente, compreendendo ainda um veículo ou adjuvante farmacêuticamente aceitável. Ainda noutro aspeto adicional da invenção, uma composição vacinal pode ser baseada numa glicoproteína compreendendo um componente ou fragmento imunogénico de uma proteína de *P. aeruginosa* da invenção e, opcionalmente, compreendendo ainda um veículo e/ou adjuvante farmacêuticamente aceitável.

É bem conhecido daqueles com conhecimentos gerais na técnica como modificar uma vacina para administração a um tipo de mamífero, por exemplo, murganhos, para administração a outro tipo de mamífero, por exemplo, humanos. Por exemplo, um especialista saberia facilmente que a eliminação da cauda de

histidina do veículo proteico de uma glicoproteína utilizada numa composição vacinal em murganhos tornaria a glicoproteína adequada para administração numa composição vacinal em humanos. Por exemplo, a eliminação da cauda de HISTIDINA (cauda de HIS) em transportadores proteicos, tais como, e. g., EPA (SEQ ID N°: 13), ClfA (SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12) e Hla (SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 16), seria reconhecida pela sua utilização numa glicoproteína para administração a um humano.

Deve ser compreendido que o melhoramento de qualquer dos sintomas de uma Gram-positiva, por exemplo, *S. aureus*, ou outra infecção bacteriana ou doença é um objetivo clínico desejável, incluindo um atenuar da dosagem de medicação utilizada para a infecção ou doença causada por uma Gram-positiva, por exemplo, uma infecção ou doença causada por *S. aureus*, ou outra infecção ou doença de causa bacteriana ou um aumento na produção de anticorpos no soro ou muco de doentes. Será evidente aos especialistas na técnica que algumas das composições vacinais da divulgação são úteis para a prevenção de uma infecção por Gram-positiva, por exemplo, uma infecção por *S. aureus*, ou outra infecção bacteriana, algumas são úteis para o tratamento de uma infecção Gram-positiva, por exemplo, uma infecção por *S. aureus*, ou outra infecção bacteriana e algumas são úteis para a prevenção e o tratamento dessas infecções.

As formas de realização da presente invenção, tais como vacinas e outros agentes farmacêuticos podem, opcionalmente, ser preparadas utilizando veículos, excipientes, diluentes e/ou adjuvantes adequados e farmaceuticamente aceitáveis, como são bem conhecidos na técnica e evidentes à luz desta descrição. Um excipiente, diluente ou adjuvante pode ser um material sólido,

semissólido ou líquido que pode servir como um veículo ou meio para o ingrediente ativo. À luz desta descrição, alguém com conhecimentos gerais na técnica no campo da preparação de composições, pode facilmente selecionar a forma e modo de administração apropriados, dependendo das características particulares do produto selecionado, da doença ou estado a ser tratado, da fase da doença ou estado e outras circunstâncias relevantes (Remington's Pharmaceutical Sciences, Mack Publishing Co. (1990)). A proporção e natureza do diluente, excipiente ou adjuvante farmacêuticamente aceitável são determinadas pela solubilidade e propriedades químicas do composto farmacêuticamente ativo selecionado da via de administração escolhida e da prática farmacêutica corrente.

Consequentemente, em formas de realização da invenção, as composições vacinais compreendem componentes ou fragmentos imunogênicos, e. g., polissacárido ou fragmento deste de *S. aureus* e/ou proteína ou fragmento desta de *S. aureus* ou *P. aeruginosa* e, opcionalmente, incluem um veículo farmacêuticamente aceitável. A expressão "veículo farmacêuticamente aceitável" refere-se a um veículo que é não tóxico. Os veículos farmacêuticamente aceitáveis adequados incluem, por exemplo, um ou mais de água, soro fisiológico, solução salina tamponada com fosfatos, dextrose, glicerol, etanol, e semelhantes, assim como as suas combinações. Os veículos farmacêuticamente aceitáveis podem ainda compreender quantidades menores de substâncias auxiliares, tais como agentes molhantes ou emulsionantes, conservantes ou tampões, que melhoram a estabilidade em armazenamento ou eficácia do anticorpo. Esses veículos farmacêuticamente aceitáveis incluem, por exemplo, diluentes líquidos, semissólidos ou sólidos que servem como veículos, excipientes ou meios farmacêuticos. Pode

ser utilizado qualquer diluente conhecido na técnica. Diluentes exemplificativos incluem, mas não estão limitados a polioxietileno monolaurato de sorbitano, estearato de magnésio, metil- e propil-hidroxibenzoato, talco, alginatos, amidos, lactose, sacarose, dextrose, sorbitol, manitol, goma de acácia, fosfato de cálcio, óleo mineral, manteiga de cacau e óleo de teobroma.

Além disso, em formas de realização adicionais da invenção, a composição vacinal pode, opcionalmente, incluir um adjuvante ou uma combinação de adjuvantes, incluindo, mas não limitados a adjuvantes particulados, tais como sais de alumínio (hidróxido de alumínio, fosfato de alumínio, sulfato de hidroxifosfato de alumínio, etc.); emulsões, tal como óleo em água (MF59, AS03); combinações de lípido e sal, tal como AS04; água em óleo (Montanida); ISCOMS, lipossomas/virossomas; nanopartículas e micropartículas, etc.; não particulados, tal como péptidos; saponinas (QS21); MPL A; citocinas; derivados de ADN; toxinas bacterianas; etc. Uma forma de realização adicional inclui adjuvantes utilizados em animais, tais como Adjuvante Completo de Freund e Adjuvante Incompleto de Freund, adjuvantes baseados em micolato (e. g., dimicolato de trealose), lipopolissacárido bacteriano (LPS), peptidoglicanos (*i. e.*, mureínas, mucopéptidos ou glicoproteínas, tais como N-Opaca, dipéptido de muramilo [MDP] ou análogos de MDP), proteoglicanos, preparações estreptocócicas (e. g., OK432), DEAE-dextrano, óleos neutros (tal como migliol), óleos vegetais (tal como óleo de amendoim), Pluronic, o sistema adjuvante de Ribí ou interleucinas, particularmente aquelas que estimulam a imunidade de mediação celular. O adjuvante utilizado dependerá, em parte, da composição e tipo da vacina de glicoconjugado. A quantidade de adjuvante a administrar vai depender do tipo e estatura do

mamífero. As dosagens ótimas podem ser facilmente determinadas por métodos de rotina.

Um aspeto adicional da presente invenção refere-se a uma composição farmacêutica, compreendendo, pelo menos, uma glicoproteína de acordo com invenção. A preparação de medicamentos compreendendo glicoproteínas é bem conhecida na técnica. O esquema de preparação para a composição farmacêutica final e o modo e pormenores da sua administração dependerão da proteína, da célula hospedeira, do ácido nucleico e/ou do vetor empregues.

Será evidente aos especialistas na técnica que a quantidade terapeuticamente eficaz de polissacárido ou glicoproteína desta invenção dependerá, *inter alia*, do programa de administração, da dose unitária de anticorpo administrado, se o polissacárido ou glicoproteína é administrado em combinação com outros agentes terapêuticos, do estado imunitário e saúde do doente e da atividade terapêutica do polissacárido ou glicoproteína particular.

As composições vacinais e/ou preparações farmacêuticas da invenção podem ser adaptadas para utilização oral, parentérica ou tópica e podem ser administradas ao doente na forma de comprimidos, cápsulas, supositórios, solução, suspensões ou qualquer outro meio ou forma de dosagem adequado. Em aspetos adicionais da invenção, as composições vacinais e/ou preparações farmacêuticas podem ser introduzidas no indivíduo a ser imunizado por qualquer método conhecido, incluindo, e. g., por injeção intravenosa, intradérmica, intramuscular, intramamária, intraperitoneal ou subcutânea; ou por distribuição oral, sublingual, nasal, anal ou vaginal. Os compostos

farmaceuticamente ativos da presente invenção, embora eficazes por si só, podem ser formulados e administrados na forma dos seus sais farmaceuticamente aceitáveis, tais como sais de adição de ácido ou sais de adição de base, para fins de estabilidade, conveniência de cristalização, solubilidade aumentada e semelhantes. As composições vacinais, numa forma de realização da invenção, são administradas parentericamente, e. g., por injeção, subcutaneamente ou intramuscularmente. Os métodos para imunização intramuscular são descritos por Wolff *et al.* (1990) *Science* 247: 1465-1468 e por Sedegah *et al.* (1994) *Immunology* 91: 9866-9870. Outros modos de administração incluem oral e transdérmica.

As vacinas da invenção podem ser administradas como um agente profilático primário, e. g., adultos ou em crianças, como uma prevenção secundária, após erradicação com sucesso de bactérias Gram-positivas, tal como *S. aureus*, num hospedeiro infetado, ou como agente terapêutico com o objetivo de induzir uma resposta imunitária num hospedeiro para prevenir infeção por uma bactéria Gram-positiva, tal como *S. aureus*. As vacinas da invenção são administradas em quantidades facilmente determinadas por pessoas com conhecimentos gerais na técnica. O tratamento pode consistir numa dose única ou numa pluralidade de doses, ao longo de um período de tempo. Por exemplo, em algumas formas de realização, é esperado que uma dosagem típica para humanos de uma vacina da presente invenção seja cerca de 1 a 25 µg do antigénio oligossacárido, o qual estará ligado ao (e não inclui a massa do) veículo proteico, em formas de realização adicionais cerca de 1 µg a cerca de 10 µg do antigénio polissacárido e ainda em formas de realização adicionais cerca de 2 µg do antigénio polissacárido. Em formas de realização adicionais, a razão de açúcar/proteína no glicoconjugado ou na

vacina é cerca de 1:5 a cerca de 1:10. Opcionalmente, uma vacina, tal como uma vacina de bioconjugado da presente invenção, pode incluir um adjuvante. Os especialistas na técnica reconhecerão que a dose ótima pode ser maior ou menor, dependendo do peso corporal do doente, doença, da via de administração e de outros fatores. Os especialistas na técnica também reconhecerão que os níveis de dosagem apropriados podem ser obtidos com base em resultados com vacinas conhecidas. O número de doses dependerá da doença, da formulação e dados de eficácia de ensaios clínicos.

As composições vacinais podem ser embaladas em formas convenientes para distribuição. São preferidas as formas de distribuição compatíveis com a entrada do componente ou fragmento imunogénico no mamífero recetor.

Uma forma de realização da divulgação é geralmente dirigida à produção recombinante de uma vacina para um organismo Gram-positivo num organismo Gram-negativo, por utilização de uma via biossintética de LPS modificada. Isto é alcançado por inserção num hospedeiro que compreende ácidos nucleicos que codificam para uma oligossacariltransferase e uma proteína e ácidos nucleicos que codificam para glicosiltransferases tendo origem em, pelo menos, dois organismos diferentes. Esta forma de realização é dirigida à manipulação genética de um organismo com base num organismo natural dentro do qual são inseridos ácidos nucleicos que codificam para (i) uma proteína; (ii) uma oligossacariltransferase e (iii) glicosiltransferases de, pelo menos, dois organismos diversos.

Num exemplo dessa forma de realização, é produzido um produto de proteína glicosilada para utilização como uma vacina

para *Staphylococcus aureus*. Os produtos vacinais da invenção são produzidos num hospedeiro de *E. coli* geneticamente modificado. *S. aureus* é uma bactéria Gram-positiva e tem uma cápsula de polissacárido. Um produto vacinal para este organismo poderia ser baseado numa proteína glicosilada cuja secção de açúcar tivesse uma estrutura semelhante a este polissacárido capsular.

Noutro aspeto, a presente divulgação é dirigida a uma nova abordagem de biomanipulação para produção de vacinas de conjugados imunogénicos que proporcionam vantagens relativamente aos métodos clássicos de conjugação química. Numa forma de realização, a abordagem envolve a produção *in vivo* de glicoproteínas em células bacterianas, por exemplo, células Gram-negativas, tal como *E. coli*.

Como conhecido de uma pessoa com conhecimentos gerais na técnica, a produção e purificação de glicoconjugado pode variar, dependendo do candidato vacinal e da combinação de plasmídeos utilizados. Por exemplo, que processo de purificação escolher é conhecido com base no veículo proteico, no componente de açúcar do glicoconjugado e na utilização pretendida do candidato vacinal purificado, por exemplo, em animais ou humanos. Para utilização em humanos, por exemplo, é conhecido que a cauda de HIS, a qual de outro modo facilitaria a purificação, seria removida.

Deve ser compreendido que o termo "ou", como aqui utilizado, denota alternativas que podem, onde apropriado, ser combinadas; ou seja, o termo "ou" inclui cada alternativa listada separadamente, assim como a sua combinação. Como aqui utilizado, salvo claramente indicado em contrário pelo contexto, as referências ao singular, tais como as formas singulares "um",

uma" e "o", incluem o plural e as referências ao plural incluem o singular.

A invenção é ainda definida por referência aos seguintes exemplos que descrevem adicionalmente as composições e métodos da presente invenção, assim como a sua utilidade.

Exemplos

Exemplo 1: Síntese de polissacárido CP5 e CP8 em células de *E. coli*

Um objetivo de uma forma de realização da invenção é produzir os polissacáridos antigénicos CP5 e CP8 em *E. coli*. Como discutido acima, a requerente explorou de um novo modo, surpreendente em consideração da técnica anterior, o facto de que as vias de produção de CP e antígeno O se sobrepõem funcionalmente, um facto que está representado na estrutura da RU (Ver as FIG. 1-4). Os glicanos capsulares de CP5 e CP8 são polímeros consistindo de RU de trissacáridos semelhantes de Ácido 2-acetamido-2-desoxi-D-manurónico (D-ManNAcA) e dois resíduos de 2-Acetamido-2,6-didesoxigalactose com configurações D e L (D- e L-FucNAc). Os resíduos de ManNAcA estão ligados de um modo diferente nos dois serotipos; adicionalmente, a ligação entre as RU no glicano polimerizado é diferente. Além disso, existe uma modificação O-acetilo imunodominante em diferentes posições nos dois antígenos (Jones, C. 2005. Estruturas revistas para os polissacáridos capsulares de *Staphylococcus aureus* tipos 5 e 8, componentes de novas vacinas de glicoconjugado. *Carbohydr Res* 340:1097-106). O antígeno O11 de LPS de *P. aeruginosa* é semelhante na sua estrutura a CP5 e CP8,

dado que o antígeno O11 de *P. aeruginosa* LPS contém [-3)- α -L-FucNAc-(1,3)- β -D-FucNAc-(1,2)- β -D-Glc-(1-] (FIG. 4). (Knirel, Y. A., V. V. Dashunin, A. S. Shashkov, N. K. Kochetkov, B. A. Dmitriev e I.L. Hofman. 1988. Antígenos somáticos de *Shigella*: estrutura da cadeia polissacárida específica de O do lipopolissacárido de *Shigella dysenteriae* tipo 7. *Carbohydr Res* 179: 51-60). As RU trissacáridas diferem apenas na medida em que o D-ManNAcA de *S. aureus* está substituído por uma unidade de glucose, não existe modificação O-acetilo em LPS de *P. aeruginosa* O11 e a diferença no tipo de ligação entre o 2° e 3° monossacárido na RU (FIG. 4).

Para produzir um sistema genético capaz de sintetizar os glicanos de CP5 e CP8 em UndPP, utilizando o método de Dean *et al.*, (Dean, C. R., C. V. Franklund, J. D. Retief, M. J. Coyne, Jr., K. Hatano, D. J. Evans, G. B. Pier e J. B. Goldberg. 1999. Caracterização do locus do antígeno O do serogrupo O11 de *Pseudomonas aeruginosa* PA103. *J Bacteriol* 181:4275-4284), a requerente modificou o aglomerado gênico do antígeno O de *P. aeruginosa* O11 da estirpe PA103. Os genes que codificam a maquinaria biossintética para a síntese da estrutura tronco consistindo de UndPP-D-FucNAc-L-FucNAc foram complementados com as enzimas de *S. aureus* requeridas para a conclusão do glicano de *S. aureus* (FIG. 1-4), o que também constituiu uma nova utilização deste processo. Deste modo, utilizando o método de Dean *et al.*, todos os elementos genéticos de *P. aeruginosa* PA103 requeridos para a biossíntese de UndPP-FucNAc-FucNAc foram expressos. O gene que codifica a glicosiltransferase que adiciona o terceiro açúcar foi eliminado e substituído pelos correspondentes genes dos aglomerados de cap5 ou 8 de *S. aureus* Mu50 (CP5) e MW2 (CP8), com ligeiras modificações.

Os genes que codificam as enzimas que sintetizam os resíduos específicos para o polissacárido capsular de *S. aureus* foram integrados, passo a passo, no contexto de O11 de acordo com as funções dos genes previstos por Sau et al. (Sau, S., N. Bhasin, E. R. Wann, J. C. Lee, T. J. Foster e C. Y. Lee. 1997. Os loci genéticos alélicos de *S. aureus* para a expressão da cápsula de serotipo 5 e 8 contêm os genes específicos de tipo flanqueados por genes comuns. *Microbiology* 143: 2395-405; O'Riordan, K. e J. C. Lee. 2004. Polissacáridos capsulares de *Staphylococcus aureus*. *Clin Microbiol Rev* 17(1): 218-34). Tais passos são explicados abaixo.

Foi predado que o produto génico *cap5I/cap8H* era a glicosiltransferase que adiciona o ManNAcA a UndPP-D-FucNAc-L-FuncNAc da RU que forma uma ligação específica para cada serotipo (Sau, S., N. Bhasin, E. R. Wann, J. C. Lee, T. J. Foster e C. Y. Lee. 1997. Os loci genéticos alélicos de *Staphylococcus aureus* para a expressão da cápsula de serotipo 5 e 8 contêm os genes específicos de tipo flanqueados por genes comuns *Microbiology* 143: 2395-405). Para provar isto, a atividade de Cap5I e Cap8H foi analisada em *E. coli* na presença de um plasmídeo conferindo a produção do antígeno O de *P. aeruginosa* O11. Células que expressam o aglomerado de O11 sintetizam o antígeno O de O11 primeiro em UndPP, de onde é transferido para o núcleo de lípido A pela enzima WaaL de *E. coli*, a ligase de antígeno O, formando o lipopolissacárido específico de O11 (LPS) (Goldberg, J. B., K. Hatano, G. S. Meluleni e G. B. Pier. 1992. Clonagem e expressão de superfície de antígeno O de *Pseudomonas aeruginosa* em *Escherichia coli*. *Proc Natl Acad Sci U S A* 89(22): 10716-20). Para sintetizar este lipopolissacárido, o aglomerado de antígeno O de O11 de *P. aeruginosa* PA103 foi clonado em pLAFR1 (SEQ ID N°: 1).

Depois, o gene *wbjA* que codifica a glucosiltransferase, a enzima que adiciona o terceiro açúcar à RU de O11, foi eliminado por mutagénese de transposição. O aglomerado mutado (O11 *wbjA::Tn50<dhfr-1>*) foi ainda modificado por recombinação homóloga, para eliminar a atividade de polimerase do gene *wzy*, formando O11 *wbjA::Tn50<dhfr-1>wzy::cat*, o que denota a SEQ ID N°: 1 mutada, em que os genes para a glicosiltransferase *wbjA* e a polimerase *wzy* do aglomerado génico de O11 foram inativados. Este aglomerado modificado foi expresso em células W3110 Δ wecA, os extratos foram tratados com proteinase K e analisados por SDS PAGE e coloração de prata, de acordo com o método divulgado em Tasi, *et al.* (Tsai, C. M. e C. E. Frasch. 1982. Uma coloração de prata sensível para deteção de lipopolissacáridos em géis de poliacrilamida. *Anal Biochem* 119:115-9). Os resultados são proporcionados na FIG. 5A, mostrando a coloração de prata dos extratos de W3110 Δ wecA que expressam o aglomerado de O11 mutado de pLAFR1, como aqui descrito. A segunda linha indica os genes expressos do plasmídeo indutível pEXT22. Os asteriscos indicam genes sintetizados e codão-otimizados. Diferentes glicoformas relevantes estão indicadas com setas).

A análise resultou em duas bandas principais nos géis (FIG. 5A, pista 1). Os sinais correspondem ao núcleo de lípido A não modificado (FIG. 5A, banda inferior) e LPS consistindo de núcleo de lípido A e dois resíduos de FucNAc, como esperado numa RU de O11 truncada. Após expressão de uma cópia de tipo selvagem de *wbjA* de um plasmídeo separado, indutível por IPTG, a banda superior mudou para uma mobilidade eletroforética mais lenta, indicando a adição de um resíduo de glucose ao LPS de O11 truncado (FIG. 5A, pista 2). Quando as glicosiltransferases de *S. aureus* previstas Cap5I (pista 4) e Cap8H (FIG. 5A, pista 3)

foram expressas em trans, em vez de WbjA, foi observada uma mudança semelhante do sinal do núcleo de lípido A glicosilado, indicativa de adição de um monossacárido, possivelmente até maior que a glucose, sendo muito provavelmente ManNAcA. Estes dados provam que as glicosiltransferases de *S. aureus* podem alongar o glicolípido UndPP-D-FucNAc-L-FuncNAc que foi sintetizado por atividade de enzimas de *P. aeruginosa*.

Deste modo, também foi confirmado que um pré-requisito para o agrupamento da RU de *S. aureus* em *E. coli* é a provisão de UDPManNAcA, porque a maquinaria biossintética está presente nos aglomerados de CP5/8 de *S. aureus* mas não no aglomerado de antígeno O de O11 de *P. aeruginosa*. Todos os outros açúcares ativados por nucleótido requeridos são proporcionados por funções de gestão de *E. coli* e do aglomerado de antígeno O de O11 de *P. aeruginosa*. Sabe-se que a *E. coli* produz UDP-ManNAcA, o substrato para a glicosiltransferase de ManNAcA, por expressão de *wecB* e *wecC*. Aqueles genes são expressos de modo constitutivo no aglomerado responsável pela biossíntese do antígeno enterobacteriano comum (ECA) (Meier-Dieter, U., R. Starman, K. Barr, H. Mayer e P. D. Rick. 1990. Biossíntese do antígeno enterobacteriano comum em *Escherichia coli*. *J Biol Chem* 265:13490-13497). Verificou-se que o homólogo funcional para a biossíntese de UDP-ManNAcA encontrado no aglomerado de CP de *S. aureus* complementa a atividades de *wecBC*, como referido anteriormente (Kiser, K. B., N. Bhasin, L. Deng e J. C. Lee. 1999. O cap5P de *Staphylococcus aureus* codifica uma UDP-N-acetilglucosamina 2-epimerase com redundância funcional. *J. Bacteriol* 181(16): 4818-24). Isto mostra que a produção dos antígenos de CP em *E. coli* depende da expressão funcional dos genes *wecBC* na estirpe hospedeira. Assim, para proporcionar UDP-ManNAcA como substrato para Cap5I e Cap8H num sistema

recombinante, foi confirmado que WecB e WecC têm que ser expressos. Num tal sistema, pode ser utilizada qualquer estirpe procariótica que expressam o antigénio enterobacteriano comum como estirpe de *E. coli* de tipo selvagem, e. g., tipos de células baseados em W3110, com ou sem uma eliminação de *wecA* e com ou sem eliminação adicional de *wzzE*.

Pensa-se que, para atividade imunológica máxima do glicano, o alongamento adicional do polissacárido capsular de *S. aureus* seja requerido. Os genes *cap5J/cap8I* codificam os homólogos *wzy* que polimerizam as unidades de repetição e *cap5K/cap8K* codificam a flipase que transloca o trissacárido ligado a UndPP do lado citoplasmático para o lado periplasmático da membrana. *Cap5H/cap8J* codificam a O-acetiltransferase que modifica o L-FucNAc na posição 3' ou o ManNAcA na posição 4' da RU (Bhasin, N., A. Albus, et al. (1998). "Identificação de um gene essencial para a O-acetilação do polissacárido capsular de *Staphylococcus aureus* tipo 5". *Mol Microbiol* 27(1): 9-21. A acetilação é um determinante importante que discrimina a reatividade imunológica do polissacárido (Fattom, A. I., J. Sarwar, L. Basham, S. Ennifar e R. Naso. 1998. Determinantes antigénicos de vacinas de polissacárido capsular tipo 5 e tipo 8 de *S. aureus*. *Infect Immun* 66:4588-92). Para mostrar que as RU poderiam ser alongadas e acetiladas, as enzimas de *S. aureus* responsáveis pela polimerização e O-acetilação foram expressas de plasmídeos separados na presença do aglomerado de O11 mutado. Os extratos de células W3110 Δ *wecA* que expressam o aglomerado *wbjA::Tn50<dhfr-1> wzy::cat* de O11 e diferentes genes do aglomerado de CP5 foram tratados com proteinase K e analisados por SDS PAGE, eletrotransferência seguida de imunotransferência utilizando um açúcar anti-CP5 (obtido de J. C. Lee no Departamento de Medicina, Brigham and Women's Hospital,

Faculdade de Medicina de Harvard, Boston, MA, EUA). A FIG. 5B mostra os resultados da imunodeteção dos extratos de *E. coli* tratados com proteinase K separados por SDS PAGE e eletrotransferidos que utilizam o antissoro anti-CP5. Todos os extratos analisados continham um aglomerado de *P. aeruginosa* O11 com eliminações dos genes *wbjA* e parcialmente (indicado por um asterisco) *wzy* expressos a partir do plasmídeo pLAFR, como aqui descrito, e mais dois plasmídeos (pEXT22, pACT3) expressando diferentes proteínas Cap5 (como indicado) que possibilitam a polimerização e O-acetilação de CP5 nestas células. Estão indicados pormenores experimentais, tais como concentrações de indutor e temperaturas de incubação de cultura de expressão.

Na FIG. 5B, os resultados mostram sinais semelhantes a escada, típicos para um polímero de antígeno O numa gama de peso molecular mais elevado. As diferentes bandas representam diferentes números de RU polimerizadas linearmente em LPS ou em UndPP, ambos os quais são estáveis à digestão de proteinase K. Foram observadas diferentes intensidades da estrutura semelhante a escada na presença ou ausência de O-acetiltransferase. Enquanto fortes sinais foram detetados na presença de *cap5H* (FIG 5B, pistas 1-4), estes estavam virtualmente ausentes em pistas sem *cap5H* (FIG 5B, pistas 5, 6). Isto significa que a O-acetilação aumenta o reconhecimento pelo antissoro específico, ou que melhora a atividade de polimerização, acelerando o deslocamento ou tornando a polimerização como tal mais eficiente ou induzindo mais produção de RU. O gene *cap5H* é funcional quando expresso a partir de diferentes plasmídeos de estrutura (FIG 5B, pistas 1, 2 e 3, 4), embora a intensidade de sinal seja mais forte quando o *cap5H* é expresso isoladamente a partir de um plasmídeo separado (comparar a pista 1 com a pista 3 da FIG. 5B e pista 2 com pista 4 da FIG. 5B). É surpreendente e notável que

quanto menos IPTG era utilizado para indução dos genes de *S. aureus*, mais fortes os sinais (comparar a pista 1 com a pista 2 da FIG. 5B e a pista 3 com a pista 4 da FIG. 5B).

Exemplo 2: Síntese de polímero CP5 e CP8 em lípido em células de *E. coli*

Dado que a expressão elevada dos genes específicos de *cap5* conduz a menor formação de polímero, foi construído um sistema de expressão alternativo para os glicanos recombinantes para abordar este problema. Em pormenor, numa nova abordagem, inesperada à luz da técnica anterior, a glucosiltransferase (*wbjA*) de *P. aeruginosa* e a polimerase (*wzy*) de O11 foram substituídas pelos genes que codificam os elementos específicos de CP5/8 do aglomerado génico capsular de *S. aureus* Mu50/MW2 (*cap5/8HIJK* e partes destes), produzindo um aglomerado génico único, quimérico, composto pelos genes CP5 ou CP8 de *P. aeruginosa* O11 e *S. aureus* (FIG. 6). A construção continha os genes específicos de *S. aureus*. Cada foi marcado para deteção de expressão e cada continha um sítio de ligação ribossómica introduzido e era seguido de uma cassete de resistência a cloranfenicol (*cat*) para seleção de clones recombinados, resultando nas SEQ ID N°: 2, SEQ ID N°: 3 e SEQ ID N°: 4, de acordo com o método de Datsenko, *et al.* (Datsenko, K. A. e B. L. Wanner. 2000. Inativação de um passo de genes cromossómicos em *Escherichia coli* K-12 utilizando produtos de PCR. *Proc Natl Acad Sci U S A* 97:6640-5).

A FIG. 6 representa uma forma de realização de uma estratégia da presente invenção para a construção dos aglomerados génicos O11/CP55 e O11/CP8 quiméricos da presente

invenção. Os aglomerados de CP de *S. aureus* CP5 e CP8 (topo) e o aglomerado *rfb* de *P. aeruginosa* PA103 (O11, meio) estão representados como publicado (Dean, C. R., C. V. Franklund, J. D. Retief, M. J. Coyne, Jr., K. Hatano, D. J. Evans, G. B. Pier e J. B. Goldberg. 1999. Caracterização do lócus do antígeno O do serogrupo O11 de *Pseudomonas aeruginosa* PA103. *J Bacteriol* 181:4275-84; (Sau, S., N. Bhasin, E. R. Wann, J. C. Lee, T. J. Foster e C. Y. Lee. 1997. Os loci genéticos alélicos de *S. aureus* para a expressão da cápsula de serotipo 5 e 8 contêm os genes específicos de tipo flanqueados por genes comuns. *Microbiology* 143 (Pt 7): 2395-405). As funções homólogas dos genes estão descritas abaixo. Diagonais para a frente completas indicam os genes responsáveis pela síntese do dissacárido D-FucNAc-L-FucNAc em UndPP nos dois organismos; pontos indicam os genes de glicosiltransferase que adicionam o terceiro monossacárido à RU. Os genes de flipase semelhantes a *Wzx* estão indicados por diagonais para a frente a tracejado, os genes da polimerase de RU semelhantes a *wzy* estão indicados por diagonais para trás a tracejado. O aglomerado de CP5 não contém um regulador de comprimento de *Wzz* (seta vazia), mas um conjunto de três genes compondo a maquinaria de exportação para polissacárido capsular que inclui a função reguladora de comprimento em *S. aureus* (setas vazias). O gene da transferase O-acetilo, indicado por diagonais para a frente completas, é único ao aglomerado de CP. Os genes requeridos para a biossíntese de UDP-ManNAcA em *S. aureus* estão indicados a negro. Não são requeridos para a produção do antígeno O de *P. aeruginosa*. Os genes responsáveis pelas diferenças estruturais dos polissacáridos O11, CP5 e CP8 estão aglomerados uns aos outros no início (O11: *wbjA* e *wzy*) ou meio (CP5/8: *cap5/8HIJK*) dos respectivos aglomerados génicos. O aglomerado de CP8 é quase idêntico ao aglomerado de CP5 considerando

comprimento e sequência de ADN, exceto na parte do meio (*cap5/8HIJK*) que confere especificidade estrutural. O aglomerado quimérico foi construído por substituição dos genes *wbjA* e *wzy* de um aglomerado de O11 gerado em plasmídeo, com a parte da especificidade do aglomerado de CP5 (ou CP8) (*cap5/8HIJK*) e uma cassette de cloranfenicol acetiltransferase representada pela seta vazia rotulada *cat* (*cat*, para seleção), por recombinação homóloga e clonagens clássicas, resultando nas SEQ ID N°: 2, SEQ ID N°: 3 e SEQ ID N°: 4. Os asteriscos nas setas a tracejado indicam sequências génicas incompletas utilizadas para recombinação homóloga. Os dois aglomerados quiméricos resultantes estão mostrados no painel inferior, representando o ADN das SEQ ID N°: 3 e SEQ ID N°: 4.

Para provar que os CP5 e CP8 quiméricos da presente invenção agrupam, surpreendentemente, a RU correta em UndPP e garantem que as unidades de repetição são polimerizadas, a digestão de proteinase K de células de *E. coli* (W3310 Δ *wecA*) contendo os aglomerados quiméricos de comprimento completo foi separada por SDS-PAGE. Especificamente, células com um plasmídeo contendo ou desprovido do aglomerado génico de CP5 quimérico (FIG. 7A) ou o aglomerado génico de CP8 quimérico (FIG. 7B) no plasmídeo pLAFR foram tratadas com Proteinase K, separadas por SDS-PAGE e os lípidos foram visualizados por coloração de prata (painel da esquerda nas FIG. 7A e 7B) ou imunodeteção com antissoro anti-CP5 ou anti-CP8 após eletrotransferência para membranas de nitrocelulose (painel da direita na FIG. 7A e B)). As construções desprovidas (SEQ ID N°: 2) e contendo (SEQ ID N°: 3) o gene *cap5K* de flipase foram testadas. Verificou-se que a última era menos ativa na produção de LPS de CP5.

Após eletrotransferência e imunodeteção com soro específico anti-CP5, extratos expressando os aglomerados de CP5 quimérico integrais mostram um sinal semelhante a escada semelhante às estruturas de antígeno O endógenas de *E. coli* sondadas com o seu soro autólogo (FIG. 7A, últimas duas pistas à direita). Isto sugere fortemente que as unidades de repetição de CP5 são polimerizadas, que existe um comprimento de polímero preferido e que o antígeno de CP5 é transferido para núcleo de lípido A nestas células. Os mesmos extratos foram visualizados por coloração de prata após SDS PAGE (FIG. 7A, no lado esquerdo da figura, as duas pistas à direita rotuladas como: CP5 quimérico (sem cap5K) e CP5 quimérico mostrando que o LPS é, de facto, formado, consistindo do núcleo de lípido A de *E. coli* decorado com a estrutura semelhante a antígeno O de CP5. Diferenças de intensidade foram obtidas de extratos que têm origem em células que expressavam o aglomerado quimérico de CP5 com ou sem o gene de flipase *cap5K*. A comparação dos dois extratos mostra que a expressão de Cap5K aumenta consideravelmente a produção do polímero (comparar as pistas do meio e da direita em ambos os painéis da FIG. 7A).

Como mostrado na FIG. 7B, os mesmos resultados foram observados com um aglomerado quimérico de CP8. Células contendo um plasmídeo contendo ou desprovido do aglomerado génico de CP8 quimérico no plasmídeo pLAFR foram tratadas com Proteinase K, separadas por SDS PAGE e os lípidos foram detetados por coloração de prata (painéis da esquerda) ou imunodeteção com antissoro anti-CP8, após eletrotransferência para membranas de nitrocelulose (painel da direita). A construção quimérica de CP8 contendo o gene de flipase *cap8K* corresponde à SEQ ID N°: 4.

Uma extensão nova e surpreendente adicional da invenção foi desenvolvida pela alteração das estruturas do plasmídeo utilizadas para manutenção e expressão do aglomerado quimérico em *E. coli*. A cassette de resistência em pLAFR1 contendo o aglomerado de CP5 quimérico foi alterada de Tet para Kan. Adicionalmente, o aglomerado quimérico de CP5 integral contendo o *cap5K* foi subclonado no plasmídeo pDOC-C, de acordo com método de Lee *et al.* (Lee, D. J., L. E. Bingle, K. Heurlier, M. J. Pallen, C. W. Penn, S. J. Busby e J. L. Hobman. 2009. Adulteração génica: um método para manipular de modo recombinante em estirpes de *Escherichia coli* de laboratório e patogénicas. *BMC Microbiol* 9: 252) e pACYC177 (acesso GeneBank N° X06402).

Como mostrado nas FIG. 8A e 8B, todos estes plasmídeos conferiram produção do polímero CP5, como analisado por SDS PAGE, eletrotransferência e imunodeteção com antissoro específico de anti-CP5. Na FIG 8A, extratos de células totais de células contendo diferentes aglomerados quiméricos foram tratadas com Proteinase K e analisadas por SDS PAGE e coloração de prata. Estão indicados os plasmídeos que contêm diferentes genes específicos de *S. aureus* e diferentes genes de resistência, utilizados para seleção de antibiótico: Tetraciclina (Tet) e HIJ, SEQ ID N°: 2; Tet HIJK, SEQ ID N°: 3, Tet e sem genes, controlo de plasmídeo vazio, os números correspondem a marcadores de peso molecular. As pistas rotuladas Canamicina (Kan) contêm uma variação da SEQ ID N°: 3, na qual a cassette de resistência a tetraciclina é substituída por um gene de resistência à canamicina.

Na FIG. 8B, a estirpe hospedeira foi *E. coli* W3110 Δ wecA, como na FIG. 8A. A pista da esquerda na FIG. 8B corresponde ao

marcador de peso molecular como na FIG. 8A. Na FIG. 8B, extratos de células totais de células contendo diferentes aglomerados quiméricos foram tratadas com Proteinase K e analisadas por SDS PAGE e coloração de prata (painel da esquerda) e por imunotransferência anti-CP5 após eletrotransferência (painel da direita). Os plasmídeos utilizados contêm o aglomerado de CP5 quimérico indicado na SEQ ID N°: 3, presente numa estrutura do plasmídeo pLAFR1 modificada contendo uma cassete de Canamicina, em vez de tetraciclina (ver a FIG. 8A), ou em pACYC contendo uma cassete de resistência a cloranfenicol.

Além disso, foram testados diferentes promotores para expressar o LPS de CP5 de O11 quimérico. Nestes testes, a estirpe hospedeira era *E. coli* W3110 Δ wecA transportando o aglomerado de CP5 quimérico. Nesta estirpe, o aglomerado quimérico substituiu os genes *wecAwzzE*. Extratos de células totais de células contendo diferentes aglomerados quiméricos expressos a partir de pLAFR1 foram tratados com Proteinase K e analisados por SDS PAGE e imunotransferência anti-CP5 após eletrotransferência. Os plasmídeos continham aglomerados de O11, onde *wbjA* e *wzy* foram substituídos por diferentes genes de especificidade de *S. aureus* (com uma cassete de *cat*), como indicado abaixo nas pistas na FIG. 9. Além disso, o ADN em frente dos genes de especificidade de *cap5* foi alterado e o efeito na glicosilação lipídica foi analisado. O efeito destas regiões promotoras diferentes foi analisado, como representado na FIG. 9. *Wzz/wzx* denota os genes originais (ver a FIG. 6) em frente dos genes *cap* após a recombinação homóloga inicial (FIG. 9 correspondente às duas primeiras pistas). Estes dois genes foram removidos (-) (FIG. 9 correspondente às três pistas no meio) e substituídos com a região de 0,6 kb (P0121) (FIG. 9 correspondente às três últimas pistas) presente a montante do

aglomerado do antígeno O de *E. coli* 0121 que codifica uma forte sequência promotora. As pistas denotadas wzz/wzx e HIJ na FIG. 9 foram derivadas de células expressando a SEQ ID N°: 2, as pistas denotadas wzz/wzx e HIJK derivam da SEQ ID N°: 3. Na FIG. 9, os marcadores de peso molecular estão indicados à esquerda da moldura do gel.

Como indicado na FIG. 9, os resultados mostraram que uma atividade promotora relevante reside no gene *wzx* (FIG. 9, primeiras duas pistas- wzz/wzx) e que pode ser funcionalmente substituída por um promotor constitutivo de *E. coli*, *e. g.*, o promotor do serovar 0121 wb (P0121, últimas três pistas na FIG. 9), sem perda de produção de LPS. Considerados em conjunto, estes resultados significam que os elementos de O11 e de *S. aureus* para a produção do antígeno O e do polímero capsular CP5 de O11, como aqui descrito, podem ser combinados em muitos sistemas de expressão de *E. coli* diferentes, resultando na produção de polissacárido de *S. aureus* recombinante.

Estes resultados mostraram, pela primeira vez, a produção em *E. coli* de uma estrutura capsular de polissacárido tendo origem num organismo Gram-positivo. Isto significa que foi possível, contrariamente à técnica anterior e expectativas convencionais, combinar as enzimas do aglomerado de O11 e as enzimas do aglomerado de cap de *S. aureus* para construir um polissacárido quimérico, *i. e.*, que a enzima trabalha conjuntamente na mesma estrutura *in vivo*.

Exemplo 3: Confirmação da estrutura molecular dos glicanos recombinantes

Para confirmar a atividade do aglomerado de CP5/O11 quimérico em *E. coli* a um nível molecular, foi desenvolvido um novo método permitindo a análise de açúcares ligados a UndPP, por utilização de marcação fluorescente do açúcar na extremidade de redução com 2-Aminobenzamida (2-AB). Para melhorar a resolução da análise, foram utilizados aglomerados quiméricos contendo eliminações que aumentaram a quantidade de RU não polimerizadas. Os glicolípidos de diferentes células de *E. coli* expressando o aglomerado quimérico contido no plasmídeo pLAFR1 e desprovido da flipase de cap5K (SEQ ID N°: 2) foram analisados como descrito abaixo.

Para extrair os glicanos ligados a UndPP, células de *E. coli* foram lavadas com NaCl a 0,9% e liofilizadas. As células secas foram extraídas, uma vez, com 30 mL de solvente orgânico (85 a 95% de Metanol = M). O sedimento celular liofilizado foi ainda extraído, duas vezes, com 5 mL de Clorofórmio: Metanol: Água (C:M:W = 10:10:3; v/v/v). O extrato (M) foi convertido com clorofórmio e água para uma razão final de 3:48:47 (C:M:W). O extrato 10:10:3 (C:M:W) foi convertido para um sistema de duas fases Bligh/Dyer (Bligh, E. G. e W. J. Dyer. 1959. Um método rápido de extração e purificação de lípidos totais. *Can J Biochem Physiol* 37(8): 911-7) por adição de água, resultando numa razão final de 10:10:9 (C:M:W). As fases foram separadas por centrifugação e a fase aquosa superior foi mantida para processamento adicional.

Para purificar os glicolípidos extraídos, a fase aquosa foi submetida a um cartucho tC₁₈ Sep-PAK. O cartucho foi condicionado

com 10 mL de metanol, seguido de equilibração com 10 mL de 3:48:47 (C:M:W). Após carregamento da amostra, o cartucho foi lavado com 10 mL de 3:48:47 (C:M:W) e eluído com 5 mL de metanol e 5 mL de 10:10:3 (C:M:W). As eluições combinadas foram secas sob N₂. As amostras de glicolípido foram hidrolisadas por dissolução das amostras secas em 2 mL de n-propanol:ácido trifluoroacético a 2 M (1:1), aquecimento a 50 °C, durante 15 min e, depois, evaporação até à secura sob N₂ (Glover, K. J., E. Weerapana e B. Imperiali. 2005. Agrupamento *in vitro* do heptassacárido ligado a UndPP para glicosilação procariótica ligada a N. *Proc Natl Acad Sci U S A* 102(40): 14255-9). As amostras secas foram marcadas com 2-AB e a limpeza do glicano foi realizada utilizando o método do disco de papel, como descrito (Bigge, J. C., T. P. Patel, J. A. Bruce, P. N. Goulding, S. M. Charles, R. B. Parekh. 1995. Marcação fluorescente não seletiva e eficiente de glicanos utilizando 2-aminobenzamida e ácido antranílico. *Anal Biochem* 230(2): 229-38; Merry, A. H., D. C. Neville, L. Royle, B. Matthews, D. J. Harvey, R. A. Dwek e P. M. Rudd. 2002. Recuperação de O-glicanos intactos marcados com 2-aminobenzamida libertados de glicoproteínas por hidrazinólise. *Anal Biochem* 304(1): 91-9). Os glicanos marcados com 2-AB foram separados por HPLC utilizando uma coluna de fase normal GlycoSep-N, de acordo com Royle *et al.*, mas modificada para um sistema de três solventes (Royle, L., T. S. Mattu, E. Hart, J. I. Langridge, A. H. Merry, N. Murphy, D. J. Harvey, R. A. Dwek, P. M. Rudd. 2002. Uma base de dados analítica e estrutural proporciona uma estratégia para sequenciar O-glicanos a partir de quantidades de micrograma de glicoproteínas. *Anal Biochem* 304(1): 70-90). O solvente A foi formato de amónio a 10 mM, pH 4,4, em acetonitrilo a 80%. O solvente B foi formato de amónio a 30 mM, pH 4,4, em acetonitrilo a 40%. O solvente C foi ácido fórmico a 0,5%. A

temperatura de coluna foi 30 °C e os glicanos marcados com 2-AB foram detetados por fluorescência (λ_{exc} de excitação = 330 nm, λ_{em} de emissão = 420 nm). As condições de gradiente foram um gradiente linear de 100% de A a 100% de B, ao longo de 160 min, a um caudal de 0,4 mL/min, seguido de 2 min de 100% de B a 100% de C, aumentando o caudal para 1 mL/min. A coluna foi lavada durante 5 min com 100% de C, regressando a 100% de A ao longo de 2 min e correndo durante 15 min a 100% de A, a um caudal de 1 mL/min, depois, restituindo o caudal para 0,4 mL/min, durante 5 min. As amostras foram injetadas em água.

As frações secas foram ressuspensas em 5 μ L de acetonitrilo (ACN) a 10%, ácido trifluoroacético (TFA) a 0,1% e misturadas 1:1 com solução de matriz (DHB a 40 mg/mL em ACN a 50%, TFA a 0,1%) na placa alvo. Os dados de MS e MS/MS foram adquiridos manualmente no modo iónico positivo num espectrómetro de massa MALDI-ToF/ToF Ultraflex-II (Bruker Daltonik GmbH, Bremen, Alemanha). MS/MS foram obtidos utilizando o método de LIFT. Uma mistura peptídica padrão (Bruker Daltonik GmbH) foi utilizada para calibração externa. Os espectros foram exportados utilizando o software Flex Analysis (Bruker Daltonik GmbH) e analisados manualmente.

Os extratos de metanol de *E. coli* W3110 Δ wecA (CP5) contendo plasmídeos com (linha grossa) ou sem (linha fina a tracejado) os aglomerados quiméricos foram purificados através de cartuchos tC18 e analisados por HPLC de fase normal. As frações correspondentes aos picos mostrados na FIG. 10A verificados a 37, 40 e 45 minutos de eluição foram analisados por MALDI-MS/MS. As amostras eluindo a 37 e 40 minutos foram identificadas como RU de CP5 recombinantes, com e sem o grupo O-acetilo conjugado, respetivamente. A amostra eluindo a

45 minutos foi identificada como estrutura de RU de *S. aureus* não acetilada alongada por uma desoxi-N-acetil-hexosamina (como mostrado na FIG. 11E). No aglomerado quimérico de CP5, *cap5HIJ* substituiu os genes *wbjA* e *wzy* do aglomerado de O11 em pLAFR. Além dos genes *cap5HIJ*, a substituição continha a cassete de *cat* (SEQ ID N°: 2).

Os extratos de metanol de *E. coli* W3110 Δ wecAwzzE contendo plasmídeos com (linha grossa) ou sem (linha fina a tracejado) o aglomerado quimérico foram purificados através de cartuchos tC18 e analisados por HPLC de fase normal. A FIG. 10B mostra os resultados da análise de HPLC de RU recombinante de CP8 produzida utilizando um aglomerado quimérico (SEQ ID N°: 4 sem polimerase). Os picos específicos para células que expressam o açúcar recombinante foram identificados a 23', 32', 38' e 45 de eluição, recolhidos e analisados por MALDI-MS e MALDI-MS/MS. No aglomerado quimérico de CP8, *cap8HJK* substituiu os genes *wbjA* e *wzy* do aglomerado de O11, *i. e.*, uma construção sem a polimerase para acumular RU única para análise. Além dos genes *cap*, a substituição continha a cassete de *cat*.

A FIG. 11A mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP5 quimérico da presente invenção em *E. coli* eluindo a 37 minutos. A massa principal $m/z=772$ ($[M+H]^+$) foi selecionada e analisada por MS/MS, a qual mostra um padrão de fragmentação consistente com a estrutura de RU de CP5 acetilada que era esperada à luz da invenção divulgada nesta descrição. As espécies O-acetiladas são caracterizadas por uma perda específica de 42 mais a massa do monossacárido FucNAc (dHexNAc(OAc)) na posição do meio da RU. Os íons de fragmentação estão indicados de acordo com nomenclatura do consórcio para a

glicômica funcional, CFG (www.functionalglycomics.org/static/consortium/Nomenclature.shtml). 2-AB, 2-aminobenzamida. A legenda para os iões de fragmentação é dada na inserção da FIG. 11A.

A FIG. 11A mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP5 quimérico da presente invenção em *E. coli* eluindo a 40 minutos. A massa principal de $m/z = 730$ ($[M+H]^+$) foi selecionada e analisada por MS/MS, a qual mostra a série iônica de fragmentação consistente com a estrutura de RU de CP5 não acetilada que era esperada à luz da invenção divulgada nesta descrição. 2-AB, 2-aminobenzamida. A legenda para os iões de fragmentação é dada na inserção da FIG. 11B.

A FIG. 11C mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP8 quimérico da presente invenção em *E. coli* eluindo a 32 minutos. Uma massa principal de $m/z = 794$ ($[M+H]^+$) foi selecionada e analisada por MS/MS, a qual mostra a série iônica de fragmentação consistente com a estrutura de RU de CP8 acetilada que era esperada à luz da invenção divulgada por esta descrição. As espécies O-acetiladas são caracterizadas por uma perda específica de 42 mais a massa do monossacárido ManNAcA (HexNAcA(OAc)) na posição mais exterior da RU. Os iões de fragmentação estão indicados de acordo com a nomenclatura do CFG. 2-AB, 2-aminobenzamida. A legenda para os iões de fragmentação é dada na inserção da FIG. 11C.

A FIG. 11D mostra os resultados da análise de MALDI-MS/MS do pico específico produzido pela expressão de uma forma de realização do aglomerado de CP8 quimérico da presente invenção

em *E. coli* eluindo a 38 minutos. A massa de $m/z = 730$ ($[M+H]^+$) foi selecionada e analisada por MS/MS, a qual mostra a série iônica de fragmentação consistente com a estrutura de RU de CP8 não acetilada que também era esperada à luz da invenção divulgada nesta descrição. A análise adicional mostrou que os picos eluindo mais tarde (mostrados na FIG. 10A a 40 min e FIG. 10B a 38 min) contêm os trissacáridos não O-acetilados das RU de CP5 e 8. Os íons de fragmentação estão indicados de acordo com a nomenclatura do CFG. 2-AB, 2-aminobenzamida. A legenda para os íons de fragmentação é dada na inserção da FIG. 11D.

Os resultados de MS mostraram que as massas e série iônica de fragmentação estão de acordo com a estrutura molecular do oligossacárido de RU de CP5 com a O-acetilação do resíduo de FucNAc do meio (*i. e.*, o pico a 37 minutos na FIG. 10A e na FIG. 11A) ou sem a O-acetilação do resíduo de FucNAc do meio (*i. e.*, o pico 40 minutos na FIG. 10A e em 11B). O sinal a 45 minutos na FIG. 10A foi identificado como um tetrassacárido, o qual é ainda analisado abaixo. A mesma análise foi repetida com o aglomerado de CP8 quimérico que não possuía o gene de polimerase. Nesses extratos, sinais consistentes com a estrutura de RU O-acetilada esperada à luz da invenção divulgada nesta descrição foram verificados a 23 e 32 minutos de eluição, como mostrado nas FIG. 10B e 11C. A presença de dois tempos de eluição diferentes para a mesma sequência de glicano que a identificada por MALDI-MS/MS indica um evento de migração de O-acetilo a ocorrer durante a preparação da amostra. As RU não acetiladas foram identificadas para os extratos de CP5 e CP8 a 40' e 38', como mostrado nas FIG. 11B e D, respectivamente. As estruturas de RU de CP5 e CP8 estavam presentes em diferentes estirpes de *E. coli*, incluindo, por exemplo, W3110, W3310 $\Delta wecA$, W3110 $\Delta wecAwzzE$ e W3110 $\Delta wecAwzzE \Delta waaL$.

Exemplo 4: Melhoramento da estrutura da unidade de repetição e sua análise.

O pico de HPLC mostrado na FIG. 10B eluindo a 45 minutos, derivado de células de *E. coli* expressando o aglomerado de CP8 quimérico (SEQ ID N°: 4), mas desprovido do gene de polimerase *wzy cap8I*, também foi analisado por MALDI-MS/MS. O íon mais intenso no MS de varrimento completo foi $m/z=939$ ($[M+H]^+$) e a análise de sequência foi realizada por MS/MS. Os resultados desta análise de MS/MS estão mostrados na FIG. 11E e apresentam uma série iônica de fragmentação consistente com a RU capsular de *S. aureus* não acetilada, alongada por uma massa de uma desoxi-N-acetil-hexosamina na extremidade não redutora, como esperado à luz da invenção divulgada nesta descrição. Os íons de fragmentação correspondentes às estruturas hipotéticas estão indicados de acordo com a nomenclatura do CFG acima dos picos. 2-AB, 2-aminobenzamida. A legenda para os íons de fragmentação é dada na inserção da FIG. 11E.

O resultado mostrado na FIG. 11E sugeriu que uma glicosiltransferase de *E. coli* era capaz de modificar o resíduo de ManNAcA da RU de CP8. Essa RU alterada, muito provavelmente, não estaria polimerizada por *cap8I*. A análise das especificidades de glicosiltransferase no hospedeiro W3110 de *E. coli* indicou que uma enzima do aglomerado de ECA pode interferir com o açúcar recombinante, especificamente o produto gênico *wecF*, uma putativa 4-N-acetilfucosaminatransferase. *WecF* adiciona naturalmente uma 4-N-acetilfucosamina a ManNAcA compreendida em ECA, muito provavelmente a enzima também poderia alongar a RU de CP8 e CP5.

Para resolver este problema, foi desenvolvida outra abordagem nova. Especificamente, os genes do aglomerado de ECA localizados a jusante do gene *wecC*, incluindo *wecF*, foram eliminados. Isto foi alcançado utilizando o método descrito por Datsenko *et al.* (Datsenko, K. A. e B. L. Wanner. 2000. "Inativação de um passo de genes cromossômicos em *Escherichia coli* K-12 utilizando produtos de PCR". *Proc Natl Acad Sci USA* 97(12):6640-6645). Diferentes hospedeiros de expressão de *E. coli* foram eliminados nas regiões gênicas de *waaL* e *rmlB-wecG* e também em algumas estirpes em *wecA-wzzECA*. Os extratos Purificados de Sep-PAK (extratos de Metanol e 10:10:3) destas células mutadas expressando o aglomerado quimérico de CP8 mutante de polimerase foram analisados por HPLC de fase normal, como descrito acima.

A FIG. 11F apresenta os resultados das análises de HPLC de extratos de metanol de células de *E. coli* W3110 $\Delta waaL$ que expressam o mutante de polimerase da SEQ ID N°: 4 (linha fina a tracejado) em comparação com células com uma eliminação adicional dos genes do aglomerado de ECA *rmlB-wecG* (W3110 $\Delta waaL \Delta rmlB-wecG::cat$) (linha grossa). Os extratos foram purificados através de cartuchos tC18 e analisados por HPLC de fase normal. Como mostrado na FIG. 11F, o pico principal surgindo a 45 minutos na FIG. 10B estava ausente, resultando em picos específicos para as RU de CP8 acetiladas e não acetiladas (FIG. 11F), indicando que uma das glicosiltransferases de ECA - muito provavelmente *wecF* - é responsável pelo fenótipo de alongamento aberrante. Resultados semelhantes foram obtidos quando o aglomerado quimérico de CP5 foi testado em diferentes estirpes. Isto implica que a eliminação de glicosiltransferases e enzimas geradas em *E. coli* requeridas para a biossíntese de açúcar ativada por nucleótido é uma estratégia possível para

otimização da qualidade e quantidade de polissacáridos produzidos recombinantemente em *E. coli*. As enzimas alvo muito provavelmente seriam codificadas no aglomerado de antígeno O, o aglomerado de ECA, e nos aglomerados de ácido colânico ou de cápsula.

Evidência adicional para a qualidade do polissacárido recombinante ligado a UndPP foi obtida de uma análise de HPLC de fase normal otimizada de extratos de glicolípidos marcados fluorescentemente, purificados por Sep-PAK, de hospedeiros de expressão cromossomicamente otimizados, como descrito acima. Para desempenho ótimo das colunas Sep-PAK para purificação de lípidos ligados a oligossacáridos e polissacáridos de CP5 e CP8 carregados, fosfato de terc-butilamônio (TBAP) foi adicionado aos extratos antes de carregamento nos cartuchos Sep-PAK. Como referido por Trent, et al., o catião deste sal melhora a ligação de coluna de compostos carregados blindando cargas negativas com cadeias de butilo hidrófobas (Trent, M. S., A. A. Ribeiro, et al. (2001). "Acumulação de um aminoaçúcar ligado a poliisopreno em *Salmonella typhimurium* e *Escherichia coli* resistentes a polimixina: caracterização estrutural e transferência para lípido A no periplasma". *J Biol Chem* 276(46): 43132-43144.). Este método otimizado foi aplicado às amostras de CP5 e CP8 obtidas por extração com metanol de células expressando aglomerados quiméricos de CP5 ou CP8 contendo uma polimerase.

A FIG. 11G proporciona os resultados da análise de HPLC que mostram o repertório de glicano de CP5 completo presente em UndPP em células de *E. coli*. Os extratos de metanol de *E. coli* W3110 $\Delta waaL \Delta wecAwzzECA \Delta rmlB-wecG::cat$ que expressam o aglomerado de CP5 quimérico SEQ 3 (linha a cheio) ou um controle

plasmídico vazio (linha a tracejado) foram extraídos em fase sólida em cartuchos Sep-PAK e tratados com ácido moderado para hidrolisar açúcares de UndPP. O material resultante foi colocado em reação com 2AB por aminação redutora, para marcar extremidades de redução dos glicanos e analisado por HPLC de fase normal. Os sinais presentes na linha a cheio, mas não na linha a tracejado, representam material específico de CP5. As letras maiúsculas indicam picos contendo polímeros da RU de CP5 acetilada e/ou não acetilada, como identificado por MALDI-MS/MS das frações recolhidas. A legenda da FIG. 11G indica as estruturas moleculares propostas, como deduzidas a partir da análise de MS/MS. Deve notar-se que os polímeros de RU acetilados e não acetilados mostrados para MS/MS confirmaram estruturas do mesmo grau de polimerização agrupadas no cromatograma, como indicado por barras grossas. As letras maiúsculas mostram os seguintes comprimentos: A e B: uma RU; C, D e E: duas RU; F e G: três RU; e H: quatro RU. O pico alargado entre 95' e 125' na FIG. 11G, muito provavelmente, representa 5 ou mais RU polimerizadas não resolvidas pela coluna.

A FIG. 11H apresenta resultados adicionais de HPLC, mostrando glicanos de CP5 acetilados e homogeneidade de RU. Para preparar esta análise de HPLC, amostras de glicano marcadas com 2AB de *E. coli* W3110 $\Delta waaL$ $\Delta wecA wzzECA$ $\Delta rmlB-wecG::cat$ expressando o aglomerado de CP5 quimérico SEQ ID N°.: 3 (preparado de acordo com os processos descritos acima fazendo referência à FIG. 11G) foram tratadas com NaOH em solução aquosa e remarcadas. Como mostrando na FIG. 11H, as amostras antes (a tracejado) e após (linha a cheio) tratamento alcalino foram analisadas por HPLC. Os números na FIG. 11H indicam os putativos números das RU nos picos correspondentes. Deve ser observado que, na FIG. 11H, os picos acetilados mostrados na FIG. 11G

unificam no sinal de polímero não acetilado e que a desacetilação resolveu as unidades de RU nos tempos de eluição após 95 minutos.

A FIG. 11I proporciona os resultados da análise de HPLC mostrando o repertório de glicano de CP8 presente em UndPP em células de *E. coli*. Os extratos de metanol de *E. coli* W3110 $\Delta waaL \Delta wecAwzzECA \Delta rmlB-wecG::cat$ expressando o aglomerado de CP8 quimérico SEQ ID N°. 4 (linha a cheio) ou um controle plasmídico vazio (linha a tracejado) foram extraídos em fase sólida em cartuchos Sep-PAK e tratados com ácido moderado para hidrolisar açúcares de UndPP. O material resultante foi colocado em reação com 2AB por aminação redutora, para marcar extremidades de redução dos glicanos e analisado por HPLC de fase normal. Os sinais presentes na linha a cheio, mas não na linha a tracejado, representam material específico de CP8. Estão indicadas as putativas estruturas de RU de CP8 acetiladas e/ou não acetiladas, como identificadas por MALDI-MS/MS das frações recolhidas. É de notar que como nos resultados de HPLC com CP5 mostrados na FIG. 11G, os polímeros de RU de CP8 acetilados e não acetilados do mesmo grau de polimerização agrupam-se no cromatograma da FIG. 11H, como indicado por barras grossas. O material detetado após 110 minutos representa polímeros de CP8 mais longos.

A FIG. 11J apresenta resultados adicionais de HPLC, mostrando a desacetilação de glicanos de CP8 e homogeneidade de RU. As amostras de glicano marcadas com 2AB de *E. coli* W3110 $\Delta waaL \Delta wecAwzzECA \Delta rmlB-wecG::cat$ expressando o aglomerado de CP8 quimérico SEQ ID N°. 4 foram tratadas com NaOH em solução aquosa e remarcadas. As amostras antes (a tracejado) e após (linha a cheio) tratamento alcalino foram analisadas por HPLC.

Os números indicam os putativos números das RU nos picos correspondentes. Deve notar-se que os picos acetilados desaparecem em grande parte e que os sinais do polímero não acetilado aumentam e que a desacetilação resolveu as unidades de RU nos tempos de eluição após 110 minutos.

As FIG. 11H e 11J mostram resultados de HPLC indicativos do característico padrão de bandas semelhante a escadas de antígenos O, quando o tratamento alcalino foi realizado nestas amostras de CP5 e CP8, para remover as modificações de acetilação dos oligossacáridos e polissacáridos. Os resultados mostram discretos picos agudos, com incrementos de tempo de eluição diminuindo constantemente. Isto implica que essas cadeias de hidratos de carbono analisadas são polímeros lineares compostos por RU idênticas. Estes dados mostram que os açúcares de CP5 e CP8 recombinantes produzidos em *E. coli* são polimerizados regularmente e parcialmente acetilados. Os polímeros CP5 e CP8 não acetilados eluem identicamente da coluna de HPLC, como esperado da sua semelhança em estrutura; contudo, a cromatografia de fase normal também revela diferenças: por exemplo, o CP5 polimeriza em menor escala do que o CP8 e a acetilação é mais frequente em CP5; nos comprimentos de RU acima de 4, o CP5 tem uma clara preferência para produzir polímeros de 7 RU, enquanto que o CP8 polimeriza num grau mais amplo; e, como indicado pelos resultados de HPLC e MS/MS, o CP5 é mais eficiente para a produção de glicano do que o CP8.

Nas vias de polimerização dependente de *wzy*, foi referido por Marolda, et al., que uma enzima específica (*wzz* ou *cld*, para determinante de comprimento de cadeia) é responsável por determinar o número médio de passos de polimerização de RU realizados (Marolda, C. L., L. D. Tatar, et al. (2006).

"Interação da translocase de Wzx e da correspondente polimerase e proteínas reguladoras do comprimento de cadeia na translocação e agrupamento periplasmático do lipopolissacárido antigénio O". *J Bacteriol* 188(14): 5124-5135.). As enzimas Wzz causam uma média do número específico de repetições, e. g., polímeros de açúcar curtos, longos e muito longos, e são conhecidas por transferirem a sua especificidade de comprimento para vias polissacáridas exógenas. Os comprimentos e quantidades dos glicolípidos de CP8 foram analisados na estirpe de produção, resultando em comprimentos maiores e quantidade menor deste açúcar. Para aumentar a quantidade de moléculas e, desse modo, a eficiência de transferência de açúcar para glicosilação proteica, foi realizada uma regulação negativa do comprimento do açúcar de CP8 utilizando uma enzima Wzz específica.

Para testar o efeito de uma proteína Wzz no tamanho e quantidades de açúcares de CP8 em lípido, foi realizada a co-expressão de Wzz de *E. coli* wzz 07 a partir de um plasmídeo separado (SEQ ID N°: 19). A FIG. 11K apresenta os resultados deste teste. Os extratos de metanol de *E. coli* W3110 $\Delta waaL$ $\Delta wecA wzzECA$ $\Delta rmlB-wecG::cat$ expressando o aglomerado de CP8 quimérico SEQ ID N°: 4 e uma cópia indutível de IPTG produzida em plasmídeo de wzz07 (SEQ ID N°: 21, linha a cheio) ou um controlo plasmídico vazio (linha a tracejado) foram extraídos em fase sólida em cartuchos Sep-PAK e tratados com ácido moderado para hidrolisar açúcares de UndPP. Os glicanos marcados com 2AB foram analisados por HPLC de fase normal. O tratamento alcalino da amostra de CP8 mostrou que mais de 85% da área entre 95 e 115' representa 7 ou 8 polímeros de RU de CP8, indicando uma ampla variedade de acetilação. Estes resultados também indicam que o aglomerado de CP8 quimérico induziu: a) uma intensificação em números de repetição do glicano mais abundante de 7 para 8 e

b) uma intensidade global mais elevada de sinal fluorescente, como avaliada a partir da área sob o cromatograma.

O tratamento alcalino confirmou a acetilação do glicano encurtado, como na FIG. 11I e 11J, indicando que o comprimento de um polissacárido recombinante pode ser regulado por uma enzima Wzz estranha. Também é possível regular o comprimento do polímero de açúcar capsular por uma enzima Wzz derivada de antigénio O. Além disso, diferentes promotores em frente do aglomerado quimérico, quando presentes num plasmídeo, causam diferentes níveis de expressão e diferentes graus de polimerização.

Exemplo 5: Glicosilação proteica com os glicanos de CP5 e CP8 e caracterização de produto

Foram testados diferentes variantes do aglomerado quimérico para a produção de bioconjugado. Os aglomerados génicos de O11/CP5 quiméricos (SEQ ID N°: 2 e 3), os quais contêm diferentes variantes de regiões de especificidade de *S. aureus* no aglomerado de antigénio O de O11, em vez de *wbjA* e *wzy*, foram expressos na estirpe hospedeira *E. coli* W3110 $\Delta waaL \Delta wecAwzzE::cat$, na presença de PglB (SEQ ID N°: 27?) e EPA (SEQ ID N°: 13). As células hospedeiras de W3110 $\Delta waaL \Delta wecAwzzE::cat$ expressaram EPA com dois sítios de glicosilação (da SEQ ID N°: 13) e PglB (SEQ ID N°: 27) a partir de plasmídeos separados, além do plasmídeo pLAFR1 com o aglomerado de antigénio O de O11, onde os genes *wbjA* e *wzy* foram substituídos com diferentes conjuntos do gene *cap5* (e a cassette de *cat*, SEQ ID N°: 2 e SEQ ID N°: 3).

A proteína EPA é expressa, contendo: a) uma sequência de péptido sinal N-terminal para exportar para o periplasma, b) duas sequências consenso de N-glicosilação bacterianas manipuladas na sequência da proteína (SEQ ID N°: 13), como apresentado no Exemplo 10 do documento WO 2009/104074, aqui incorporado por referência na sua totalidade e c) uma cauda de hexa his para purificação. As células foram cultivadas em frascos Erlenmeyer de 5 L em meio de LB. Uma cultura de um dia para o outro foi diluída para D.O._{600 nm} = 0,05. A D.O._{600 nm} cerca de 0,5, a expressão de PglB foi induzida por adição de IPTG a 1 mM e a expressão de EPA foi induzida por adição de arabinose (concentração final de 0,2%). As células foram cultivadas durante 4 horas, a indução foi repetida e as células foram cultivadas durante cerca de 16 horas adicionais. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. O bioconjugado de EPA-CP5 sem e com o gene *cap5K* de flipase de *S. aureus* (SEQ ID N°: 2 e 3) foi eluído por imidazole a 0,5 M e os picos eluído foram agregados e analisados por SDS PAGE e corados por Coomassie e prata (FIG. 12).

A FIG. 12 apresenta os resultados de SDS PAGE. O painel esquerdo mostra a coloração de Coomassie e o painel da direita mostra a coloração de prata. Os números no meio indicam os tamanhos do marcador de peso molecular. As letras por baixo das pistas indicam os genes que estavam presentes no aglomerado quimérico expresso nas estirpes utilizadas para produção de bioconjugado. A estirpe hospedeira foi *E. coli* W3110 $\Delta waaL \Delta wecAwzzE::cat$. Os resultados mostram sinal de proteína a 70 kDa

(mobilidade eletroforética) correspondendo, muito provavelmente, a EPA não glicosilada e uma escada de bandas acima (100-170 kDa). A escada corresponde, provavelmente, à proteína EPA glicosilada com o glicano CP5 de *S. aureus* recombinante. Além disso, os resultados indicam que a inclusão do gene de flipase no sistema aumenta o rendimento da glicoproteína (pistas do meio e da direita).

Numa análise separada, o bioconjugado de CP5-EPA foi produzido em *E. coli* W3110 $\Delta waaL \Delta wecA wzzE::cat$ por co-expressão do aglomerado génico de CP5 quimérico (SEQ ID N°: 3), PglB (SEQ ID N°: 27) a partir do plasmídeo pEXT21 e EPA (contendo dois sítios de glicosilação, SEQ ID N°: 13) de plasmídeos separados. Para se obter um processo mais controlado para produção de bioconjugado, as células foram cultivadas num biorreator de 2 L até uma D.O._{600 nm} = 30 a 37 °C e a expressão de PglB e EPA foi induzida pela adição de IPTG a 1 mM e arabinose a 0,2%. As células foram cultivadas durante 18 h, a 37 °C, sob condições limitantes de oxigénio. As células foram sedimentadas por centrifugação, lavadas e ressuspensas em tampão de sacarose a 25%, a uma D.O._{600 nm} = 200, após incubação de 30 min., a 4 °C, a suspensão foi sedimentada e lisada por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas presentes no sobrenadante foram carregadas numa cromatografia de afinidade de Ni²⁺. As EPA glicosiladas e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M e carregadas numa coluna de troca aniónica SourceQ. A EPA glicosilada foi separada da EPA não glicosilada por aplicação de um gradiente de concentração crescente de NaCl.

Como mostrado na FIG. 13A, a EPA glicosilada purificada (CP5-EPA) foi separada por SDS PAGE e corada por Coomassie (pista da esquerda) ou transferida para membranas de nitrocelulose e incubada com anticorpos anti-CP5 (pista do meio) ou anticorpos anti-EPA (pista da direita). O bioconjugado purificado foi reconhecido pelos anticorpos específicos de EPA (pista da direita), assim como o antissoro policlonal específico de CP5 (pista do meio). A indica a posição no gel de onde um pedaço foi cortado e utilizado para tripsinização e análise de glicopéptidos por MALDI-MS/MS. A FIG. 13B apresenta as massas MALDI-MS/MS de M/Z verificadas para o sítio de glicosilação no péptido tripsinizado DNNNSTPTVISHR ligado N-glicosidicamente à massa de RU O-acetilada ($m/z=2088$ ($[M+H]^+$)). A análise de MS/MS da $m/z=2088$ mostra fragmentação parcial da unidade de açúcar, como indicado. A inserção ilustra a estrutura de RU conjugada ao péptido derivado de tripsinização de CP5-EPA purificado da FIG. 13A. As perdas sequenciais de ManNAcA (HexNAcA, 217 Da) e FucNAc acetilado (dHexNAc(OAc), 229 Da) apoiam a estrutura esperada do glicano. A FIG. 13C apresenta as massas MALDI-MS/MS de M/Z verificadas para o sítio de glicosilação no péptido tripsinizado DQNR ligado N-glicosidicamente à massa de RU O-acetilada ($m/z=1165$ ($[M+H]^+$)). A análise de MS/MS de $m/z=1165$ mostra a série iônica de fragmentação completa do ião Y, consistente com a estrutura de RU de CP5. A inserção ilustra a estrutura de RU conjugada ao péptido derivado de tripsinização de CP5-EPA purificado da FIG. 13A. As perdas sequenciais de ManNAcA (HexNAcA, 217 Da), FucNAc acetilado (dHexNAc(OAc), 229 Da) e FucNAc (dHexNAc, 187 Da) estão mostradas, confirmando a estrutura esperada do glicano no péptido DQNR ($m/z=532$ Da ($[M+H]^+$)).

Na FIG. 13D, o bioconjugado de CP8 em *E. coli* foi produzido utilizando a mesma estratégia que a produção do bioconjugado de CP5. O bioconjugado de CP8-EPA foi produzido em *E. coli* por co-expressão do aglomerado génico de CP8 quimérico (SEQ ID N°: 4), PglB (no plasmídeo pEXT21 (SEQ ID N°: 27)) e EPA contendo dois sítios de glicosilação (SEQ ID N°: 13). As células foram cultivadas num biorreator, com um volume de partida de 7 L em meio semidefinido contendo, como fontes de C, glicerol, peptona e extrato de levedura. As células foram cultivadas a 37 °C em modo descontínuo ou descontínuo pulsado até uma D.O._{600 nm} de 30 e a expressão de PglB e EPA foi induzida pela adição de IPTG a 1 mM e arabinose a 10%. Após indução, as células foram ainda cultivadas em modo semicontínuo, durante um período de 15 horas, sob condições limitantes de oxigénio. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. As EPA glicosiladas e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M e carregadas numa coluna de troca aniónica SourceQ. A EPA glicosilada foi separada da EPA não glicosilada por aplicação de um gradiente de concentração crescente de NaCl.

Como representado na FIG. 13D, a proteína purificada foi separada por SDS PAGE e corada por Coomassie (pista da esquerda) ou transferida para membranas de nitrocelulose e incubada com anticorpos anti-CP8 (pista da direita) ou anticorpos anti-EPA (pista do meio).

Foram testadas diferentes estratégias para melhoramento adicional do sistema de glicosilação. Numa estratégia para reduzir o número de plasmídeo no sistema de produção para diminuir a carga de um antibiótico adicional, assim como manter um plasmídeo extra, a cassette de expressão para *pglB* foi clonada no plasmídeo contendo os aglomerados quiméricos para CP5 (SEQ ID N°: 17) e CP8 (SEQ ID N°: 18). A cassette de expressão é composta pela região intergênica presente entre *galF* e *wbqA* do genoma de *E. coli* 0121 (para uma sequência promotora) e a sequência de *pglB* a jusante desta. Esta cassette de expressão foi clonada imediatamente a jusante dos aglomerados quiméricos de CP5 e CP8. A requerente testou *E. coli* W3110 $\Delta waaL$ $\Delta wecA wzzECA::cat$ contendo o aglomerado de CP5 quimérico (SEQ ID N°: 3) e *pglB* (SEQ ID N°: 27) em plasmídeos separados ou no mesmo plasmídeo (SEQ ID N°: 17). Além disso, a EPA (SEQ ID N°: 13) foi expressa a partir de um plasmídeo sob o controlo de um promotor indutível de arabinose. As células foram cultivadas em frascos Erlenmeyer de 5 L em meio de LB, a 37 °C. Uma cultura foi diluída, de um dia para o outro, para D.O._{600 nm} = 0,05. A D.O._{600 nm} cerca de 0,5 de expressão de *PglB* foi induzida por adição de IPTG a 1 mM e a expressão de EPA foi induzida por adição de arabinose (concentração final de 0,2%). As células foram cultivadas durante 4 horas, a indução foi repetida e as células foram cultivadas durante cerca de 16 horas adicionais. A cultura foi sedimentada por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. O EPA-CP5 foi eluído por imidazole a 0,5 M e os picos eluídos foram agregados e analisados por SDS PAGE e por Coomassie. A FIG. 13E representa os resultados de SDS PAGE. São

mostradas as células contendo 3 (esquerda) ou 2 plasmídeos (pista da direita) para produção de glicoconjugado. Os resultados mostram que a produção de glicolípido e conjugado para CP5-EPA foi mantida.

Uma otimização adicional do sistema foi a integração da sequência proteica *wzz* (reguladora do comprimento do polímero) nos plasmídeos utilizados para a glicosilação proteica. Exemplificado pelo sistema produzindo CP8-EPA, *wzz* foi integrado no aglomerado de CP8 quimérico plasmídeo gerado em plasmídeo (SEQ ID N°: 19) e a *jusante* do gene *epa* no plasmídeo de expressão para a proteína transportadora (SEQ ID N°: 20). O bioconjugado de CP8-EPA foi produzido em *E. coli* W3110 $\Delta waaL$ $\Delta wecAwzzECA$ $\Delta rmlB-wecG::cat$ compreendendo 2 plasmídeos: um plasmídeo continha além do aglomerado génico de CP8 quimérico uma cópia do gene *wzz* 07 e uma cassete de ADN para a expressão constitutiva do gene *pglB* (SEQ ID N°: 19); o segundo plasmídeo continha, em primeiro lugar, o gene para expressão e secreção da proteína EPA destoxificada contendo dois sítios de glicosilação e, em segundo lugar, uma cópia de *wzz*07 sob o controlo do mesmo promotor (SEQ ID N°: 20). A estirpe resultante, *E. coli* W3110 $\Delta waaL$ $\Delta wecAwzzECA$ $\Delta rmlB-wecG::cat$, contendo os plasmídeos mencionados foi cultivada num biorreator, com um volume de partida de 7 L em meio semidefinido contendo, como fontes de C, glicerol, peptona e extrato de levedura. As células foram cultivadas em modo descontínuo ou descontínuo pulsado até uma D.O._{600 nm} de 30 e a expressão de PglB e EPA foi induzida. Após indução, as células foram ainda cultivadas em modo semicontínuo, durante um período de 15 horas, sob condições limitantes de oxigénio e recolhidas por centrifugação. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e

lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni^{2+} . As EPA glicosiladas e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M. A formação do glicoconjugado CP8-EPA é mostrada na FIG. 13F por Coomassie e transferência de western utilizando antissoros anti-his e anti-CP8. A FIG. 13F mostra os resultados da separação de SDS PAGE da proteína purificada e análise por coloração de Coomassie (pista da esquerda) ou transferida para membranas de nitrocelulose e sondada com anticorpos anti-cauda de his (pista do meio) ou anticorpos anti-CP8 (pista da direita).

A caracterização do glicoconjugado de CP5-EPA foi ainda refinada por diversos métodos analíticos. O CovalX (Schlieren, Suíça) realizou a análise de MALDI de Elevada Massa de uma amostra de CP5-EPA purificada, produzida utilizando o sistema de 3 plasmídeos como utilizado nas análises representadas na FIG. 13A em *W3110 ΔwaaL ΔwecAwzzECA::cat*. A FIG. 14A representa os resultados de MALDI de Elevada Massa. A^+ e B^+ apontam na direção das espécies de proteína de massa ($[\text{M}+\text{H}]^+$) correspondendo a EPA não glicosilada e EPA glicosilada, respetivamente. As formas oligoméricas podem estar presentes com peso molecular mais elevado e os sinais na área de baixo PM são contaminantes ou produtos de degradação. Os resultados apresentados na FIG. 14A mostram que a preparação de proteína acima continha uma população proteica em grande parte monomérica que é 4 kDa maior que a proteína EPA isoladamente, indicativo de um comprimento médio de açúcar de 5,2 unidades de repetição. Isto está em concordância com o comprimento de açúcar de 5-7 da principal forma de glicoconjugado na preparação, como analisado por SDS-PAGE, coloração de azul brilhante de Coomassie e contagem

das unidades de repetição na principal forma de conjugado (ver as FIG. 7, 8 e 13A).

O CP5-EPA foi ainda caracterizado por cromatografia de exclusão por tamanho (SEC-HPLC). A requerente utilizou o sistema de 3 plasmídeos em W3110 *ΔwaaL ΔwecAwzzECA::cat*, como utilizado nas análises representadas na FIG. 13A. A amostra foi purificada por cromatografia de troca aniônica para remover EPA não glicosilada. A análise foi realizada numa coluna G2000SWXL Supelco TSK. A FIG. 14B mostra os resultados da análise de SEC-HPLC da amostra de CP5-EPA purificada. É mostrado o traçado de UV medido a 280 nm. A linha grossa a cheio deriva da análise de 3,25 µg de CP5-EPA purificado, a linha fina foi obtida de 5 µg de EPA purificada, não glicosilada. Um pico homogêneo principal a 11,5 minutos de eluição está mostrado, enquanto que a EPA não glicosilada eluiu a 12,9 minutos (FIG. 14B). O cálculo dos raios hidrodinâmicos das duas moléculas resultou num tamanho de 42 kDa para EPA não glicosilada e 166 kDa para EPA glicosilada. Isto indica que a EPA glicosilada surge como uma proteína monomérica alongada em solução, como esperado devido à estrutura linear do glicano.

As presentes análises confirmaram, deste modo, que o bioconjugado de CP5-EPA consiste da proteína EPA e da correta estrutura do glicano O-acetilado. Com base nestes resultados, também poderia ser predado que o bioconjugado de CP8-EPA consistia da proteína EPA e da estrutura correta do glicano O-acetilado.

Exemplo 6: Glicosilação da proteína de *S. aureus* e caracterização do produto

Para provar a versatilidade da glicosilação "in vivo", para produzir candidatos vacinais de glicoconjugado, foram utilizadas várias proteínas veículo como substrato para serem glicosilados com CP5. Para aumentar adicionalmente a resposta imunitária da vacina de bioconjugado contra *S. aureus*, a proteína transportadora EPA é trocada pela forma AcrA de *C. jejuni* e duas proteínas de *S. aureus*: Hla e ClfA. Para serem utilizadas como proteínas veículo, Hla e ClfA foram modificadas pela inserção dos sítios de N-glicosilação bacterianos. O processo foi realizado como descrito no documento WO 2006/119987, produzindo quatro versões para Hla H351: SEQ ID N°: 6, SEQ ID N°: 7, SEQ ID N°: 8, SEQ ID N°: 16 e três para ClfA: SEQ ID N°: 10, SEQ ID N°: 11, SEQ ID N°: 12.

Para glicosilação do sítio 130 de Hla H351, foram utilizadas células de *E. coli* (W3110 $\Delta waaL$ $\Delta wecA$ $\Delta wzzE$ $\Delta rmlB$ $\Delta wecG$), compreendendo dois plasmídeos de expressão: um para a produção de Hla H351 (SEQ ID N°: 16), no qual a expressão de Hla H35L contendo o péptido sinal N-terminal para secreção periplasmática, um sítio de N-glicosilação e uma cauda de hexa HIS para purificação está sob o controlo do promotor ParaBAD e, em segundo, lugar, um para expressão do aglomerado quimérico de CP5 e pglB (SEQ ID N°.: 17). Este sistema corresponde ao sistema de expressão de 2 plasmídeos otimizado de antemão de CP5-EPA, com um plasmídeo de expressão de veículo proteico trocado. As células foram cultivadas num biorreator de 12 L em meio rico até uma D.O._{600 nm} = 30, a expressão de Hla foi induzida pela adição de arabinose a 0,2%. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol

de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas no sobrenadante foram carregadas numa cromatografia de afinidade de Ni^{2+} . As Hla glicosiladas (CP5-Hla) e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M e carregadas numa cromatografia de troca aniônica. As proteínas foram eluídas com um gradiente linear de 0 a 0,7 M de NaCl para separar CP5-Hla de Hla. A proteína resultante foi separada por SDS PAGE e corada por Coomassie, ou transferida para membranas de nitrocelulose e sondada com antissoros anti-His, anti-Hla ou anti-CP5, como indicado (FIG. 14C). Os resultados na FIG. 14C mostram a formação de glicoconjugado (CP5-Hla) por Coomassie (pista da esquerda) e transferência de western utilizando antissoros anti-His (pista da esquerda do meio) e anti-Hla (direita do meio) e anti-CP5 (direita).

A identidade de Hla H351 com um sítio de glicosilação 130 manipulado foi confirmada por tripsinização em gel e MALDI-MS/MS.

Para mostrar adicionalmente que a proteína transportadora é permutável para glicosilação por CP5 e CP8, a proteína AcrA de *C. jejuni* foi utilizada como um aceitador de glicosilação (ver a FIG. 14D). Utilizando o sistema de 3 plasmídeos (SEQ ID N°: 3, SEQ ID N°: 15 e SEQ ID N°: 27), a estirpe de produção para este conjugado foi W3110 $\Delta waaL$ albergando o aglomerado quimérico de CP5 (SEQ ID N°: 3), a proteína PglB induzida por IPTG (SEQ ID N°: 27) e a AcrA (SEQ ID N°: 15) sob indução de arabinose em plasmídeos separados. As células foram cultivadas num biorreator, com um volume de partida de 7 L em meio semidefinido contendo, como fontes de C, glicerol, peptona e

extrato de levedura. As células foram cultivadas em modo descontínuo ou descontínuo pulsado até uma D.O._{600 nm} de 30 e a expressão de PglB e AcrA foi induzida pela adição de IPTG a 1 mM e arabinose a 10%. Após indução, as células foram ainda cultivadas em modo semicontínuo, durante um período de 15 horas, sob condições limitantes de oxigénio e recolhidas por centrifugação. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. As glicoproteínas CP5-AcrA foram eluídas da coluna de afinidade por imidazole a 0,5 M. A proteína purificada foi separada por SDS PAGE e corada por Coomassie, ou transferida para membranas de nitrocelulose e sondada com antissoros anti-AcrA ou anti-CP5, como indicado na FIG. 14D.

A inserção dos sítios de N-glicosilação bacterianos em ClfA foi realizada como descrito no documento WO 2006/119987, produzindo as SEQ ID N°: 10; SEQ ID N°: 11; SEQ ID N°: 12. As proteínas veículo foram expressas em células de *E. coli* a partir de promotores indutíveis de arabinose. Os genes foram concebidos para produzir um péptido sinal N-terminal para secreção periplasmática, vários sítios de N-glicosilação e uma cauda de hexa HIS para purificação. A purificação foi iniciada a partir de extratos periplasmáticos de células de *E. coli*.

Para glicosilação de ClfA 327, foram empregues os sistemas de expressão otimizados de antemão de CP5-EPA. Utilizando o sistema de 2 plasmídeos (SEQ ID N°: 17 e SEQ ID N°: 11), células de *E. coli* (W3110 Δ wecAwzzE Δ rmlB-wecG Δ waaL) compreendendo o aglomerado quimérico de CP5 e *pglB* (cassete de expressão

constitutiva), assim como o plasmídeo de expressão para ClfA 327 (sob controlo do promotor ParaBAD), foram cultivadas em frascos Erlenmeyer de 1 L em meio de LB. Uma cultura de um dia para o outro foi diluída para D.O. $_{600\text{ nm}} = 0,05$. A D.O. $_{600\text{ nm}}$ de cerca de 0,5, foi induzida a expressão de ClfA por adição de arabinose (concentração final de 0,2%). As células foram cultivadas durante 20 horas. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni^{2+} . O ClfA-CP5 foi eluído por imidazole a 0,5 M, foi separado por SDS PAGE e corado por Coomassie ou transferido para membranas de nitrocelulose e sondado com antissoros anti-His ou anti-CP5. A FIG. 14E mostra os resultados obtidos utilizando o variante ClfA com o sítio de glicosilação inserido próximo da posição de aminoácido 327 da proteína (SEQ ID N°: 11). Eles mostram a formação de ClfA por coloração de Coomassie e transferência de western anti-His e glicoconjugado (CP5-ClfA) por transferência de western utilizando antissoros anti-CP5.

Exemplo 7: Atividade de CP5-EPA como vacina de glicoconjugado

Células W3110 $\Delta waaL$ $\Delta wecAwzzECA::cat$ compreendendo aglomerado quimérico de CP5 (SEQ ID N°: 3) com cap5K no interior, a proteína PglB (SEQ ID N°: 27) e EPA com 2 sítios de glicosilação de sinal em pEC415 (SEQ ID N°: 13) foram cultivadas em frascos Erlenmeyer de 1 L, em meio de LB. Uma cultura foi diluída, de um dia para o outro, para D.O. $_{600\text{ nm}} = 0,05$. A D.O. $_{600\text{ nm}}$

de cerca de 0,5, foi induzida a expressão de EPA e PglB por adição de arabinose (concentração final de 0,2%) e IPTG a 1 mM, respetivamente. As células foram cultivadas durante 20 horas. As células foram sedimentadas por centrifugação; as células foram lavadas e suspensas em 0,2 vol de tampão de sacarose, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. As EPA glicosiladas e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M e carregadas numa coluna de troca aniónica SourceQ. A EPA glicosilada foi separada da EPA não glicosilada por aplicação de um gradiente de concentração crescente de NaCl. As quantidades de proteína eluída foram determinadas pelo ensaio de BCA e, com base no tamanho das bandas obtidas em SDS PAGE corado por Coomassie, a massa teórica do açúcar foi calculada. Juntamente com a determinação de proteína, a quantidade de antigénio polissacárido foi estimada na preparação. Esta quantificação estimada foi confirmada pelo método de MS de maldi de elevada massa (ver a FIG. 14A).

Para medir a imunogenicidade de CP5-EPA em animais vivos, 1 µg do glicoconjugado purificado foi injetado em murganhos pela via IP (intraperitoneal), na presença de Hidróxido de alumínio como adjuvante, nos dias 1 (primeira injeção), 21 (segunda injeção) e 56 (terceira injeção). Após 35 e 61 dias, que foram duas semanas após a segunda e terceira injeções, respetivamente, a resposta de IgG foi medida por ELISA utilizando um CP5 modificado por poli-L-lisina para revestimento (Gray, B.M. 1979. Metodologia de ELISA para antigénios polissacáridos: ligação proteica de polissacáridos para adsorção sobre tubos de plástico. *J. Immunol.* 28:187-192). Sangue de murganhos

imunizados com bioconjugado de CP5 foi analisado para anticorpos específicos de IgG contra polissacárido capsular de CP5. A FIG. 15A apresenta os títulos de IgG deduzidos por CP5-EPA em murganhos. As placas de ELISA foram revestidas com CP5 modificado por poli-L-lisina, a resposta de IgG em murganhos imunizados duas vezes (segunda barra (vazia) a cada diluição) ou três vezes (primeira barra (diagonais para a frente) a cada diluição) com CP5-EPA foi medida em triplicados. Os sinais obtidos com os soros pré-ímmes como controlo estão indicados pela terceira barra (diagonais para trás) a cada diluição. A resposta de IgG dos murganhos foi medida com proteína G conjugada a fosfatase alcalina. Como mostrado na FIG. 15A, o bioconjugado de CP5-EPA deduziu um título de anticorpo sérico de $6,4 \times 10^3$. Os resultados apresentados na FIG. 15A mostram que o CP5-EPA deduz anticorpos específicos de CP5 em murganhos. Esta experiência mostra que o bioconjugado produzido em *E. coli* é imunogénico em murganhos.

Uma experiência semelhante foi realizada em coelhos como o organismo hospedeiro. CP5-EPA (15 µg de CP5) foi injetado em coelhos intradermicamente na presença de adjuvante completo de Freund no dia 1 e subcutaneamente na presença de adjuvante incompleto de Freund nos dias 20, 30 e 40. Após 61 dias, a resposta de IgG foi medida por ELISA utilizando um CP5 modificado por poli-L-lisina para revestimento (Gray, B.M. 1979. Metodologia de ELISA para antigénios polissacáridos: ligação proteica de polissacáridos para adsorção sobre tubos de plástico. *J. Immunol.* 28:187-192). A FIG. 15B apresenta títulos de IgG deduzidos por CP5-EPA em coelhos. Os resultados apresentados na FIG. 15B mostram que o CP5-EPA deduz anticorpos específicos de CP5 em coelhos. A resposta imunitária a bioconjugado de CP5-EPA é a segunda barra (diagonais para a

frente) a cada diluição. Os soros de controlo incluem soros absorvidos específicos de CP5 deduzidos para *S. aureus* neutralizado (extratos de WC, primeira barra (pontos) a cada diluição) e soro pré-imune (terceira barra (vazia) a cada diluição). O soro de coelhos imunizados com diversos antigénios foi analisado para anticorpos específicos para CP5 purificados. As placas foram revestidas com CP5 modificado por poli-L-lisina. Os sinais obtidos com os soros pré-ímmes como controlo estão indicados pela terceira barra (diagonais para trás) a cada diluição. A resposta de IgG dos coelhos foi medida com proteína G conjugada a fosfatase alcalina em triplicados. O bioconjugado de CP5-EPA deduziu um título de 1×10^6 , o qual foi 4 vezes superior ao título de soros de controlo (preparados por imunização com *S. aureus* neutralizado intacto e, depois, absorvidos com Wood 46 e um mutante acapsular isogénico tripsinizado, de modo a que o antissoro fosse tornado específico de CP5). Esta experiência mostra que o bioconjugado foi capaz de deduzir uma resposta de IgG específica de CP5 de elevado título.

Exemplo 8: Atividades funcionais de anticorpos de CP5

Atividade *in vitro*

O antissoro policlonal de coelho deduzido como descrito no Exemplo 7 foi purificado por coluna de afinidade de Proteína A, para enriquecimento para anticorpos específicos de IgG. IgG de coelhos imunizados com o bioconjugado de *S. aureus* CP5-EPA foi testado para atividade funcional num ensaio de neutralização opsonofagocítica clássica *in vitro* (Thakker, M., J.-S. Park, V. Carey e J. C. Lee. 1998. O polissacárido capsular de serotipo 5 de *Staphylococcus aureus* é antifagocítico e melhora a virulência

bacteriana num modelo de bacteriemia murina. *Infect Immun* 66:5183-5189). O *S. aureus* foi cultivado durante 24 h em ágar Columbia + NaCl a 2%. As bactérias foram suspensas em meio mínimo essencial + BSA a 1% (MEM-BSA). Os PMN (neutrófilos polimorfonucleares) foram isolados de sangue humano fresco, lavados contados e suspensos em MEM-BSA. As preparações de IgG purificadas de coelhos imunizados com CP5-EPA de *S. aureus* ou como preparações de IgG purificadas de controlo de coelhos imunizados com 01-EPA de *Shigella* que foi purificado como descrito no documento WO 2009/104074, foram adicionadas ao ensaio em diluições seriais de fator 10 preparadas em MEM-BSA. Soro de cobaio (Pel-Freez) foi utilizado como uma fonte de C'. Cada ensaio (volume total de 0,5 mL) continha $\sim 5 \times 10^6$ PMN, 1×10^6 CFU de *S. aureus*, 0,5% a 1% de soro de cobaio e concentrações variáveis de IgG, variando entre 140 µg/mL e 1 µg/mL. As amostras de controlo continham 1) *S. aureus* incubado com C' e PMN, mas sem anticorpo; 2) *S. aureus* incubado com IgG e C', mas sem PMN; e 3) *S. aureus* isoladamente. As amostras foram rodadas pelas extremidades (12 rpm), durante 2 h, a 37 °C. As diluições das amostras foram agitadas em vórtice em água estéril e a neutralização bacteriana foi estimada por plaqueamento das amostras diluídas, em duplicado, em TSA. A percentagem de neutralização foi definida como a redução em CFU/mL após 2 h em comparação com aquela no tempo zero.

No primeiro conjunto de experiências, foi testada a neutralização opsonofagocítica da estirpe Reynolds de *S. aureus* sensível a meticilina (MSSA), o isolado protótipo de CP5, e os resultados estão mostrados na FIG. 16A. A atividade opsónica de anticorpos para CP5-EPA deduzidos em coelho foi testada contra a estirpe Reynolds de *S. aureus* serotipo 5. Os anticorpos de CP5-EPA eram opsónicos até uma concentração de 1,4 µg/mL,

enquanto que os anticorpos de 01-EPA mostraram pouca atividade opsonica a 140 µg/mL. Um soro de controlo positivo deduzido contra extratos de células intactas de *S. aureus* (obtido de J. C. Lee no Departamento de Medicina, Brigham and Women's Hospital, Faculdade de Medicina de Harvard, Boston, MA, EUA) mostrou atividade semelhante ao soro anti-CP5-EPA (antissoro WC, 1%).

Como mostrado na FIG. 16A, entre 65-75% de *S. aureus* Reynolds foi neutralizado pelas PMN quando incubado com anticorpos para CP5-EPA e soro de cobaio a 1% com atividade de complemento. O antissoro foi utilizado a 1% final no ensaio e 89% do inóculo de *S. aureus* foi neutralizado sob estas condições. Foi observada pouca neutralização quando o *S. aureus* foi opsonizado por C' isoladamente (soro de cobaio a) ou anticorpos e C' sem as PMN. Os dados mostrados são as médias de 2 a 5 experiências. Todas as amostras em gráfico incluíam soro C' de cobaio e, na ausência de C', não foi observada neutralização. Nem os anticorpos isoladamente nem o complemento isoladamente eram opsonicos e esta particularidade é característica de patógenos bacterianos encapsulados. Em contraste, anticorpos deduzidos pela vacina de controlo (antigénio de *Shigella* 01 ligado a EPA) não mostraram atividade opsonica, mesmo na presença de C'. Como um controlo positivo neste ensaio, a requerente também testou antissoro de coelho específico de CP5 (obtido de J. C. Lee no Departamento de Medicina, Brigham and Women's Hospital, Faculdade de Medicina de Harvard, Boston, MA, EUA). Estes dados mostram que os anticorpos deduzidos para o bioconjugado de CP5-EPA mostraram atividade opsonica contra *S. aureus* encapsulado que é comparável a anticorpos de CP5 com atividade opsonica documentada (Thakker, M., J.-S. Park, V. Carey e J. C. Lee. 1998. O polissacárido

capsular de serotipo 5 de *Staphylococcus aureus* é antifagocítico e melhora a virulência bacteriana num modelo de bacteriemia murina. *Infect Immun* 66:5183-5189).

A atividade opsônica de anticorpos para CP5-EPA testou contra a estirpe USA100 de MRSA de CP5-EPA. A FIG. 16B apresenta os resultados da atividade opsônica de IgG e C' testada contra a estirpe USA100 de *S. aureus*, um isolado de CP5+, e é denominada NRS382. Os dados mostrados são as médias de 2 a 5 experiências. Todas as amostras em gráfico incluíam soro C' de cobaio e, na ausência de C', não foi observada neutralização. Como mostrado na FIG. 16B, ~60% do inóculo de USA100 foi neutralizado pelas PMN incubadas com complemento de cobaio a 0,5% e concentrações de IgG de CP5-EPA variando entre 100 a 1 µg/mL. Foi observada neutralização mínima na ausência das PMN ou quando a IgG foi omitida do ensaio. Não foi alcançada neutralização quando a IgG deduzida para a vacina conjugada de O1-EPA foi adicionada às PMN + C' (nesta amostra, as bactérias multiplicaram-se). Foi observada pouca neutralização quando o *S. aureus* foi opsonizado por C' isoladamente ou anticorpos e C' sem as PMN. Assim, os anticorpos de CP5-EPA eram opsônicos a concentrações que variam entre 100 e 1 µg/mL, enquanto que os anticorpos de O1-EPA mostraram pouca atividade opsônica a 100 µg/mL. Esta experiência mostra que os anticorpos de CP5-EPA exibem atividade opsônica contra as estirpes de MSSA e de MRSA.

Atividade *in vivo*

Para determinar se a atividade opsônica de IgG deduzida para a vacina do bioconjugado CP5-EPA conferia proteção num modelo de murganho de infecção estafilocócica, foram realizadas

experiências de imunização passiva. Nos estudos iniciais, murganhos Swiss-Webster macho (~6 semanas de idade) foram injetados IV (veia da cauda) com 1,4 a 2 mg de IgG de coelhos imunizados com CP5-EPA ou O1-EPA de *Shigella*. Após 24 h, os murganhos foram estimulados pela via intraperitoneal (IP) com $\sim 3,6 \times 10^7$ CFU de *S. aureus* Reynolds. Os níveis de bacteriemia foram medidos 2 h após o estímulo para avaliar a eliminação mediada por anticorpo da bacteriemia. O limite inferior de detecção pela cultura era 5 CFU/mL de sangue. A FIG. 17A mostra os níveis de bacteriemia resultantes. Cada ponto representa uma cultura de sangue quantitativa realizada por perfuração da veia da cauda num murganho individual 2 h após inoculação bacteriana. As linhas horizontais representam valores medianos de CFU/mL. Os círculos vazios são amostras de sangue de murganhos que obtiveram anticorpos anti-CP5-EPA, os círculos cheios a negro são amostras de animais que receberam uma preparação de anticorpo de controlo que foi deduzida contra EPA conjugado a um glicano diferente (O1 de *S. dysenteriae*). Os resultados da FIG. 17A mostram que os murganhos que receberam anticorpos de CP5 mostraram uma redução significativa ($P = 0,0006$ pela análise de Mann-Whitney) nos níveis de bacteriemia, em comparação com murganhos que receberam os anticorpos específicos de O1. De facto, a redução em CFU/mL de sangue foi 98% em murganhos imunizados passivamente com IgG de CP5-EPA vs. murganhos que receberam IgG de O1-EPA.

Em experiências subsequentes de imunização passiva, os murganhos foram estimulados IP com um inóculo inferior ($\sim 5,5 \times 10^6$ CFU/murganho) de *S. aureus* Reynolds. A imunização passiva com anticorpos de CP5-EPA foi testada em murganhos estimulados IP com $5-6 \times 10^6$ CFU de *S. aureus* Reynolds. Os murganhos foram injetados intravenosamente (IV) com 2 mg de IgG

de CP5-EPA ou IgG de O1-EPA, 24 h antes do estímulo bacteriano. A FIG. 17B mostra os níveis de bacteriemia resultantes. Cada ponto representa uma cultura de sangue quantitativa realizada por perfuração da veia da cauda num murgancho individual 2 h após inoculação bacteriana. As linhas horizontais representam valores medianos de CFU/mL. Os círculos vazios são amostras de sangue de murganchos que obtiveram anticorpos anti-CP5-EPA, os círculos cheios a negro são amostras de animal que recebeu uma preparação de anticorpo de controlo que foi deduzida contra EPA conjugado a um glicano diferente (O1 de *S. dysenteriae*). Como mostrado na FIG. 17B, os murganchos que receberam 2 mg de IgG de CP5-EPA tiveram níveis de bacteriemia significativamente ($P < 0,0001$ pela análise de Mann-Whitney) inferiores aos animais que receberam 2 mg de IgG de O1-EPA. De facto, 6 de 7 murganchos imunizados passivamente com anticorpos de CP5-EPA tiveram culturas de sangue estéreis (limite inferior de deteção 6 a 30 CFU/mL de sangue, dependendo do volume de sangue recolhido e plaqueado de cada murgancho). A redução nos níveis de bacteriemia atribuível aos anticorpos de CP5 foi 98%, em comparação com murganchos de controlo que receberam IgG de O1-EPA.

Para determinar se a proteção contra bacteriemia poderia ser conferida por um nível inferior de IgG, foi realizada uma experiência subsequente em que os murganchos foram imunizados passivamente pela via IV com 300 µg de IgG de CP5-EPA ou O1-EPA. Após 24 h, os murganchos foram inoculados IP com 6×10^6 CFU de *S. aureus* Reynolds. O limite inferior de deteção pela cultura era 13-67 CFU/mL de sangue. A FIG. 17B mostra os níveis de bacteriemia resultantes. Cada ponto representa uma cultura de sangue quantitativa realizada por perfuração da veia da cauda num murgancho individual 2 h após inoculação bacteriana. As linhas horizontais representam valores medianos de CFU/mL. Os

círculos vazios são amostras de sangue de murganhos que obtiveram anticorpos anti-CP5-EPA, os círculos cheios a negro são amostras de animal que recebeu uma preparação de anticorpo de controlo que foi deduzida contra EPA conjugado a um glicano diferente (O1 de *S. dysenteriae*). Como na FIG. 17B, os resultados da FIG. 17C mostram que a proteção mediada por anticorpo de CP5 contra bacteriemia foi alcançada com esta dose inferior de anticorpo. Foi alcançada uma redução de 98% nos níveis de bacteriemia por anticorpos deduzidos pela vacina de bioconjugado de CP5 e 8 de 9 murganhos tiveram culturas de sangue estéreis, em comparação com 0 de 8 murganhos que receberam anticorpos de O1-EPA de *Shigella*.

Exemplo 9: Imunização ativa em murganhos

Para mostrar que a vacinação de murganhos com o bioconjugado CP5-EPA medeia proteção contra estímulo bacteriano, como no ensaio de imunização passiva, foram realizados estudos de imunização ativa.

O bioconjugado de CP5-EPA foi produzido em W3110 $\Delta waaL \Delta wec \Delta wz z E::cat$ por co-expressão do aglomerado génico de CP5 quimérico (SEQ ID N°: 3), PglB (SEQ ID N°: 27) a partir do plasmídeo pEXT21 e EPA (contendo dois sítios de glicosilação, SEQ ID N°: 13) de plasmídeos separados. As células foram cultivadas num biorreator, com um volume de partida de 7 L em meio semidefinido contendo, como fontes de C, glicerol, peptona e extrato de levedura. As células foram cultivadas em modo descontínuo ou descontínuo pulsado até uma D.O._{600 nm} de 30 e a expressão de PglB e EPA foi induzida pela adição de IPTG a 1 mM e arabinose a 10%. Após indução, as células foram ainda

cultivadas em modo semicontínuo, durante um período de 15 horas, sob condições limitantes de oxigénio e recolhidas por centrifugação. As células foram lavadas e suspensas em tampão de sacarose a 25%, a uma D.O._{600 nm} = 200, sedimentadas e lisadas por choque osmótico. Os esferoplastos foram sedimentados por centrifugação e as proteínas periplasmáticas foram carregadas numa cromatografia de afinidade de Ni²⁺. As EPA glicosiladas e não glicosiladas foram eluídas da coluna de afinidade por imidazole a 0,5 M e carregadas numa coluna de troca aniónica SourceQ. A EPA glicosilada foi separada da EPA não glicosilada por aplicação de um gradiente de concentração crescente de NaCl.

Pretende-se que CP5-EPA seja utilizado como uma vacina conjugada para proteger contra estirpes de *S. aureus* de CP5. Para testar se essa imunização ativa é funcional, a requerente imunizou diferentes grupos de murganhos Swiss Webster fêmea com três doses diferentes de CP5-EPA e analisou a imunização utilizando um modelo de bacteriemia. Três doses foram injetadas subcutaneamente nos dias 0, 14 e 28. Os murganhos foram estimulados intraperitonealmente no dia 42 com a estirpe JL278 de *S. aureus*, como mostrado na FIG. 18. Cinco grupos de murganhos foram imunizados com três doses diferentes de CP5-EPA, como indicado debaixo do eixo x (círculos ponteados; círculos vazios; e diagonais para trás em círculos). Dois grupos de controlo receberam adjuvantes (diagonais para a frente em círculos) ou PBS (círculos cheios a negro) isoladamente. Cada ponto representa uma amostra de sangue de um único murganho. A dose mais baixa de vacina (0,2 µg) induziu proteção em todos os murganhos do grupo. Duas horas após estímulo, as amostras de sangue foram analisadas para formação de cfu e anticorpos anti-CP5 por ELISA, utilizando um CP5 modificado por poli-L-lisina para revestimento (Gray et al. (1979)). Em todos

os grupos imunizados com CP5-EPA, foi observada uma redução média de cfu no sangue. Contudo, apenas no grupo que recebeu a dose mais baixa de vacina existiu uma proteção geral da bacteriemia em todos os cinco murganhos. A análise de sangue para anticorpos anti-CP5 resultou numa correlação positiva de proteção e títulos médios de ELISA nos diferentes grupos de murganho. Os resultados apresentados na FIG. 18 indicam que os anticorpos induziram a proteção da bacteriemia em murganhos imunizados.

Estes estudos indicam que a vacina de bioconjugado de CP5-EPA induziu anticorpos que opsonizaram *S. aureus* para neutralização fagocítica pelas PMN humanas e protegeram murganhos contra bacteriemia em estudos de imunização positiva e ativa. Estes dados proporcionam forte evidência de que o bioconjugado apresentado irá proteger contra as doenças provocadas por múltiplas estirpes de *S. aureus*.

LISTAGEM DE SEQUÊNCIAS

<110> GLYCOVAXYN AG

<120> VACINAS DE BIOCONJUGADO DE BACTÉRIAS GRAM-POSITIVAS
CAPSULARES

<130> P51971WO

<140>

<141>

<150> 61/332,170

<151> 2010-05-06

<160> 27

<170> PatentIn version 3.5

<210> 1

<211> 13369

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido
Sintético"

<400> 1

aattcacatg ttgcccatcc acgaaaccac cttatcgccg tggaacgcac ctggatcgac	60
agccccagca aagcagtcgc ttccctggtcc ggcaccggaa acatcgtacg gagaaaacaa	120
aaaaggccgc taggcggcct tttccggaga acgatgactc agggttctcg ccgcctctgg	180
cgatagatcc agtcgacgat ttcaccgtca ggcgcatagc cgctgacggg ttcccgcagc	240
aactggcgaa cccgcgagta gtcgtccttc tccacggcgg ccagcaactg ctccagcacg	300
accttgaagg cctcccagct caggtgttcc tcggtggccc gcatgatcat cggatggtcg	360
gtgggattca cgttgtcacc gatcagcagc tcttcgtaga gcttctcgcc aggacgcagg	420
ccactgaact cgatggcgat gtcaccatgg ggcgaacgct cggaacgcac gctcaggccg	480
gacaggtgga tcattcttctc ggcgagctcc aggatcttca ccggcgggcc catgtccagc	540
acgaatacat ctccgccttg ccccatcgaa ccggcctgga tgaccaactg cgccgcctcg	600
ggaatggtca tgaagtaacg ggtgatgctc ggggtgggtga ccgtcaccgg gccgccgcgc	660
ttgatctgct cgcggaacag cggaatgacc gaaccggacg aaccgaggac gttgccgaag	720
cggaccatgg tgaacgggt cttgttgacg tgatgcacgt cttccggtc gccgaacagc	780
accggcgccg attcgttgct gagcgctga aggaccattt ccgccaggcg cttggtgctg	840
cccatcacat tggtcggccg caccgccttg tcggtggaaa tcagcacgaa gttctgcacg	900
ccgacctgca ccgcggcctg caccgcatgc aaggtgccta tcacgttggt gagaacgccc	960
tcggcgatgt tgtgctcgac gatcggcaca tgcttgtagg ccgccgcatg gtagacggta	1020
ttgaccttcc aggtacgcat cacgtccacc aggcgctcgg gattgcgcac cgaaccgagg	1080
atcggcaaca ggttcaccga aagcgactcg cgcttgatcc gacgctccag ttccctgatgg	1140
atgctataga ggttgatttc gctgtgttcg aacaggatca gcacgctagg cgaacaactc	1200
atgatctgcc gacagagttc cgaaccgata gagccgcccg ccccggtcac catcaccacc	1260
tgaccgcgga tgcaccgttc cagcagctcc ttgcgcggtg cgacgctgtc gcgcccagc	1320
aggtcagcga tgtccacctc ctgcaggtca tccaccttga cccggccgct ggcaggtcc	1380
atgaagccgg gcatgctgcg cacgtgcagc gggaaacggct ccagggactc gagaatctct	1440
cggcgccggg cccgagtggc ggaaggaatc gccaggagaa cctcctgcgc gccgctctcg	1500
tcgatcatct ggcggatatg cttggcggtg tagaccgcga gaccggcaat gaccgggttg	1560
gcgatctgct tgtcgtcatc gatgaacgcc accggacgca tcgcccgacc gagacgcaac	1620
gccgcaacca actggttgcc ggccgcccc ccccataga taaccaccct gggcaggcca	1680

tcctggcgggt tgagaaatgg taccgactgc acagcagagt accagtgcgc catgaaatac	1740
tggcgcatgg ccagacgcaa gccgccgatac agcagcatgc tcaaccacca gtagttgaac	1800
accaggggaac gcggcaccgg cgccggcgcg ccacgatacc agtacaccac cagcgacagc	1860
accagagccg agatggtcac cgcccttggcg atggcgatca atgcgtcgtt accgagatag	1920
cgcatacccg cgcgatacat gccgaagcga atgaatagtg gaatggcgat gaccggcgca	1980
gtgatgaaaa gccatgcatg ctgcgcgaac acgtcgatca tatcgtctgt gcctagacgg	2040
accacaaaag cgagccacag agacagccat accagaagga tatccgtagc cacttgaagc	2100
aaacgtttcc agcgacgagg catggataac aactttactc ttaaaccgat tagcattccc	2160
ctactccttt aggcctgagc ggacgactct agcttaccgc ccttgaacct cactgccaga	2220
aaaacgagag ggacgtaggc tatgacgatg cccatcaaag gctctagaga cccactcccg	2280
accaacaagg ctataggag gagccagatg acgttcaatg ccgtgacgcc aatcgtaaca	2340
ggagcatgct ttccatagta tcggcttgcg tattgatagg catggcttcg atgagcctca	2400
tacaccttgt cccctctcag caagcgacga atcaggggat aggtcgcatc gacgatgaaa	2460
acgcctaaca ggaccaacca tgcccagaaa aaattcgtat tcatccacat ggcatgaatg	2520
gaaagaattc ctaaaacaat cccagaaga ccaactaccg catctcccat gaaaattttt	2580
gggggtggaa aattccagaa caagaatcca aaaacggcaa aagctagcga taaggggagc	2640
aaagcctgog tcagttggcc attcagccag tataataatg cccaccaac acagacaaaa	2700
atggcctgaa ggctagcaag tccatcgatc ccatccatga agttatagag attcagcaac	2760
cacacgagat agaaagcaaa gagaagtctt ccgaaccacc ccaggtccca actgacacca	2820
acaatctgaa atgggtggtat gccattcaaa aagtacaatg agaccaaggc tgcactaaaa	2880
tgtcccagca gacgccagcg tcgggcgata tgaccgtgat catccatgaa tccaataatg	2940
gcaaccccg cgcagcgag aagtaaagcc caggacacgg cccatgagat atttcccagt	3000
acagcccaaa taggcagcat gagacaaaaa gtaactacaa tggcgacccc tcctccgcgt	3060
ggagtgggaa cgacatggga acttcgagag ttaggggtgt caagtaagct cctcgctaag	3120
gcataacgac gcaagaggcc tgtaagcagt cccgaaaccc cagctgcagc gagtaacaaa	3180
taccattctt ccatttataa atttatcccc aaaaatactc tcacccttc aacagggttc	3240
tattcgattg ttctgcgtgg tacgccatgc tagtagcaaa gcgcctgcag ataatcattc	3300
agtacgtttg acatatattc taccaacatc ttccaacgcg gaaaggggtgc tcttgcgggg	3360

ttcccaacca agcaaaagcc tgccttttoga cgagtcgacc tctagtgagc agcataactg	3420
agtgtataac ccaccctttc ctactaattt aagaaacctc agaataaacc taggaacagg	3480
ccacatgatg gggcgacgcc ccattcccgcc agcaagcgca gtcaccagtt gcttggtaga	3540
aatctcctga ccatcgata ccaaaaacag ttgcgcggca gcagaagggt gcgatcataca	3600
gcaagctaga aagtcaacta aattatccag agaaacaaaa cttcgtcggg tatctatgca	3660
accaaagga agaggaagtc ccgaagcaac cagcttcaac aatcgcgaga aatttccagg	3720
agctttccag tcgtaaacga gtggaggcct gacgataaca agttctgttg aggaatgctt	3780
gaaaagctcc cgaagcgcta cttcagcctc aaactttgaa atcgcatatt ctgcatgagg	3840
agccggcttg gagtctcat cgaagggtt ttctttggtt aaagcgccat ttacaccaat	3900
agaactaaca aaaatgaaac gctttacaga tgcttcgac gccgcccag caagcgccag	3960
agtagcatcg cgattcacct tccgaaaaat atctagtga tcaagctgcc ttccaaagat	4020
atgggctcgt ccagctagat gaactacaca ttcaacaccc cgcagcgag catcaagctt	4080
agtgtctctt ttcagctccg ctogaacata ctcaaccccc gtaacgggat tgtacaggga	4140
tcgtacttgt ccgacaacct gaaagggggc ggcagcaagc gacctgcaa ggcactccc	4200
gacaaaaccg ctagcccggt ttaccagcac cttcatcata ttttcgactc agacagaagc	4260
ggtcgaacac aggcagcgaa tttttccaat gaaatatatt cagaataatt ttgctgaagg	4320
acactccgag ggactccgcc taaactagac aagtcgagtt cacaaatctc atcgatcaat	4380
cttgccaaag caaccggatc attcggcgga cagttccaac cgattccggg ctcatctatt	4440
atacgggaga tttcagcccc tttttccatg acagctagaa tcgggtttgtc tgctgccatg	4500
gagaaatatg ccttgctggg aaccccaagc ccgaacattc cttcttctaa ggtaactaag	4560
gcaacgtcac aggcagccaa accaaaattc ttttcggcta atggcagcct tccaaaatac	4620
ctcaaccgag cacactgac ttccagcgcg tgttttttta cactgtogac caaggcacca	4680
tctccaataa aagcaaaagc cgccttctcg tttttaacca actgaatagc agaaagtatg	4740
ttttctatac cttgtaatcg accgacatta ccaaaaaatt ggaaaaccct tttacctttc	4800
cattcaggaa tattgataaa aggagcatcc tctctcggtc ctgggaaaac ctctttctca	4860
caggcccaat tcgaaataaa gaccaaagat cgcgggtcat tcacctctc tttcataaga	4920
gcttccatat cgcgccctat tacgactaga cgatcagcgg atgagtaa atgaaagaaaag	4980
agacgacgta gaagccggta ggcaatacta tctttcttca gaacgcccgc cggcaccaag	5040

ttctcgggaa acacatcatg caccagcagc acccacttga aaccgagggc ataccttagc	5100
aagggaaaacg tcatcagtag aagagcaggg ttggttccac tcaataccac atctcctcgt	5160
ctggcacgag aagtcaattt tactgagaac agaaacgcct gaaaaatctg cgccaatcct	5220
ctagaaagga gcctattctt attgctcctc ggaaacttaa agcattcttg ctcaacggaa	5280
ggtgaaacca caggcggggg ttcacctgcg gtcagcggga aaattacggt tagcccacca	5340
aactcccgct gcctctttcc tattatcttc tcccaatagt atcccgaggga gttctgattg	5400
gcaccgacat actcagaaac cacaatatc cttgccatca actccacgcc tcaagttaat	5460
atTTTTTcca gacagttcgc atgacatagt cccgatagct atgaactatg cgaacaatct	5520
tttcagaaac gttgggcatg ctatagtcag cgaccaagcg taacatgcgc tctgcgtcgc	5580
gcctctgtcc ctccaacacc tcgagtgcct gtagtactcg atccgaatcc agtccgacca	5640
tcatcaccac agcctcttcc atgccttctg ggcgttcag agcctcacga atattcaaag	5700
cgggaaaatt cagtattgaa gactcctcac tgatgggtccc actgtcagaa ataaactgcct	5760
tggctgtaat ttgcagttta ttgtaatcct taaagccgag gggtttcagc agtttaatac	5820
cctcgtgaaa ctttgcctcc gtcgcctcaa ttctcttttt tgttctaggg tgagttgata	5880
cgatgacggg cagcgagtac ttttctgcca cagcgttgag catagaaacc aacttcaaga	5940
aattcttatc cgaatctatg ttttctctc ggtgcgcact cacgacaaag aaccgctctg	6000
ttttcaacc cagcctttca agaatatcg aggactcgat cccgtcacga tagtgctcga	6060
gaacttcgaa catagggcta ccagtcttga taaccatgtc tggagaaagt ccttcacgca	6120
agagataatc acgcgcaatt gtactatagg tcaaatttac atcagctgta tgatcgacaa	6180
tgcgccgatt tatctcttca ggcacacgca tatcgaaaca gcgattgcct gcttccatat	6240
gaaaggctcg tatcttacgc cgttttgag gtagtaccgc catacaacta ttggtatcac	6300
ccagcacgag cagcgcatcg ggatctatct cgcgcagaa acgatcgact gcgattatta	6360
cattccctat cgtttcagcc ccggaagacc cggcggcgtt tagaaaataa tccggctttc	6420
ttataccgag gtcttgaaaa aatatttcat taagttcgta atcataattc tgtccagtat	6480
ggacaagtac atgatcgag tactgatcaa gcttcgccat gacctagac aagcgaataa	6540
tctcaggacg agttccaaca accgtaacga cttttagctt ctgcattgtt atctcactat	6600
accttacgca ccttcgccta ctgaacaagc gtaggtatcc ggattttccc gatcaaatac	6660
ttcgttagcc cacaacatga ctaccatata gtcagttccg acattagtaa tgtcatgagt	6720

ccatccaggt	actgttttoga	caattttctgc	ctttttcacca	ttagtgcaaa	tttcgtaaaa	6780
tgccccgggc	aggatgtttc	taaacttgaa	acgtgccatc	cctttgataa	ccagaaaactt	6840
ttcggttttc	gagtgatggg	aatgcccggc	cctggtaaca	cctggatgag	ccgtaaaaaa	6900
cgagaactgg	ccagagtccg	cggctctcag	catctcgacg	aatgtgccac	gcggatccga	6960
atgcattggc	acgtcgtaac	taaaactatc	ttctggcaag	aaacttagat	aagtcgagta	7020
caaggcgcgc	gtcaatcccg	agccaaccct	tgcggtagtc	agtgaacttc	gactattacg	7080
aaactcatat	aattgttctg	cgagctcacc	aacagaaatc	tgatactggg	gctcgacctg	7140
tagtgaaact	gcattggata	gcttcccatc	catgaacttc	atgaagggtc	gaaccacatc	7200
atctatgtat	acaagagtga	tctctgccga	ggaattgtta	atttgaatcg	gaatatctcg	7260
aataatatta	tgacaaaaag	tgcgaaccgc	tgaattataa	ttcggacgcg	accatttacc	7320
gaatacatta	ggaaggcgaa	atatgtagac	aggacaacca	atatacctac	ctagcacttg	7380
gagatgtctt	tctgoggctc	gcttgcttaa	accgtactca	ttatccacct	cagcctgaat	7440
ggatgaagta	taaagaagtg	gtatggctcg	tccattggac	cttaccgcct	cacacagagc	7500
atacgtgagt	toggaattcc	cgatcttaaa	ctcttctggg	ttttccggac	gattgacccc	7560
ggcaagatga	aaaataaaat	cgacggaacg	aattagctca	ggcaaattac	caacactact	7620
ctcgogggtg	aatggcacca	cctcgatacc	accccgctct	gcaagatgag	cgcacagatt	7680
ccttccaaca	aatccattcg	cgccagttac	aagaactttc	atcgtttatt	cctctggact	7740
ggcactctcg	ccacgctgaa	tagcacgaat	gaaatccaac	ttcagcaaca	gctttttcat	7800
tccttcgata	tcagacggtt	tggtattatg	agaattatag	tcctctgtat	gagtaatttt	7860
ttcctcgctt	tgctccacaa	acttactata	gttcagatca	cgcaaactctg	gggggatacg	7920
ataatagtca	cccatgtctt	cagcacaggc	catttcctct	cgactaagaa	gcgcctcata	7980
aagcttctct	ccatgacgcg	tacctattac	attgatagga	taaccattct	tgccaagcaa	8040
ttgagtaagc	gcatgagcca	gcacctcgat	ggttgacgcc	ggtgctttct	gtacaaaaag	8100
atctccattg	gtaccatgct	cgaaagcata	aagcacaagg	tcacggcat	ccgtaagcgt	8160
catcatgaaa	cgtgtcatgt	ttggatcagt	gattgtgaga	ggctggcctg	atcgcatttg	8220
ctcgataaag	agaggaatga	ccgagcccct	tgaagccatg	acgttaccat	aacgggtgcc	8280
acaaattacg	gtaggagtgc	gttccagggt	tcgagacttg	gagaccatga	ccttttccat	8340
catggccttt	gaaataccca	tggcattgat	tgggtaaact	gccttatccg	tactcagaca	8400
aacgactttt	ttgacgccat	tctggatagc	agattcgagg	acattttccg	ttccgatgac	8460
attggtcttc	acagcctcca	tcgggtagaa	ctcacaagag	ggaacctgtt	tcaatgcagc	8520
cgcattgaaa	atgtagtcca	caccgcgagt	agcattcaga	gtgctttgat	agtcgcggac	8580

atctccaata	taaaacotca	acttggggtg	agcatagcac	ttacgcatat	catcttgctt	8640
cttctcatcc	cgactgaata	caggtatttc	accaatatct	gtatccagaa	aacgcttcaa	8700
aacggcattt	ccaaaggaac	cagttccacc	ggtaattaac	agaacagagt	tcttatccat	8760
acaccacctc	tttactaatg	tgttcaacgt	cocttttccg	cctcaaaata	taaagaaagc	8820
gaaaaacaac	gagccagaac	aatgcagccc	ctaccggatg	gaaaagccca	ccgcccgtaa	8880
agacataatt	aacgaaccag	attataaatg	cagagcataa	gaagcctaac	tctgtacaag	8940
ctgtttttct	tcgccggaag	acaagcaaaa	atacagagta	gacaaaaagc	aacaatagca	9000
aaagaacagc	caagccacca	aaaatgaaac	catagatgca	aagcagctcg	acaataccat	9060
tgtgcgccat	tggtatttgc	gcatacccta	gaaccccccc	tccccaaaaa	acgcccgcgc	9120
ccaaaacaag	tatctcgcta	ccgtagaaaag	acaatgtttc	caaaggcagt	gcagttaccc	9180
gatogaatat	ggattcgaag	aatggaacat	cttgcgagac	tgaggaatcg	gaggtgagga	9240
gcccgcacaat	cagtccgatgc	acggtttcca	gactgccagc	cccgaagtct	gcaaggaaaa	9300
atacagataa	accagccct	gcaagaagga	aggtgaaaaa	caagtatgtc	tttttgatca	9360
gttttagcca	ataggcaaag	aataagacaa	tgacgatatt	ggccagagca	gctttctgca	9420
gcgacagcat	cgcccccaag	caaagaagaa	ggaaacatac	cacgcgcgcc	caacccttta	9480
gataaaagag	tgaagcaagt	gcgggaactc	ccaccataac	accatatgcg	gtcaaacttc	9540
ctaccaatga	tgcaaaacgg	gcgccccccag	cacgctcact	ggcttcagcg	aaccattcta	9600
taggtccaaa	tacatactgc	cagagaaaag	aaagagccgc	cagactgaag	aaaaatgaaa	9660
taagttgcat	gggcactttt	agattatctc	tggcaagaac	catagcaaca	aagagaagag	9720
gtaatgatgc	ccacaacctg	aaaatcccta	tggaaatagtc	tccatataga	tcgaaccata	9780
acaccgaaac	gagcgaggca	tacgtccaga	agaaaataac	aaatatacct	aatgcgctat	9840
atttaatatc	cctccagtta	ttaataaata	agagaaaaaa	cagcccacaa	aaccagcaa	9900
gaactaaata	atacgaccgg	gatagattcc	ggagaatatg	aaagagaaac	acaaaaagga	9960
gaaccagagc	aacgcttctc	tgcttcatga	gtaccaccca	tctttgtttg	ctttctcttt	10020
ggaaaagtttc	gactttctat	cgagaacccc	cttccgatag	gatgaggcca	aaacgagcct	10080
ataagcagct	ctagagtaga	gccatogaat	agagtccgtt	ttctctagag	aatgaataag	10140
aatatttgaa	aactttgact	ccagcgcaag	acacagcact	gctctgaata	gactctgaaa	10200
cttgcttgac	aaagcaaagc	ccggctccat	ccatcgtata	ggaccaagag	tctcaaggtc	10260
atcagttggt	gcaactttca	tgaattccgg	cattccgtca	cgggcccgc	gatagtactt	10320
gcgacagaac	ttcgaaatac	tgcttccgg	ttcgacatgc	tcgatttcag	cccgcacaag	10380
aagaacagaa	taaccaagct	tgccaaccct	taacccaaac	tccgtatctt	caaagccgta	10440
cccaacaaat	cgctcatcaa	aaagagcatt	atcttggagc	agaatatccc	tctcagcgag	10500

catattcatg gtaacaatag ttttataccc caaggagttg ggcagcggct tgtcagcaga	10560
gaagtgacaa ctatccctat agogatagta attactcgtc gacaccaatg agcaagaaaa	10620
tctgacacca ccgcaccaga taacgcgttc cccccttcga gagttagcag cgttcagaaa	10680
attcgagaga tgagcaggat ctggcacaca atcatcatca agaagcacta cgtagcgccc	10740
cttcgctaaa cgaacgcctt tatttcgttt tgcactagcc gaattcagct cggcctgctc	10800
aagacggacg ttgaagccct gaaccgcttc gaactcggg accacatcgg gtgtttcatc	10860
accactatta ctgtcgacaa caataacttc aaaatcctta aaagccagag actgcgctac	10920
aaggccggcc aggaccaatt ttagttcctt gggacgacgc caagtactaa cgacgacact	10980
caacaacatc aaagtccctt tttctctctc aacttaaccg tggccaggaa aaacattccc	11040
atcccccagg agacctctgt aacgaccagc gtccacgctc caaaaagctc atcccccaa	11100
agcgccaaga ggataaatgt taataccccg cccaattccg caaaaaaat tgcctcgaaa	11160
tatacaccat catacccaga aggaacaaga gtcaaccctc catacagaac accaatacaa	11220
gcaaacactg gcactatcga aaacattctt attacaacag ttagacctct ggactgttcc	11280
gggaatagga gatatgaaat gtattcagaa aataagaaaa gaaataaaca actaggaata	11340
gotattagta acatcagacc tagcgctttc cttcttaaac tgcgactgc cgggtcggca	11400
ctcgcataca tctgctaaa agtcggaaat agagcgctag ctatgggtga agtagcagcc	11460
gcgatccgc gaaggaattt atccgcagta gaaagcactc cggccgcgga ggctccacca	11520
acaacaccaa ccgcgcgaac gagcacttgc atgtgcaagc taagaaaagc cagagaaaga	11580
aaagatcgag caccgtcacg caatatatcg agaattctgt ctttctcgag aaccgggcgc	11640
caccgtattc ccatagaaaa taaaatacaa cagagagcac ttccacctat gacataggaa	11700
aaaccaaacc ccaacgaagc cagcacaagg tcagaatcct tcgtgacaaa aaaaacgacc	11760
aataggaaat aaaatacctt agacaaaaaa ttggtaagcg ccaaccatcc aaacagcgct	11820
cttccctgaa agaaccatac tgcttgaaga taatttccaa ctactgccgg aagagcagcc	11880
gcaaccaata caagcaaggg gattggtaaa atagaagata cagccaaaat ggccagcact	11940
aaaagtgaag gcagcaacaa taagaatcta gcaactctgt caacagaaaa gaaactagac	12000
agttcaactt tgttatcgat aatggcagcc tttcttgatc ctgccagaat aaatccaaag	12060
tctaccagtt gacatagaat gacagccacg gcctgggcaa tcaccaattg gccaaacgct	12120
tcgcttgaca atgttctcgt gagaaaagga atcgagcaa gaggcaaaag atagttgctg	12180
cccatggata taccggagta gaaaacgccc cgccttattg acattctact cgataccccc	12240
taatacaatt caacaactac aacaagtaag ccctgatgcc agtaagtggc atcagggttt	12300
agatcaaac ttagcgaaga gagccatcgc tacgaagctt ccttataaaa ccagcgagca	12360

ctgcgagtag aattccaatt atcaatcctg ccaaagtacc tatagtaact ataagaatct	12420
tcttcgggctt aatgggttga tttgaaaaag agagtcacct gtcttccttg tagacagcca	12480
ccgcatcaga atccacagac aaactggagt tccaagatag tttctcttgg agagttctca	12540
actcaggaat gaatggagca totacactac gcgactcaag attgttgatt tcagcgcgca	12600
gcgcccttagc tectcgcatg tacatcaagt caccatccat gatogaggag agttgttgct	12660
cggacgcccc ttctattaat ggcgggccat ctatcttgag cgactcgcga atcagcaatg	12720
cctccttcaa acgtgcaatt ctatcatcac ggcggccctt cgccatattc tgcagcacgg	12780
ttatgcggct ctgcattgca gcattttctta cctggaaatc tctacctgca ctatcaataa	12840
cctcatgcac ggcccgatcc gcagccaaac gcacgaaagc ttgtgcccac gtagcaagaa	12900
cctctcgctt cgtgcctctc acaattaccg tataacggct tgcattctggc ttgttagcag	12960
gatcaatctt tacctctttg gagaacttct tataaaactc ctcttgctca tcttcgcttt	13020
ccgctccctc acccacctgg ggaaggtata tcttatagaa gaactctttt ttattctcat	13080
ccgaaagcag attgcgogaa aagatcgcat agatacttct aacagtatat gcatctaggc	13140
cattctccct tctaccaaca ttgaaacctt cgatagacct aagagcagga ggcactactg	13200
caaccctata ttcatataca ggcttactca gatacgcata ggtaaaagac ccgattaatg	13260
caagaagagt agtcagaaga atcagaacct tgttaaccca aagctccttg accagcttca	13320
ccaggtcaac ctccatca gccgtcatca aagaagaatt gcctcaggg	13369

<210> 2

<211> 15244

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400>

2

aattcacatg ttgcccattcc acgaaaccac cttatcgccg tggaacgcac ctggatcgac	60
agccccagca aagcagtcgc ttcttggtcc ggcaccggaa acatcgtagc gagaaaacaa	120
aaaaggccgc taggcggcct ttcccgga acgatgactc agggttctcg ccgcctctgg	180
cgatagatcc agtcgacgat ttcaccgtca ggcgcatagc cgtcgacggg ttcccgacgc	240
aactggcgaa cccgcgagta gtcgtccttc tccacggcgg ccagcaactg ctccagcacg	300
accttgaagg cctcccagct caggtgttcc tcgttggccc gcatgatcat cggatggctg	360
gtgggattca cgttgtcacc gatcagcagc tcttcgtaga gcttctcgcc aggacgcagg	420
ccactgaact cgatggcgat gtcaccatgg ggcgaacgct cggaaacgcac gctcaggccg	480
gacaggtgga tcatcttctc ggcgagctcc aggatcttca ccggcggccc catgtccagc	540
acgaatacat ctccgcctcg ccccatcgaa ccggcctgga tgaccaactg cgcgcctcg	600
ggaatggta tgaagtaacg ggtgatgctc ggggtgggta ccgtcaccgg gccgcgcgc	660
ttgatctgct cgcggaacag cggaaatgacc gaaccggacg aaccgaggac gttgccgaag	720
cggaccatgg tgaacgggt cttgttgacg tgatgcacgt ccttcgggtc gccgaacagc	780
accggcgccg attcgttgct gagcgctga aggaccattt ccgccaggcg cttggtgctg	840
cccatcacat tggtcggccg caccgccttg tcggtggaaa tcagcacgaa gttctgcacg	900
ccgacctgca ccgcggcctg caccgcatgc aagggtgccta tcacgttggt gagaacgccc	960
tcggcgatgt tgtgctcgac gatcggcaca tgcttgtagg ccgccgcatg gtagacggta	1020
ttgaccttc aggtacgcat caggtccacc aggcgctcgg gattgcgcac cgaaccgagg	1080
atcggaaca ggttcaccga aagcgactcg cgttgatcc gacgctccag ttctgatgg	1140
atgctataga ggttgatttc gctgtgttcg aacaggatca gcacgctagg cgaacaactc	1200
atgatctgcc gacagagttc cgaaccgata gagccgcccg ccccggtcac catcaccacc	1260
tgaccgcgga tgcaccgttc cagcagctcc ttgcgcggtg cgaacgtgtc gcgcccagc	1320
aggtcagcga tgtccacctc ctgcaggtca tccacctga cccggccgct ggccaggtcc	1380
atgaagccgg gcatgctgag cacgtgcagc gggaaacggt ccagggactc gagaatctct	1440
cggcgccggg cccgagtggc ggaaggaatc gccaggagaa cctcctgcgc gcccgctcgc	1500
tcgatcatct ggcggatatg cttggcggta tagaccgcga gaccggcaat gaccgggtg	1560
gcgatctgct tgtcgtcatc gatgaacgcc accggacgca tcgcccgacc gagacgcaac	1620
gccgcaacca actgggtgcc ggccgcccc gccccataga taaccacctt gggcaggcca	1680

tcctggcggt tgagaaatgg taccgactgc acagcagagt accagtcgcc catgaaatac	1740
tggcgcatgg ccagacgcaa gccgccgatc agcagcatgc tcaaccacca gtagttgaac	1800
accaggggaa ggggcaccgg cgccggcgcg ccacgatacc agtacaccac cagcgacagc	1860
accagagccg agatgggtcac cgccttggcg atggcgatca atgcgtcggt accgagatag	1920
cgcatacccg cgcgatacat gccgaagcga atgaatagtg gaatggcgat gaccggcgca	1980
gtgatgaaaa gccatgcatg ctgcgcgaac acgtcgatca tatcgtctgt gcttagaagg	2040
accacaaaag cgagccacag agacagccat accagaagga tatccgtagc cacttgaagc	2100
aaacgtttcc agcgacgagg catggataac aactttactc ttaaaccgac tagcattccc	2160
ctactccttt aggcctgagc ggacgactct agcttaccgg ccttgaacct caatgccaga	2220
aaaacgagag ggacgtaggc tatgacgatg cccatcaaag gctctagaga cccactcccg	2280
accaacaagg ctataggagg gagccagatg acgttcaatg ccgtgacgcc aatcgtaaca	2340
ggagcatgct ttccatagta tcggcttgcg tattgatagg catggcttcg atgagcctca	2400
tacaccttgt cccctctcag caagcgacga atcaggggat aggtcgcatc gacgatgaaa	2460
acgcctaaca ggaccaacca tgcccagaaa aaattcgtat tcatccacat ggcataatg	2520
gaaagaattc ctaaaacaat ccccagaaga ccactacccg catctcccat gaaaattttt	2580
gggggtggaa aattccagaa caagaatcca aaaacggcaa aagctagcga taaggggagc	2640
aaagcctgcg tcagttggcc attcagccag tataataatg cccaccaaac acagacaaaa	2700
atggcctgaa ggctagcaag tccatcgatc ccatccatga agttatagag attcagcaac	2760
cacacgagat agaaagcaaa gagaagtctt ccgaaccacc ccaggtccca actgacacca	2820
acaatctgaa atggtggtat gccattcaaa aagtacaatg agaccaaggc tgcactaaaa	2880
tgtcccagca gacgccagcg tgcggcgata tgaccgtgat catccatgaa tccaataatg	2940
gcaaccccg cgcagcgag aagtaaagcc caggacacgg cccatgagat atttccagt	3000
acagcccaaa taggcagcat gagacaaaaa gtaactacaa tggcgacccc tcctccgct	3060
ggagtgggaa cgacatggga acttcgagag ttaggggtgt caagtaagct cctcgctaag	3120
gcataacgac gcaagaggcc tgtaagcagt cccgaaaccc cagctgcagc gagtaacaaa	3180
taccattctt ccatttataa atttatcccc aaaaatactc tcaccccttc aacagggttc	3240
tattcgattg ttctgcgtgg taagccatgc tagtagcaaa gcgcctgcag ataatcattc	3300
agtacgttg acatatattc taccaacatc ttccaacgcg gaaaggtgc tcttgcgggg	3360

ttcccaacca agcaaaagcc tgccttttga cgagtcgacc totagtgagc agcataactg	3420
agtggtataac ccaccctttc ctactaattt aagaaacctc agaataaacc taggaacagg	3480
ccacatgatg gggcgacgcc ccattccgcg agcaagcgca gtcaccagtt gcttggtaga	3540
aatctcctga ccatcggata ccaaaaacag ttgcgcggca gcagaagggt gcgtcataca	3600
gcaagctaga aagtcaacta aattatccag agaaacaaaa cttcgtcggg tatctatgca	3660
accaaagga agaggaagtc ccgaagcaac cagcttcaac aatcgcgaga aatttccagg	3720
agctttccag tcgtaaacga gtggaggcct gacgataaca agttctgttg aggaatgctt	3780
gaaaagctcc cgaagcgcta cttcagcctc aaactttgaa atcgcatatt ctgcatgagg	3840
agccggcttg gagttctcat cgaagggtt ttctttggtt aaagcgccat ttacaccaat	3900
agaactaaca aaaatgaaac gctttacaga tgcttcgata gcctgccgag caagcgccag	3960
agtagcatcg cgattcacct tccgaaaaat atctagtga tcacgctgcc ttccaaagat	4020
atgggctcgt ccagctagat gaactacaca ttcaacaccc cgcagcgag catcaagctt	4080
agtgtctctt ttcagctccg ctcgaaacata ctcaaccccc gtaacgggat tgtacaggga	4140
tcgtacttgt ccgacaacct gaaagggggc ggcagcaagc gacctgcaa gcgcactccc	4200
gacaaaaccg ctagccccgg ttaccagcac cttcatcata ttttcgactc agacagaagc	4260
ggtcgaacac aggcagcgaa tttttccaat gaaatatatt cagaataatt ttgctgaagg	4320
acactccgag ggactccgcc taaactagac aagtcgagtt cacaaatctc atcgatcaat	4380
cttgccaaag caacoggatc attcggcgga cagttccaac cgattccggt ctcatctatt	4440
atacgggaga tttcagcccc tttttccatg acagctagaa tcggtttgtc tgctgccatg	4500
gagaaatatg ccttgctggg aaccccaagc ccgaacattc cttcttctaa ggtaactaag	4560
gcaacgtcac aggcagccaa accaaaattc ttttcggcta atggcagcct tccaaaatac	4620
ctcaaccgag cacactgata ttccagcgcg tgttttttta cactgtcgac caaggcacca	4680
tctccaataa aagcaaaagc cgccttctcg tttttaacca actgaatagc agaaagtatg	4740
ttttctatac cttgtaatcg accgacatta ccaaaaaatt ggaaaaccct tttacctttc	4800
cattcaggaa tattgataaa aggagcatcc tctctcggta ctgggaaaac ctctttctca	4860
caggcccaat tcgaaataaa gaccaaagat cgcgggtcat tcaccttctc tttcataaga	4920
gcttccatat cgcgccctat tacgactaga cgatcagcgg atgagtaaag gaaagaaaag	4980

agacgacgta gaagccggta ggcaatacta tctttcttca gaacgcccgc cggcaccaag	5040
ttctcgggaa acacatcatg caccagcagc acccacttga aaccgagggc ataccttagc	5100
aagggaaacg tcatcagtag aagagcaggg ttggttccac tcaataccac atctcctcgt	5160
ctggcacgag aagtcaattt tactgagaac agaaacgcct gaaaaatctg cgccaatcct	5220
ctagaaagga gcctatttct attgctcctc ggaaacttaa agcatttctg ctcaacggaa	5280
ggtgaaacca caggcgggggt ttcacctgcg gtcagcggga aaattacggt tagccccacca	5340
aactcccgtc gcctctttcc tattatcttc tccaatagt atcccgaggga gttctgattg	5400
gcaccgacat actcagaaac cacaatatc cttgccatca actccacgcc tcaagttaat	5460
attttttcca gacagttcgc atgacatagt ccgatagct atgaactatg cgaacaatct	5520
tttcagaaac gttgggcatg ctatagtcag cgaccaagcg taacatgcgc tctgcgtcgc	5580
gcctctgtcc ctccaacacc tcgagtgtt gtagtactcg atccgaatcc agtcogacca	5640
tcatcaccac agcctcttcc atgcctctg ggcgttcag agcctcacga atattcaaag	5700
cgggaaaatt cagtattgaa gactcctcac tgatgggtccc actgtcagaa ataactgcct	5760
tggctgtaat ttgcagttta ttgtaatcct taaagccgag gggtttcagc agtttaatac	5820
cctcgtgaaa ctttgccctc gtgcctcaa ttctctttt tgttctaggg tgagttgata	5880
cgatgacggg cagcgagtag tttctgcca cagcgttgag catagaaacc aacttcaaga	5940
aattcttctc cgaatctatg ttttctctc ggtgcgcact cagcagaaag aaccgctctg	6000
ttttcaaccg gagcctttca agaatatcgg aggactcgat ccggtcacga tagtgctoga	6060
gaacttcgaa catagggcta ccagtcttga taacctgtc tggagaaagt ctttcacgca	6120
agagataatc acgcgcaatt gtactatagg tcaaatttac atcagctgta tgatcgacaa	6180
tgcgccgatt tatctcttca ggcacacgca tatcgaaaca gcgattgcct gcttccatat	6240
gaaagggtcg tatcttacgc cgttttgtag gcagtaccgc catacaacta ttggtatcac	6300
ccagcacgag cagcgcatcg ggatctattt cgcccagaac acgatcgact gcgattatta	6360
cattccctat cgtttcagcc ccggaagacc cggcggcggt tagaaaataa tccggctttc	6420
ttataccgag gtctgaaaa aatatttcat taagttcgta atcataattc tgtccagtat	6480
ggacaagtac atgatcgtag tactgatcaa gcttcgccat gaccctagac aagcgaataa	6540
tctcaggacg agttccaaca accgtaacga cttttagctt ctgcattgtt atctcactat	6600
accttaogca ccttcgccta ctgaacaagc gtaggtatcc ggattttccc gatcaaatac	6660

ttcgttagcc cacaacatga ctaccatata gtcagttccg acattagtaa tgtcatgagt	6720
ccatccaggt actgttttga caattttctgc cttttcacca ttagtgcaaa tttcgtaaaa	6780
tgccccggtc aggatgtttc taaacttgaa acgtgccatc cctttgataa ccagaaactt	6840
ttcggttttc gagtgatggt aatgccccgc cctggtaaca cctggatgag ccgtaaaaaa	6900
cgagaactgg ccagagtccg cggctttcag catctcgacg aatgtgccac gcggatccga	6960
atgcattggc acgtcgtaac taaaactata ttctggcaag aaacttagat aagtcgagta	7020
caaggcgcgc gtcaatcccg agccaaccct tgcggtagtc agtgactttc gactattacg	7080
aaactcatac aattgtttctg cgagctcacc aacagaaatc tgatactggg gctcgacctg	7140
tagtgaaaact gcattggata gcttcccata catgactttc atgaagggtc gaaccacatc	7200
atctatgtat acaagagtga tctctgccga ggaattgtta atttgaatcg gaatatctcg	7260
aataatatta tgacaaaaag tcgcaaccgc tgaattataa ttcggacgcg accatttacc	7320
gaatacatta ggaaggcgaa atatgtagac aggacaacca atatcctcac ctagcacttg	7380
gagatgctct tctgcggctc gcttgcttaa accgtactca ttatccacct cagcctgaat	7440
ggatgaagta taaagaagtg gtatggctcg tccattggac cttaccgcct cacacagagc	7500
atacgtgagt tcggaattcc cgatcttaaa ctcttctggt tttccggac gattgacccc	7560
ggcaagatga aaaataaaat cgacggaacg aattagctca ggcaaattac caacactact	7620
ctcgcgggtg aatggcacca cctcgatacc accccgctct gcaagatgag cgcacagatt	7680
ccttccaaca aatccattcg cgccagttac aagaactttc atcgtttatt cctctggact	7740
ggcactctcg ccacgctgaa tagcacgaat gaaatccaac ttcagcaaca gctttttcat	7800
tccttcgata tccagacgtt tggattatg agaattatag tcctctgtat gagtaatttt	7860
ttcctcgcct tgctccacaa acttactata gttcagatca cgcaaactcg gggggatacg	7920
ataatagtca cccatgtctt cagcacaggc catttcctct cgactaagaa ggcctcata	7980
aagcttctct ccatgacgcg tacctattac attgatagga taaccattct tgccaagcaa	8040
ttgagtaagc gcatgagcca gcacctcgat gggtgcagcc ggtgctttct gtacaaaaag	8100
atctccattg gtacocatgt cgaaagcata aagcacaagg tctacggcat ccgtaagcgt	8160
catcatgaaa cgtgtcatgt ttggatcagt gattgtgaga ggctggcctg atcgcatttg	8220
ctcgataaag agaggaatga ccgagcccct tgaagccatg acgttaccat aacgggtgcc	8280
acaaattacg gtaggagtgc gttccaggtt tcgagacttg gcgaccatga ccttttccat	8340
catggccttt gaaataccca tggcattgat tgggtaaact gccttatccg tactcagaca	8400

aacgactttt ttgacgccat tctggatagc agattcgagg acattttccg ttccgatgac	8460
attggtottc acagcctcca tcgggtagaa ctccacaagag ggaacctggt tcaatgcagc	8520
cgcatggaaa atgtagtcca caccgcgagt agcattcaga gtgctttgat agtcgcggac	8580
atctccaata taaaacctca acttgggggtg agcatagcac ttacgcatat catcttgctt	8640
cttctcatcc cgactgaata cagctatttc accaatatct gtatccagaa aacgcttcaa	8700
aacggcattt ccaaaggaac cagttccacc ggtaattaac agaacagagt tcttatccat	8760
acaccacctc tttacgtgta ggctggagct gcttcgaagt tctatactt tctagagaat	8820
aggaacttcg gaataggaac ttcatttaaa tggcgcgctt tacgccccgc cctgccactc	8880
atcgcagtac tgttgtattc attaatgcat tgcgcgacatg gaagccatca caaacggcat	8940
gatgaacctg aatcgccagc ggcatcagca ccttgctgcc ttgcgtataa tatttgccca	9000
tgggtgaaaac gggggcgaag aagttgtcca tattggccac gtttaaatca aaactggtga	9060
aactcaccga gggattggct gagacgaaaa acatattctc aataaacctt ttagggaaat	9120
aggccagggt ttcaccgtaa cagccacat cttgcgaata tatgtgtaga aactgccgga	9180
aatcgtcgtg gtattcactc cagagcgatg aaaacgtttc agtttgctca tggaaaacgg	9240
tgtacaagg gtgaacacta tcccatatca ccagctcacc gtctttcatt gccatacgta	9300
attccggatg agcattcatc aggcgggcaa gaatgtgaat aaaggccgga taaaacttgt	9360
gcttattttt ctttacggtc tttaaaagg ccgtaatatc cagctgaacg gtctgggtat	9420
aggtacattg agcaactgac tgaaatgcct caaatgttc tttacgatgc cattgggata	9480
tatcaacggg ggtatatcca gtgatttttt tctccatttt agcttcctta gctcctgaaa	9540
atctcgacaa ctcaaaaaat acgcccggta gtgatcttat ttcattatgg tgaaagttgg	9600
aacctcttac gtgccgatca acgtctcatt ttgcgcaaaa gttggcccag ggcttcccgg	9660
tatcaacagg gacaccagga tttatttatt ctgcgaagtg atcttccgtc acaggtaggc	9720
gcgccgaagt tctatactt tctagagaat aggaacttcg gaataggaac taaggaggat	9780
attcatatgg tgcacggaag tttaaactta tttatcatca tcatctttat aatcaccatg	9840
atgaagccgt ctttttgtaa taaaatagaa caacacaaac caaaattcac ttacaactaa	9900
ataatttgaa ctaaactta ataatgtgat tggatagatt gctataaaca taacgagtaa	9960
atctatagtg tttggatcat agttcctaata cattttataa actagtagca aaatgcaaat	10020
cattatgata aaaaacccta ataagccaaa tgatagaatc aactcaataa tgatgttatg	10080
tgggtatattt ccgattagtt tataatagtt aaatggccca tagcctaata acggactttg	10140

ttgaataaag taaatacctt tttcataaat cggtcctctt ccagaagtac cttctaaatt	10200
aagtgttcca ccttgtagat atgaaaatgt tctagtattc gaaccttttg taaaaagaaa	10260
gtaaatcaat acactagata tgcttaatgc aaaaatatac ataatgcctt ttactgcaat	10320
aggtattcct cttttaaacg taataagtat aaatgcaaat aagccgtaaa gaattaataa	10380
aatagcacct ccacgccttc ctggtataaa cacaataggg atatcaatta ttgtaaatag	10440
aacatatatc cacttatgtt tcaactgaacc tttcataatg aaataaatgc ctaatccggc	10500
agtaaagtct gaaaggtagc aagcgttttg atagttcata agtccaaaat tgatatagct	10560
aggtatctca cctgtaagtt ttggtattaa aattacaaaa ataatgaaa tagaaaatat	10620
gaaaaatact aattttaaaaa atctttctac cgtagccttg tttatatatt taatataaat	10680
accactaatt gccgctggaa ctgccatgt taaaaagaat agaattattt ttttagctag	10740
tttctcttcc ttatctggtg aaaaataata aaaagctaga taaagcaaac atatagcaat	10800
taataatatt aaccctcgcg gaatttcttg cgtgacaatt atcttataaa tagcaaacac	10860
cgtagtaatt aatgctatac caaccatggt tgagtaatac accggctcta tagggaaccc	10920
taatacttct ttagtaaatg tagagattac tataaatatg ttcattgctga taattgcaca	10980
aagtacaaaa aatttcatag ttctctctgc tagaggatcc ccgggtcctt acagatcctc	11040
ttctgagatg agtttttggt caaatttata taattctact aatcgttcac tttctatttg	11100
ccaattcaaa atttttagacg ccttaattgc attttgacgt aaatgattaa acaaatcggtg	11160
attatctctt aatttttctaa ccgccttttc aattttctaac ggcgtaactt cctttaaaac	11220
aatgccaaat ttatatTTTT cattgagata aatatgctct ttgacaggag ataaaattac	11280
tggtaaacca gcatgtatac attcaaaaat tttattagat actgtatatt caaaaattaat	11340
agatacaggt ttctgcaaga taacaccaac attactttct gctaacttat caaccaattc	11400
ttttacttca actgggtttat ccaacctaat attttccgag ttataactaa tcagttcttt	11460
tatcacttct tcatgcgga ccaacctctg aattatgaat gaaggagcat tttgtttaaa	11520
agctgatgaa gcaataataa actcttcata tctctgtcc attacaattt gaccttgata	11580
tacgatttct ttaaagtttt cgatttcttt aaattctctg ctatcattta aaatagggtgc	11640
attcgtaata acattcgctt ccttcttata tcctttagat tgataatatt cttttgctgc	11700
atgacttact gttacgaagg cattaacacg atgttttact atgtgttttt ctatactttc	11760
tacaaacttt gaaataagtg gaactttatt aataaaggca tttttcgcat atatttcatg	11820
cgcatacata acaatattag cttttttata attgcttaaa tagaccatta ataatacgtc	11880

gaaatcattt	gcatgaatca	cgtcagggtt	aaaagcttta	atttctcgga	taacacctgt	11940
tgcaaatctt	atacgcttaa	ttaatttaga	aagaatattt	ttgggatcta	ccttgctacc	12000
taacaaaoga	taattacaat	ctaaattttc	caatcgcta	ttagtagctt	gtgaattatt	12060
cattccaaca	attttataat	cattcgtaac	gcctttaatt	gtttctattt	gtttaagtac	12120
ccttggtgt	tgaacaatat	tactcgatac	aatatTTaaa	attctcattg	tacacctcct	12180
gctcattgta	cacctcctgt	tacgcataat	cgggcacatc	atacggataa	ctagtatcct	12240
ttttatttaa	atattcaacc	gaaaatcctt	ttagtttgtc	aggcgttttc	tcccaccact	12300
tgctttocaa	aagtttttca	attgttttat	tgtcaaatcg	cttcttaatc	acttttgtag	12360
gaacccccacc	aacaacctca	tatgctccta	catttttagt	aacaactgag	cgggctgcta	12420
tgactgcacc	agtattttatt	gttaatccat	ccataataat	tacatttgca	ccaatccaca	12480
catcattttt	aattgttgta	cggcttggtt	ggtcattaaa	gtctataaac	ttttgcttta	12540
tgtaaatgg	attattatta	gaataaaaaa	tgggtgatga	gctaaaaaag	tgtgtaggat	12600
gttttcttaa	cccaattttt	acatccgaag	atatcgaaac	atatcttcct	acttctacat	12660
tattaaaatc	actaccaaat	ccaatataac	tgtattcacc	aatgtgagaa	ttcctgattt	12720
tacaccatct	atctatatag	ttattgccat	caaattttga	gtttgtaata	tacgccaaagc	12780
gatgaatctt	aacattcgat	tctttagagg	actggttttt	cagcaaacca	attatctttt	12840
caatcgctat	cctcatctta	agatcaaagt	ccccctttct	ctctcaactt	aacogtggcc	12900
aggaaaaaca	ttcccatccc	caaagagacc	tctgtaacga	ccagcgctca	cgctccaaaa	12960
agctcatccc	cccaaagcgc	caagaggata	aatgttaata	ccccgcccaa	ttccgcacaaa	13020
aaaattgctc	gcaaatatac	accatcatac	ccagaaggaa	caagagtcaa	ccctccatac	13080
agaacaccaa	tacaagcaaa	cactggcact	atcgaaaaca	ttcttattac	aacagttaga	13140
cctctggact	gttcggggaa	taggagatat	gaaatgtatt	cagaaaataa	gaaaagaaat	13200
aaacaactag	gaatagctat	tagtaacatc	agacctagcg	ctttccttct	taaactgccg	13260
actgccgggt	cggcactcgc	atacatcctg	ctaaaagtcg	gaaatagagc	gctagctatg	13320
ggtgaagtag	cagccgcgat	cccgcgaagg	aatttatccg	cagtagaaag	cactccggcc	13380
gcgagggtc	caccaacaac	accaaccgcc	gcaacgagca	cttgcatgtg	caagctaaga	13440
aaagccagag	aaagaaaaga	tcgagcacgg	tcacgcaata	tatcgagaat	tctgtctttc	13500
tcgagaaccg	ggcgccaccg	tattcccata	gaaaataaaa	tacaacagag	agcacttcca	13560
cctatgacat	aggaaaaacc	aaacccaac	gaagccagca	caaggtcaga	atccttcgtg	13620
acaaaaaaaa	cgaccaatag	gaaataaaat	accttagaca	aaaaattggg	aagcgccaac	13680
catccaaaca	gcgctcttcc	ctgaaagaac	catactgctt	gaagataatt	tccaactact	13740

gccggaagag	cagccgcaac	caatacaagc	aaggggattg	gtaaaataga	agatacagcc	13800
aaaatggcca	gcactaaaag	tgaaagcagc	aacaataaga	atctagcact	ctgtacaaca	13860
gaaaagaaac	tagacagttc	aactttgtta	tcgataatgg	cagcctttct	tgatcctgcc	13920
agaataaatc	caaagtctac	cagttgacat	agaatgacag	ccacggcctg	ggcaatcacc	13980
aattggccaa	acgcttcgct	tgacaatgtt	ctcgtgagaa	aaggaatcgc	agcaagagggc	14040
aaaagatagt	tgctgcccac	ggatataccg	gagtagaaaa	cgccccgcct	tattgacatt	14100
ctactcgata	ccccctaata	caattcaaca	actacaacaa	gtaagccctg	atgccagtaa	14160
gtggcatcag	ggtttagatc	aaaacttagc	gaagagagcc	atcgctacga	agcttcctta	14220
taaaaccagc	gagcactgcg	agtagaattc	caattatcaa	tcctgccaaa	gtacctatag	14280
taactataag	aatcttcttc	ggcttaatgg	gttgatttga	aaaagagagt	ccctcgtctt	14340
ccttgtagac	agccaccgca	tcagaatcca	cagacaaact	ggagttccaa	gatagtttct	14400
cttgagaggt	tctcaactca	ggaatgaatg	gagcatctac	actacgcgac	tcaagattgt	14460
tgatttcagc	gcgagcgcc	ttagctcttc	gcatgtacat	caagtcacca	tccatgatcg	14520
aggagagttg	ttgctcggac	gccccttcta	ttaatggcgg	gccatctatc	ttgagcgact	14580
cgcgaatcag	caatgcctcc	ttcaaacgtg	caattctatc	atcacggcgg	cccttcgcca	14640
tattctgcag	cacggttatg	cggtctcgca	ttgcagcatt	tcttacctgg	aaatctctac	14700
ctgcactatc	aataacctca	tgcaaggccc	gatccgcagc	caaacgcacg	aaagcttggtg	14760
cccatgtagc	aagaacctct	cgcttcgtgc	cctccacaat	taccgtataa	cggtctgcat	14820
ctggcttggt	agcaggatca	atctttacct	ctttggagaa	cttcttataa	aactcctcct	14880
gctcatcttc	gctttccgct	ccctcaccca	cctggggaag	gtatatctta	tagaagaact	14940
cttttttatt	ctcatccgaa	agcagattgc	gcgaaaagat	cgcatagata	cttctaacag	15000
tatatgcac	taggccattc	tcctttctac	caacattgaa	accttcgata	gacccaagag	15060
caggaggcac	tactgcaacc	ctatatctcat	atacaggctt	actcagatac	gcataggtaa	15120
aagacccgat	taatgcaaga	agagtagtca	gaagaatcag	aaccttggtta	acccaaagct	15180
ccttgaccag	cttcaccagg	tcaacctcac	catcagccgt	catcaaagaa	gaattgcctc	15240
aggg						15244

<210> 3

<211> 16483

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 3

tcacatgttg cccatccacg aaaccacctt atcgccgtgg aacgcacctg gatcgacagc	60
cccagcaaag cagtcgcttc ctggtccggc accggaaaca tcgtacggag aaaacaaaaa	120
aggccgctag gcggcctttt ccggagaacg atgactcagg gttctcgccg cctctggcga	180
tagatccagt cgacgatttc accgtcaggc gcatagccgc tgacggtttc ccgcagcaac	240
tggcgaaccc gcgagtagtc gtctttctcc acggcggcca gcaactgctc cagcacgacc	300
ttgaaggcct cccagctcag gtgttcctcg ttggcccgca tgatcatcgg atggtcggtg	360
ggattcacgt tgtcaccgat cagcagctct tcgtagagct tctcgccagg acgcaggcca	420
ctgaactcga tggcgaatgc accatggggc gaacgctcgg aacgcacgct caggccggac	480
aggtggatca tcttctcggc gagctccagg atcttcaccc gcggcccat gtccagcacg	540
aatacatctc cgccctgcc catcgaaccg gcctggatga ccaactgcgc cgcctcggga	600
atggtcatga agtaacgggt gatgctcggg tgggtgaccg tcaccgggdc gccgcgcttg	660
atctgctcgc ggaacagcgg aatgaccgaa ccggacgaac cgaggacgtt gccgaagcgg	720
accatggtga aacgggtctt gttgacgtga tgcacgtcct tccggtcgcc gaacagcacc	780
ggcgcgatt cgttgctgag cgcctgaagg accatctccg ccaggcgctt ggtgctgccc	840
atcacattgg tcggccgcac cgccttgctg gtggaaatca gcacgaagtt ctgcacgccg	900
acctgcaccg cggcctgcac cgcattgcaag gtgcctatca cgttggtgag aacgcctcgc	960
gcgatgttgt gctcgacgat cggcacatgc ttgtaggccg ccgcatggta gacggtattg	1020
accttccagg tacgcatcac gtccaccagg cgctcgggat tgcgcaccga accgaggatc	1080

ggcaacaggt tcaccgaaag cgactcgcgc ttgatccgac gctccagttc ctgatggatg	1140
ctatagaggt tgtattcgct gtgttcgaac aggatcagca cgctaggcga acaactcatg	1200
atctgccgac agagttccga accgatagag ccgcccgcgc cggtcacat caccacctga	1260
ccgcggatgc accgttccag cagctccttg cgcggtgcga cgctgtcgcg cccagcagg	1320
tcagcgatgt ccacctcctg caggtcaccc accttgaccc ggccgctggc caggtcacatg	1380
aagccgggca tgctgcgcac gtgcagcggg aacggctcca gggactcgag aatctctcgg	1440
cgccggggcc gagtggcgga aggaatgcc aggagaacct cctgcgcgcc cgtctcgtcg	1500
atcatctggc ggatatgctt ggcggtatag acccgcagac cggcaatgac ccggttggcg	1560
atctgcttgt cgtcatcgat gaacgccacc ggacgcacgc cccgaccgag acgcaacgcc	1620
gcaaccaact ggttgccggc cgcggccgcc ccatagataa ccacctggg caggccatcc	1680
tggcggttga gaaatggtac cgactgcaca gcagagtacc agtcgcccac gaaatactgg	1740
cgcatggcca gacgcaagcc gccgatcagc agcatgctca accaccagta gttgaacacc	1800
agggaaacgc gcaccggcgc cggcgcgcca cgataccagt acaccaccag cgacagcacc	1860
agagccgaga tggtcacccg cttggcgatg gcgatcaatg cgtcgttacc gagatagcgc	1920
atcaccgcgc gatacatgcc gaagcgaatg aatagtggaa tggcgatgac cggcgcagtg	1980
atgaaaagcc atgcatgctc gccgaacacg tcgatcatat cgtctgtgcc tagacggacc	2040
acaaaagcga gccacagaga cagccatacc agaaggatat ccgtagccac ttgaagcaaa	2100
cgtttccagc gacgaggcat ggataacaac tttactctta aacgatctag cattccccta	2160
ctcctttagg cctgagcggc cgactctagc ttaccgcct tgaacctcac tgccagaaaa	2220
acgagaggga cgtaggctat gacgatgcc atcaaaggct ctagagaccc actcccagcc	2280
aacaaggcta tagggaggag ccagatgacg ttcaatgccg tgacgccaat cgtaacagga	2340
gcatgctttc catagtatcg gcttgctat tgataggcat ggcttogatg agcctcatac	2400
accttgctcc ctctcagcaa gcgacgaatc agggatatagg tcgcatcgac gatgaaaacg	2460
cctaacagga ccaaccatgc ccagaaaaaa ttcgtattca tccacatggc atgaatggaa	2520
agaattccta aaacaatccc cagaagacca ctaccgcct ctcccatgaa aatttttggg	2580
ggtggaaaat tccagaacaa gaatccaaa acggcaaaag ctagcgataa ggggagcaaa	2640
gcctgcgtca gttggccatt cagccagtat aataatgcc caccaacaca gacaaaaatg	2700
gcctgaaggc tagcaagtcc atcgatccca tccatgaagt tatagagatt cagcaaccac	2760
acgagataga aagcaaagag aagtcctccg aaccacccca ggtcccaact gacaccaaca	2820

atctgaaatg gtggtatgcc attcaaaaag tacaatgaga ccaaggctgc actaaaatgt	2880
cccagcagac gccagcgtgc ggcgatatga ccgtgatcat ccataaatcc aataatggca	2940
accccgccgc cagcgagaag taaagcccag gacacggccc atgagatatt tcccagtaca	3000
gcccaaatag gcagcatgag acaaaaagta actacaatgg cgacccctcc tccgcgtgga	3060
gtgggaacga catgggaact tcgagagtta ggggtgtcaa gtaagctcct cgctaaggca	3120
taacgacgca agaggcctgt aagcagtccc gaaaccccag ctgcagcgag taacaaatac	3180
cattcttcca ttataaatt tatccccaaa aatactctca ccccttcaac agggttctat	3240
tcgattgttc tgcgtggtac gccatgctag tagcaaagcg cctgcagata atcattcagt	3300
acgtttgaca tatattctac caacatcttc caacgcggaa aggggtgctct tgcgggggttc	3360
ccaaccaagc aaaagcctgc ctttcgacga gtgcacctct agtgagcagc ataactgagt	3420
gtataaccca cccctttcta ctaatttaag aaacctcaga ataaacctag gaacaggcca	3480
catgatgggg cgacgcccc a tcccgacgc aagcgcagtc accagttgct tggtagaaat	3540
ctcctgacca tcggatacca aaaacagttc gccggcagca gaagggtgcg tcatacagca	3600
agctagaaag tcaactaaat tatccagaga acaaaaactt cgtcggttat ctatgcaacc	3660
aaatggaaga ggaagtcccg aagcaaccag cttcaacaat cgcgagaaat ttccaggagc	3720
tttccagtcg taaacgagtg gaggcctgac gataacaagt tctgttgagg aatgcttgaa	3780
aagctcccgga agcgtactt cagcctcaaa ctttgaaatc gcatattctg catgaggagc	3840
cggcttgag tctcatoga agggcttttc tttggttaaa ggcgcattta caccaataga	3900
actaacaaaa atgaaacgct ttacagatgc ttcgatcgcc tgccgagcaa ggcgcagagt	3960
agcatcgga ttcaccttc gaaaaatac tagtgaatca cgctgccttc caaagatatg	4020
ggctcgtcca gctagatgaa ctacacattc aacacccgc agcgcagcat caagcttagt	4080
gctctctttc agctccgctc gaacatactc aaccccgta acgggattgt acagggatcg	4140
tacttgtccg acaacctgaa agggggcggc agcaagcgac ctgcaaagcg cactcccgac	4200
aaaaccgcta gcccgggtta ccagcacctt catcatattt tgactcaga cagaagcggc	4260
cgaacacagg cagcgaatth ttccaatgaa atatattcag aataattttg ctgaaggaca	4320
ctccgcggga ctccgcctaa actagacaag tcgagttcac aaatctcatc gatcaatctt	4380
gccaaagcaa ccggatcatt cggcggacag ttccaaccga ttccggcttc atctattata	4440
cgggagattt cagcccttt ttccatgaca gctagaatcg gtttgtctgc tgccatggag	4500

aaatatgcct	tgctgggaac	cccaagcccg	aacattcctt	cttctaaggt	aactaaggca	4560
acgtcacagg	cagccaaacc	aaaattcttt	toggctaattg	gcagccttcc	aaaatacctc	4620
aaccgagcac	actgatcttc	cagcgcggtgt	ttttttacac	tgtcgaccaa	ggcaccatct	4680
ccaataaaaag	caaaagccgc	cttctcgttt	ttaaccaact	gaatagcaga	aagtatgttt	4740
tctatacctt	gtaatcgacc	gacattacca	aaaaattgga	aaaccctttt	acctttccat	4800
tcaggaatat	tgataaaagg	agcatcctct	ctcggtactg	ggaaaacctc	tttctcacag	4860
gcccattcg	aaataaagac	caaagatcgc	gggtcattca	ccttctcttt	cataagagct	4920
tccatatcgc	gccctattac	gactagacga	tcagcggatg	agtaaagaa	agaaaagaga	4980
cgacgtagaa	gccggtaggc	aatactatct	ttcttcagaa	cgcccgccgg	caccaagttc	5040
tcgggaaaca	catcatgcac	cagcagcacc	cacttgaaac	cgagggcata	ccttagcaag	5100
ggaaacgtca	tcagtagaag	agcagggttg	gttccactca	ataccacatc	tcctcgtctg	5160
gcacgagaag	tcaattttac	tgagaacaga	aacgcctgaa	aaatctgcgc	caatcctcta	5220
gaaaggagcc	tattcttatt	gctcctcgga	aacttaaagc	attcttgctc	aacggaaggt	5280
gaaaccacag	gcgggggttc	acctgcggtc	agcgggaaaa	ttacgggttag	cccaccaaac	5340
tcccgtgca	tctttcctat	tatcttctcc	caatagtatc	ccgtggagtt	ctgattggca	5400
ccgacatact	cagaaaccac	aaatatcctt	gccatcaact	ccacgcctca	agttaatatt	5460
ttttccagac	agttcgcatg	acatagtccc	gatagctatg	aactatgcga	acaatctttt	5520
cagaaacggt	gggcatgcta	tagtcagcga	ccaagcgtaa	catgcgctct	gcgtcgcgcc	5580
tctgtccctc	caacacctcg	agtgtttgta	gtactcgatc	cgaatccagt	ccgaccatca	5640
tcaccacagc	ctcttccatg	ccttctgggc	gttcatgagc	ctcacgaata	ttcaaagcgg	5700
gaaaattcag	tattgaagac	tcctcactga	tgggtccact	gtcagaaata	actgccttgg	5760
ctgtaatttg	cagttttattg	taatccttaa	agccgagggg	tttcagcagt	ttaataccct	5820
cgtgaaactt	tgcttcgctc	gcctcaattc	tcttttttgt	tctagggtga	gttgatacga	5880
tgacgggcag	cgagtacttt	tctgccacag	cgttgagcat	agaaaccaac	ttcaagaaat	5940
tcttatccga	atctatgttt	tcctctcggt	gcgcactcac	gacaaagaac	cgctctgttt	6000
tcaacccgag	cctttcaaga	atatcggagg	actcgatccc	gtcacgatag	tgctcgagaa	6060
cttcgaacat	agggctacca	gtcttgataa	ccatgtctgg	agaaagtcct	tcacgcaaga	6120
gataatcacg	cgcaattgta	ctataggtoa	aatttacatc	agctgtatga	tcgacaatgc	6180
gccgatttat	ctcttcaggc	acacgcatac	cgaaacagcg	attgcctgct	tcctatgaa	6240

aggtcggtat	cttacgcggt	tttgcaggca	gtaccgccat	acaactattg	gtatcaccca	6300
gcacgagcag	cgcacgcgga	tctatttcgc	ccagaacacg	atcgactgcg	attattacat	6360
tccctatcgt	ttcagccccg	gaagaccggg	cggcgtttag	aaaataatcc	ggctttctta	6420
taccgaggtc	ctgaaaaaat	atttcattaa	gttcgtaatc	ataattctgt	ccagtatgga	6480
caagtacatg	atcgcagtac	tgatcaagct	tcgccatgac	cctagacaag	cgaataatct	6540
caggacgagt	tccaacaacc	gtaacgactt	ttagcttctg	cattgttatc	tcactatacc	6600
ttacgcacct	tcgcctactg	aacaagcgta	ggtatccgga	ttttcccgat	caaatacttc	6660
gtagccccc	aacatgacta	ccatatcgtc	agttccgaca	ttagtaatgt	catgagtcca	6720
tccaggtact	gtttcgacaa	tttctgcctt	ttcaccatta	gtgcaaattt	cgtaaaatgc	6780
cccggtcagg	atgtttctaa	acttgaaacg	tgccatccct	ttgataacca	gaaacttttc	6840
ggttttcgag	tgatggtaat	gcccgcgccct	ggtaacacct	ggatgagccg	taaaaaacga	6900
gaactggcca	gagtcgcggg	tcttcagcat	ctcgacgaat	gtgccacgcg	gatccgaatg	6960
cattggcacg	tcgtaactaa	aactatcttc	tggcaagaaa	cctagataag	tcgagtacaa	7020
ggcgcgcgtc	aatccccgagc	caacccttgc	ggtagtcagt	gactttcgac	tattacgaaa	7080
ctcatacaat	tgttctgcga	gtcaccaac	agaaatctga	tactggggct	cgacctgtag	7140
tgaaactgca	ttggatagct	tcccatccat	gactttcatg	aagggtgcgaa	ccacatcatc	7200
tatgtatata	agagtgatct	ctgccgagga	attgttaatt	tgaatoggaa	tatctcgaat	7260
aatattatga	caaaaagtgc	caaccgctga	attataattc	ggaocgcgac	atttacgaa	7320
tacattagga	aggcgaaata	tgtagacagg	acaaccaata	tcctcaccta	gcacttgag	7380
atgctcttct	gcggctcgct	tgcttaaacc	gtactcatta	tcacacctag	cctgaatgga	7440
tgaagtataa	agaagtggta	tggctcgtcc	attggacctt	accgootcac	acagagcata	7500
cgtgagttcg	gaattcccgga	tcttaaactc	ttctggtttt	tccggacgat	tgaccccggc	7560
aagatgaaaa	ataaaatcga	cggaacgaat	tagctcaggc	aaattaccaa	cactactctc	7620
gcgggtgaat	ggcaccacct	cgataccacc	ccgctctgca	agatgagcgc	acagattcct	7680
tccaacaaat	ccattcgcgc	cagttacaag	aactttcatc	gtttattcct	ctggactggc	7740
actctcgcca	cgctgaatag	cacgaatgaa	atccaacttc	agcaacagct	ttttcattcc	7800
ttcgatatcc	agacgtttgg	tattatgaga	attatagtcc	tctgtatgag	taattttttc	7860
ctcgccttgc	tccacaaact	tactatagtt	cagatcacgc	aaatctgggg	ggatacgata	7920
atagtcaccc	atgtcttcag	cacaggccat	ttcctctcga	ctaagaagcg	cctcataaag	7980

cttctctcca tgacgcgtac ctattacatt gataggataa ccattcttgc caagcaattg	8040
agtaagcgca tgagccagca cctcgatggg tgcagccggg gctttctgta caaaaagatc	8100
tccattggta ccatgctcga aagcataaag cacaagggtct acggcatccg taagcgatcat	8160
catgaaacgt gtcattgttg gatcagtgat tgtgagaggc tggcctgac gcatttgctc	8220
gataaagaga ggaatgaccg agccccttga agccatgacg ttaccataac ggggtgccaca	8280
aattacggta ggagtgcgtt ccagggttctg agacttggcg accatgacct tttccatcat	8340
ggcctttgaa ataccatgg cattgattgg gtaaactgcc ttatccgtac tcagacaaac	8400
gacttttttg acgccattct ggatagcaga ttccaggaca ttttccgttc cgatgacatt	8460
ggtcttcaca gcctccatcg ggtagaactc acaagaggga acctgtttca atgcagccgc	8520
atggaaaatg tagtccacac cgcgagtagc attcagagtg ctttgatagt cgcggacatc	8580
tccaatataa aacctcaact tggggtgagc atagcactta cgcataatcat cttgcttctt	8640
ctcatcccga ctgaatacac gtatttcacc aatatctgta tccagaaaac gcttcaaaac	8700
ggcatttcca aaggaaccag ttccaccggg aattaacaga acagagttct tatccatata	8760
ccacctcttt acgtgtaggc tggagctgct togaagttcc tatactttct agagaatagg	8820
aacttcggaa taggaacttc atttaaatgg cgcgccttac gccccgcct gccactcatc	8880
gcagtactgt tgtattcatt aagcatctgc cgacatggaa gccatcacia acggcatgat	8940
gaacctgaat cgccagcggc atcagcacct tgtcgccttg cgtataatat ttgcccatgg	9000
tgaaaacggg ggcgaagaag ttgtccatat tggccacggt taaatcaaaa ctggtgaaac	9060
tcacccaggg attggctgag acgaaaaaca tattctcaat aaacccttta gggaaatagg	9120
ccagggtttc accgtaacac gccacatctt gogaatatat gtgtagaaac tgccggaaat	9180
cgtcgtggta ttactccag agcgatgaaa acgtttcagt ttgctcatgg aaaacggtgt	9240
aacaaggggtg aacactatcc catatcacca gctcaccgtc tttcattgcc atacgtaatt	9300
ccggatgagc attcatcagg cgggcaagaa tgtgaataaa ggccggataa aacttgtgct	9360
tatttttctt tacggctctt aaaaaggccg taatatccag ctgaacggtc tggttatagg	9420
tacattgagc aactgactga aatgcctcaa aatgttcttt acgatgccat tgggatatat	9480
caacgggtgg atatccagtg attttttctt ccatttttagc ttcccttagct cctgaaaatc	9540
togacaactc aaaaaatagc cccggtagtg atcttatttc attatgggtga aagttggaac	9600
ctcttacgtg ccgatcaacg tctcattttc gccaaaagtt ggcccagggc ttcccgggat	9660
caacagggac accaggattt atttattctg cgaagtgatc ttccgtcaca ggtaggcgcg	9720

cogaagttcc tatactttct agagaatagg aacttcggaa taggaactaa ggaggatatt	9780
catatgggtgc acttacgcat aatccggcac atcatacgga taactatcga cgtccctttt	9840
attaatgaat ttagacttta aacctataat atttaaaatc acataactta taatgttgct	9900
gaacagcgtt gtccaaaaga atccatttag accaaacgca attgtcatta aaatagttat	9960
gaatataaat gtaatcgtgt gaagcgtcat ataattcgtt tgtaatgtta ttgatgcgtg	10020
ttttgtattt aaagtttggg tcataatact cactgcattg aataaaacac ctatattacc	10080
taaaaataata aacttcgaat ataaacttga atcaatgtta tacagtaaag agacaataat	10140
taatgttaatt ggataacata taatcatgac taaacatagt gcagctatag caaatagatt	10200
agttttcaaa tattgcttct ttatatgttc gctttcattt acagaaatat atgaaagtac	10260
tacattatta atcggataga gaaatgtagc taacattttc ccaataaatg ttgaaagaaa	10320
tgatatagtt acagctgttc cacctataat tggtaataag attaatctat ctaagtagag	10380
attcaaatta ttaaggctat ttgtactcag tagcatcaca taatctttga ccacattatt	10440
atcttcactt tgatactgc ctatagttaa tcccgtaat ttaaccaatg tatatatcgt	10500
tgcaaacaat tcactggtaa taaaacaaac aatccagttt tggattaaat aatatagaaa	10560
tagtcctatc agcaaacccta aaaattgaat aagagcaata tacaaaatct gattatattt	10620
taaagtcatc ctaaaaaata cattcagata aatccttaaa cacattaaaa tattaagtag	10680
aattaaaaag ataatatcga tgggtgttcaa attaaaaaag taaagaaata caataattaa	10740
agctatactc tcaatcagaa ttgaaattaa aagtatcgac acaaatttcc agtagtaatg	10800
attggatttg tatagattca tattaattaa tcgtatattg ttaagcgtat tgcctaatac	10860
aacactcgtg attgttatta ttgtataaat cgttaaaata gaaccaaacg cttcattacc	10920
tactcgttga ttaataattg gataagctaa aaattgtaat cctaaagcta taagcaacgt	10980
accaattatt gtctttacgc tatccataat aaaaactttc ttaaccatga tgcctcctgt	11040
ttaaacttat ttatcatcat catctttata atcaccatga tgacgccgtc cttttgtaat	11100
aaaatagaac aacacaaacc aaaattcact tacaactaaa taatttgaac taaacattaa	11160
taatgtgatt ggatagattg ctataaacat aacgagtaaa tctatagtgt ttggatcata	11220
gttcctaatac attttataaa ctagtagcaa aatgcaaatac attatgataa aaaaccctaa	11280
taagccaaat gatagaatca actcaataat gatgttatgt ggtatatttc cgattagttt	11340

ataatagtta aatggcccat agcctaataa cggactttgt tgaataaagt aaataccttt	11400
ttcataaatc ggtcctcttc cagaagtacc ttctaaatta agtggtccac cttgtagata	11460
tgaaaatggt ctagtattcg aaccttttgt aaaaagaaag taaatcaata cactagatat	11520
gcttaatgca aaaatatata taatgctttt tactgcaata ggtattcctc ttttaaactg	11580
aataagtata aatgcaaata agccgtaaag aattaataaa atagcacctc cacgccctcc	11640
tggtataaac acaataggga tatcaattat tgtaaataga acatatatcc acttatgttt	11700
cactgaacct ttcataatga aataaatgcc taatccggca gtaaagtctg aaaggtagca	11760
agcgttttga tagttcataa gtccaaaatt gatatagcta ggtatctcac ctgtaagttt	11820
tggtattaaa attacaaaaa taaatgaaat agaaaatatg aaaaatacta atttaaaaaa	11880
tctttctacc gtagccttgt ttatatattt aatataaata ccactaattg ccgctggaac	11940
tgcccatggt aaaaagaata gaattattt tttagctagt ttctcttcct tatctggtga	12000
aaaataataa aaagctagat aaagcaaaca tatagcaatt aataatatta accctcgcgg	12060
aatttcttgc gtgacaatta tcttataaat agcaaacacc gtagtaatta atgctatacc	12120
aaccatggtt gagtaatata ccggctctat agggaaccct aatacttctt tagtaaatgt	12180
agagattact ataaatatgt tcatgctgat aattgcacaa agtacaaaaa atttcatagt	12240
tcctcctgct agaggatccc cgggtcctta cagatcctct tctgagatga gttttgttc	12300
aaatttatat aattctacta atcgttcact ttctatttgc caattcaaaa ttttagacgc	12360
cttaattgca ttttgacgta aatgattaaa caaatcgtga ttatctctta attttctaac	12420
cgccttttca atttctaacg gcgtaacttc ctttaaaaca atgccaaatt tatatttttc	12480
attgagataa atatgctctt tgacaggaga taaaattact ggtaaaccag catgtatata	12540
ttcaaaaatt ttattagata ctgtatattc aaaattaata gatacagggt tcgtcaagat	12600
aacaccaaca ttactttctg ctaacttatc aaccaattct tttacttcaa ctggtttatc	12660
caacctataa ttttcogagt tataactaat cagttctttt atcacttctt catgcggacc	12720
aaaccctoga attatgaatg aaggagcatt ttgtttaaaa gctgatgaag caataataaa	12780
ctcttcatat cctctgtcca ttacaatttg accttgatat acgatttctt taaagttttc	12840
gatttcttta aattctctgc tatcatttaa aatagggtgca ttcgtaataa cattcgcttc	12900
cttcttatat ccttttagatt gataatattc ttttgctgca tgacttactg ttacgaaggc	12960
attaacacga tgttttacta tgtgtttttc tatactttct acaaactttg aaataagtgg	13020
aactttatta ataaaggcat ttttcgcata tatttcatgc gcatcataaa caatattagc	13080
ttttttataa ttgcttaaat agaccattaa taatacgtcg aaatcatttg catgaatcac	13140
gtcagggttta aaagctttta tttctcggat aacacctggt gcaaatctta tacgcttaat	13200

taatttagaa agaataat	ttt tgggatctac	cttgctacct	aacaaacgat	aattacaatc	13260	
taaat	tttcc aatcgocctat	tagtagcttg	tgaattattc	attccaacaa	ttttataatc 13320	
attcgtaacg	cctttaattg	tttctatttg	tttaagtacc	cttgggtctt	gaacaatatt 13380	
actcgataca	atatttaaaa	ttctcattgt	acacctcctg	ctcattgtac	acctcctggt 13440	
acgcataatc	cggcacatca	tacggataac	tagtatcctt	tttatttaaa	tattcaaccg 13500	
aaaatccttt	tagtttgtca	ggcgttttct	cccaccactt	gctttccaaa	agtttttcaa 13560	
ttgttttatt	gtcaaatogc	ttcttaatca	cttttgcagg	aacccaccca	acaacctcat 13620	
atgctcctac	at	tttttagta	acaactgagc	cggtctgtat	gactgcacca	gtatttattg 13680
ttaatccatc	cataataatt	acatttgac	caatccacac	atcattttta	attgttgtag 13740	
ggcttggttg	gtcattaaag	tctataaact	tttgctttat	gttaaattgga	ttattattag 13800	
aataaaaaat	cggtgatgag	ctaaaaaagt	gtgtaggatg	ttttcctaac	ccaattttta 13860	
catocgaaga	tatogaacaa	tatcttccta	cttctacatt	attaaaaatca	ctaccaaaatc 13920	
caatataact	gtattcacca	atgtgagaat	tcttgatttt	acaccatcta	tctatatagt 13980	
tattgccatc	aaat	tttgag	tttgtaatat	acgccaaagc	atgaatctta	acattcgatt 14040
cttttagagga	ctgg	tttttc	agcaaacc	aa	aatcgctatc	ctcatcttaa 14100
gatcaaagtc	cccttttctc	tctcaactta	accgtggcca	ggaaaaacat	tcccatcccc 14160	
aaagagacct	ctgtaacgac	cagcgctccac	gctccaaaaa	gctcatcccc	ccaaagcgcc 14220	
aagaggataa	atgttaatac	cccgcccaat	tccgcaaaaa	aaattgctcg	caaataataca 14280	
ccatcatacc	cagaaggaac	aagagtcaac	cctccataca	gaacaccaat	acaagcaaac 14340	
actggcacta	togaaaacat	tcttattaca	acagttagac	ctctggactg	ttccgggaat 14400	
aggagatatg	aatgtattc	agaaaataag	aaaagaaata	aacaactagg	aatagctatt 14460	
agtaacatca	gacctagcgc	tttcttctt	aaactgccga	ctgccggggtc	ggcactcgca 14520	
tacatcctgc	taaaagtcgg	aatagagcg	ctagctatgg	gtgaagtagc	agccgcgac 14580	
ccggaagga	atttatccgc	agtagaaagc	actccggccg	cggaggctcc	accaacaaca 14640	
ccaaccgccg	caacgagcac	ttgcatgtgc	aagctaagaa	aagccagaga	aagaaaagat 14700	
cgagcaccgt	cacgcaatat	atcgagaatt	ctgtctttct	cgagaaccgg	gcgccaccgt 14760	
attcccatag	aaaataaaat	acaacagaga	gcacttccac	ctatgacata	ggaaaaacca 14820	
aaccccaacg	aagccagcac	aaggtcagaa	tccttcgtga	caaaaaaac	gaccaatagg 14880	
aaataaaata	ccttagacaa	aaaattggt	agcgccaacc	atccaaacag	cgctcttccc 14940	
tgaagaacc	atactgcttg	aagataat	ccaactactg	ccggaagagc	agccgcaacc 15000	
aatacaagca	aggggattgg	taaaatagaa	gatacagcca	aaatggccag	cactaaaagt 15060	
gaaagcagca	acaataagaa	tctagcactc	tgtacaacag	aaaagaaact	agacagttca 15120	

actttgttat	cgataatggc	agcctttctt	gatcctgcc	gaataaatcc	aaagtctacc	15180
agttgacata	gaatgacagc	cacggcctgg	gcaatcacca	attggccaaa	cgcttcgctt	15240
gacaatgttc	tcgtgagaaa	aggaatcgca	gcaagaggca	aaagatagtt	gctgcccattg	15300
gatataccgg	agtagaaaac	gccccgcctt	attgacattc	tactcgatac	cccctaatac	15360
aattcaacaa	ctacaacaag	taagccctga	tgccagtaag	tggcatcagg	gtttagatca	15420
aaacttagcg	aagagagcca	tcgctacgaa	gcttccttat	aaaaccagcg	agcactgcga	15480
gtagaattcc	aattatcaat	cctgccaaaag	tacctatagt	aactataaga	atcttccttcg	15540
gcttaatggg	ttgatttgaa	aaagagagtc	cctcgtcttc	cttgtagaca	gccaccgcatt	15600
cagaatccac	agacaaaactg	gagttccaag	atagtttctc	ttggagagtt	ctcaactcag	15660
gaatgaatgg	agcatctaca	ctacgcgact	caagattggt	gatttcagcg	cgcagcgctt	15720
tagctcctcg	catgtacatc	aagtcaccat	ccatgatcga	ggagagttgt	tgctcggacg	15780
ccccttctat	taatggcggg	ccatctatct	tgagcgactc	cgcaatcagc	aatgcctcct	15840
tcaaacgtgc	aattctatca	tcacggcggc	ccttcgccat	attctgcagc	acggttatgc	15900
ggctctgcat	tgcagcattt	cttacctgga	aatctctacc	tgcactatca	ataacctcat	15960
gcacggccccg	atccgcagcc	aaacgcacga	aagcttggtc	ccatgtagca	agaacctctc	16020
gcttcgtgcc	ctccacaatt	accgtataac	ggtctgcac	tggcttggtt	gcaggatcaa	16080
tctttacctc	tttggagaa	ttcttataaa	actcctcctg	ctcatcttcg	ctttccgctc	16140
cctcaccac	ctggggaagg	tatatcttat	agaagaactc	ttttttattc	tcacccgaaa	16200
gcagattgcg	cgaaaagatc	gcatagatac	ttctaacagt	atatgcatct	aggccattct	16260
cccttctacc	aacattgaaa	ccttcgatag	acccaagagc	aggaggcact	actgcaaccc	16320
tatattcata	tacaggctta	ctcagatacg	cataggtaaa	agacccgatt	aatgcaagaa	16380
gagtagtcag	aagaatcaga	accttggtta	cccaaagctc	cttgaccagc	ttcaccaggt	16440
caacctcacc	atcagccgtc	atcaaagaag	aattgcctca	ggg		16483

<210> 4

<211> 16660

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 4

aattccctga ggcaattctt ctttgatgac ggctgatggg gaggttgacc tggggaagct	60
gggtcaaggag ctttgggtta acaagggttct gattcttctg actactcttc ttgcattaat	120
cgggtctttt acctatgcgt atctgagtaa gcctgtatat gaatataggg ttgcagtagt	180
gcctcctgct cttgggtcta tcgaagggtt caatgttggg agaagggaga atggcctaga	240
tgcataatact gttagaagta totatgcgat cttttcgcg c aatctgcttt cggatgagaa	300
taaaaaagag ttcttctata agatatacct tcccaggtg ggtgaggag cggaaagcga	360
agatgagcag gaggagtttt ataagaagtt ctcaaagag gtaaagattg atcctgctaa	420
caagccagat gcagaccgtt atacggtaat tgtggagggc acgaagcgag aggttcttgc	480
tacatgggca caagctttcg tgcgtttggc tgcggatcgg gccgtgcatg aggttattga	540
tagtgcaggt agagatttcc aggtaagaaa tgctgcaatg cagagccgca taaccgtgct	600
gcagaatatg gcgaagggcc gccgtgatga tagaattgca cgtttgaagg aggcattgct	660
gattgcggag tcgctcaaga tagatggccc gccattaata gaaggggcgt ccgagcaaca	720
actctcctcg atcatggatg gtgacttgat gtacatgcga ggagctaagg cgctgcgcgc	780
tgaaatcaac aatcttgagt cgcgtagtgt agatgctcca ttcattcctg agttgagaac	840
tctccaagag aaactatctt ggaactccag tttgtctgtg gattctgatg cgggtggctgt	900
ctacaaggaa gacgaggac tctctttttc aaatcaaccc attaagccga agaagattct	960
tatagttact ataggtactt tggcaggatt gataattgga attctactcg cagtgcctgc	1020
tggttttata aggaagcttc gtagcgatgg ctctcttcgc taagttttga tctaaaccct	1080
gatgccactt actggcatca gggcttactt gttgtagttg ttgaattgta ttagggggta	1140
tcgagtagaa tgtcaataag gcggggcggt ttctactccg gtatatccat gggcagcaac	1200
tatcttttgc ctcttgctgc gattcctttt ctcacgagaa cattgtcaag cgaagcggtt	1260
ggccaattgg tgattgcca ggccgtggct gtcattctat gtcaactggg agactttgga	1320
tttattctgg caggatcaag aaaggctgcc attatcgata acaaagttga actgtctagt	1380

ttcttttctg ttgtacagag tgctagattc ttattgttgc tgctttcact tttagtgtctg	1440
gccatttttg ctgtatcttc tattttacca atccccttgc ttgtattggg tgcggctgct	1500
cttcoggcag tagttggaaa ttatcttcaa gcagtatggg tctttcaggg aagagcgtctg	1560
tttggtatggg tggcgcttac caattttttg tctaagggtat tttatttcct attgggtcgtt	1620
ttttttgtca cgaaggattc tgaccttggt ctggcttcgt tgggggttgg tttttcctat	1680
gtcataggtg gaagtgtctt ctgttgtatt ttattttcta tgggaatacg gtggcgccccg	1740
gttctcgaga aagacagaat tctcgatata ttgcgtgacg gtgctcgatc ttttctttct	1800
ctggcttttc ttagcttgca catgcaagtg ctcggtgcgg cgggttggtgt tgttggtgga	1860
gcctccgcgg ccggagtgtt ttctactgcg gataaattcc ttcgcgggat cgcggctgct	1920
acttcaccca tagctagcgc tctatttcgg acttttagca ggatgtatgc gagtgcgcac	1980
ccggcagtog gcagtttaag aaggaaagcg ctaggtctga tgttactaat agctattcct	2040
agttgtttat ttcttttctt attttctgaa tacatttcat atctcctatt cccggaacag	2100
tccagaggtc taactgttgt aataagaatg ttttcgatag tgccagtgtt tgcttgtatt	2160
ggtgttctgt atggaggggt gactcttggt ccttctgggt atgatgggtgt atatttgcca	2220
gcaatttttt ttgcggaatt gggcggggta ttaacattta tcctcttggc gctttggggg	2280
gatgagcttt ttggagcgtg gacgctgggc gttacagagg tctctttggg gatgggaatg	2340
tttttcctgg ccacgggtta gttgagagag aaaaggggac tttgatctta aggcgatcgc	2400
taggaggaca gctatgcgtt ttgcgattct gggcgcgacc aacattaaac atatgagcct	2460
gctgagccat tatctgaacc atattgatct gaacattaac gaagtggata ttatttatac	2520
cgataaatat gatattgaag aacatattca gggcatcaac aactactaca aatacaaagt	2580
ggatatcaaa gaagattgga ccttcatcaa gaaagcgatt gcgtattatc gttttcgtcc	2640
gtatgcgatg aaaattctga aagaaaaccg ttatgatttt gtgatttgtt ggggcagcta	2700
caccggccat ctgttcaaaa gctttctgga aaaacattac aaaaacaaat tcctcctgaa	2760
catccgtgat tactttttctg aaaacaacaa actgattaaa tatcgtatga agaaaatcgt	2820
ggatgcgagc cgtgtgacca ccctgagcag cgaaggcttt ctgaaattcc tgccgaaaag	2880
cgaaaaatac cgtatcatct acagctacaa catgagcatc atccgtgaaa gcaacgtgac	2940
cgatggcttt aaaaaacggt ggccgattaa cattggcttt attggcaacg tgcgttttaa	3000
cgaaattaac cagaaactga ttaaagaact ggcgaacgat agccgttttc atatgcagta	3060

ttttggcacc ggcagcgaaa aactggaagt gtttgcgcgt gaaaacttta ttaacaacat	3120
taccttttagc ggcggctttg atctgaaaga aaccccgaaa tatctgaacg aaattgatat	3180
tctgaacaac ctgttttgca accagaacat tgcgctggat accgcgctga gcattcgtat	3240
gtattatgcg ctgtttctga acaaaccgat tattaccacc gatgatacct ttaccgcgac	3300
cgaagcgaac aaatttggcc tgggcttttag cattaaccog gaaaacctga aaggcattgg	3360
cgatgaactg atggattggt ataacaacct ggatgtgatg gatattaacc ataaacgtga	3420
agcgtatcgt aacgatgtga ttgaaaacaa caaacagttt tatcaggaaa ttggccgtat	3480
ttttaacgaa gaacagaaac tgattagcga agaagatctg taacgtttaa acaggaggac	3540
agctatgaac aaaattttata acgtgaccag ctatgtgatt gcgattctga tgtttccgtg	3600
cctgatgctg ggcgataaac cgctgctggt tctggcgccg attagctatg gcgtgggcaa	3660
actgttcacg agcttcagca acaacccgaa cttcaaattc agcaaaatcg tgtaocgatgt	3720
gctgggcttt ctgcgtctgg tgtttattcc ggcgatgatt gtgtttttcc aggatagcac	3780
cattgataac ctgccgctgg gccaggcgta ttttaaccag gcggtgattt atatgagcgt	3840
ggaatttatt attggcagcc tgtttattct gattctgagc aaactgttca agcatgaagt	3900
tgtgagccgt aacagcttta ccctgagcgg cagcagcatt tattatattg tgtttggcct	3960
ggtgatttgc ggcatttttg tggcgtttcc ggaagtgcgt aaaaacatta gctttctgat	4020
tattaaaacc gatgogatgg gccgtggcac cgaagcgacc agcggcctga acgtgctgtt	4080
tgtgatgctg tttcagctgg cgctggcgct gctgtttctg atcatcgcgt acgogagcta	4140
caaaaaatac aaagaaaacc cgaaaatcat ctacgtggtg ctgccgctgg cgattggcat	4200
tctgaacatt agcctgattg tgggcgaacg tcgtagctat cagctgtata ccatggtggc	4260
ggtgctgacc gttgtgagca tctgttttag caaacataaa cgtcgtatca acatcatcat	4320
catcagcgtg ggcattcttcg tgctggcgct gatgaccctg tataaagaac tgtatgtgtt	4380
taactatagc agctatagcg aagcgctgaa cagcaccagc gtgagcaacc tgaaaattgt	4440
ggataccctg cagagctatt tttatggccc gagcaacatt gggcgagca ttgattatct	4500
gaactattat aacggcagct ttaaacagta tctgtttgat aacacccgtg cggtgtttgg	4560
ctttaacttt ttctgggata aaaaacagct gattaccagc cagctgttta accagctgat	4620
ttatggcagc aaacagctga ccggccatct gattagcagc gcgggctatg gcattattta	4680
ttttggcccg ctgtttttct acctgaacct gattgcgaac atcttttttcg cgtttctgag	4740

cgaatacatc atccgtaaaa gccatagcct ggaagtgate ttcacggca cctacatcta	4800
catgcgtctg attaccagca ttttttagcca tccgaccccg ctgattacco tgattagcat	4860
gattctgggtg gtgtatgtga ttgcgatcat cccgggcatc atcatcaaga aattcaccaa	4920
aaaagtgggc atcgaagatt acaaagatga tgatgataaa taacgtttta acaggaggac	4980
agctatgatt gtgaaaacct ttatgaaaag caaaatTTTT cgtctgatga acaccccgct	5040
gctgctgttt tataaaaaag aatatctgac cggctattat tttgaaaaca aagtggcggg	5100
ctggctgtgg gcgtggaaag cggtgccatt caagctgctg ggcattaaca ccagcctgcc	5160
gtttccggcg gatattaccg tgcgtatgca taaccogaac aacattgtgt ttgataaaaa	5220
cgatattcat atttttcaga gcccgggcac ctattttaac aacttttagcg cggtgattta	5280
tattggccgt ggcggtgtata ttgcgccgaa cgtgggcatt attaccgca accataacat	5340
taaaaacctg aaaagccatg cgccgggcca agatgtgaaa attggcaact atagctggat	5400
tggcatgaac agcgtgattc tgccgggctg ggaactgggc gaacatacca ttgtgggcgc	5460
gggcagcgtg gtgacaaaaa gctttccgga aggcaacgtg gtgattggcg gcaacccggc	5520
gaaaattatt aagaaaatca gctatccgta tgatgtgccg gattatgctg aattaattaa	5580
ccagggtgcac gaagaaaatt atgagattaa ataaatttat tggcgattcg tttttaatga	5640
ttttaagcag tggcatcgct caagtcatat taatcatcac taccccaatt attacaagac	5700
tatattcacc tacagaattt ggtgagttaa caattttttc aaatatcgca atgattttaa	5760
taccaataat aaatgcaaga tacgatttgt tgattgtgaa taccaaaaat gaccgtagt	5820
ctaataact ttcacaaatc agttttttga tatcattgct tattttatta atactgatac	5880
caatatttgc gattagtga tgtttatacc caaactttat attagatttt attttcatta	5940
ttattatgtt gtttttggta agtttaacaa acatttttac aaattatcta aataaggaaa	6000
gaaagtataa agtgtaaagt ttgattaatg tgttttagagc tggatcaatg gctttacttc	6060
aaatcatttt cggactttta gcattaggaa gtttaggatt aattattggg ttttcattat	6120
cctatatcgc aggcattaca ctaggatata aaacgtttta aaagcacttt aatattgtga	6180
gagataaaga agaaactaaa gcattatttt tagaaaataa aaatcagtta gtttattcaa	6240
caccatcaat attattaaat agtttgtctt tctcggttgt tgtgttcttt ataggtattt	6300
tgtataccaa tacagaagtg ggtatttatg gtatggccat aagagtacta ggcataccag	6360
tgacaattat ttcattaggg ttatcaaaaa tatttatgca acaagccaat gactattata	6420
ttgaacatgg taacttccga aattttattac ttaaatttag ttccatactg gttatagttt	6480

ctataattct ttatgtgcc	ctttatttgt tcagtgaaga	attagtcaat atattattag	6540
gacatagctg ggttgacgc	attacagtta taaaaattgt	tatcccatta tttgttataa	6600
ggctgattgt atcaacggta	tcactttctg tgattgtatt	acaaaaacaa cagttagaat	6660
taatactaca agcgttattt	ttaataggta ctactgcaac	at ttgttata tcaaaaatgc	6720
ttaatttaac ttttttaaac	tttgtatcta ttaatacaat	tg ttttaatc gtatcgtaca	6780
tgatattttt catagcactc	tattattttg ctaaaaataa	acagttcaaa aattctagtt	6840
atccgtatga tgtgccggat	tatgcgtaag tgcaccatat	gaatatcctc cttagttcct	6900
attccgaagt tcctattctc	tagaaagtat aggaacttcg	gcgcgcctac ctgtgacgga	6960
agatcacttc gcagaataaa	taaatcctgg tgtccctggt	gataccggga agccctgggc	7020
caacttttgg cgaaaatgag	acgttgatcg gcacgtaaga	ggttccaact ttcaccataa	7080
tgaaataaga tcactaccgg	gcgtattttt tgagttgtcg	agattttcag gagctaagga	7140
agctaaaatg gagaaaaaaa	tcactggata taccaccggt	gatatatccc aatggcatcg	7200
taaagaacat tttagggcat	ttcagtcagt tgctcaatgt	acctataacc agaccgttca	7260
gctggatatt acggcctttt	taaagaccgt aaagaaaaat	aagcacaagt tttatccggc	7320
ctttattcac attcttgccc	gcctgatgaa tgctcatccg	gaattacgta tggcaatgaa	7380
agacggtgag ctggtgatat	gggatagtgt tcacccttgt	tacaccgttt tccatgagca	7440
aactgaaacg ttttcatcgc	tctggagtga ataccacgac	gatttccggc agtttctaca	7500
catatatctg caagatgtgg	cgtgttaacg tgaaaacctg	gcctatttcc ctaaagggtt	7560
tattgagaat atgtttttcg	tctcagccaa tccctgggtg	agtttcacca gttttgattt	7620
aaacgtggcc aatatggaca	acttcttcgc ccccgttttc	accatgggca aatattatac	7680
gcaaggcgac aaggtgctga	tgccgctggc gattcagggt	catcatgccg tttgtgatgg	7740
cttccatgtc ggcagatgct	taatgaatac aacagtactg	cgatgagtgg cagggcgggg	7800
cgtaaggcgc gccatttaaa	tgaagttcct attccgaagt	tcctattctc tagaaagtat	7860
aggaacttcg aagcagctcc	agcctacacg taaagagggtg	gtgtatggat aagaactctg	7920
ttctgttaat taccggtgga	actggttcct ttggaaatgc	cg ttttgaag cg ttttctgg	7980
atacagatat tggtgaaata	cgtgtattca gtcgggatga	gaagaagcaa gatgatatgc	8040
gtaagtgcta tgctcaccoc	aagttgaggt tttatattgg	agatgtccgc gactatcaaa	8100

gcactctgaa tgctactogc ggtgtggact acattttcca tgcggctgca ttgaaacagg	8160
ttccctcttg tgagttctac ccgatggagg ctgtgaagac caatgtcatc ggaacggaaa	8220
atgtcctcga atctgctatc cagaatggcg tcaaaaaagt cgtttgtctg agtacggata	8280
aggcagttta cccaatcaat gccatgggta tttcaaaggc catgatggaa aaggatcatg	8340
tcgccaagtc tcgaaacctg gaacgcactc ctaccgtaat ttgtggcacc cgttatggta	8400
acgtcatggc ttcaaggggc tcggtcattc ctctctttat cgagcaaagc cgatcaggcc	8460
agcctctcac aatcactgat ccaaacatga cacgtttcat gatgacgctt acggatgccg	8520
tagaccttgt gctttatgct ttcgagcatg gtaccaatgg agatcttttt gtacagaaaag	8580
caccggctgc aaccatcgag gtgctggctc atgogcttac tcaattgctt ggcaagaatg	8640
gttatcctat caatgtaata ggtacgcgtc atggagagaa gctttatgag gcgcttotta	8700
gtcgagagga aatggcctgt gctgaagaca tgggtgacta ttatcgtatc ccccagatt	8760
tgcgtgatct gaactatagt aagtttgtgg agcaaggcga ggaaaaaatt actcatacag	8820
aggactataa ttctcataat accaaacgtc tggatatoga aggaatgaaa aagctgttgc	8880
tgaagttgga tttcattcgt gctattcagc gtggcgagag tgccagtcca gaggaataaa	8940
cgatgaaagt tcttgtaact ggcgcgaatg gatttgttgg aaggaatctg tgcgctcatc	9000
ttgcagagcg ggggtggtatc gaggtggtgc cattcaccgc cgagagtagt gttggttaatt	9060
tgcctgagct aattcgttcc gtcgatttta tttttcatct tgccgggggc aatcgctccg	9120
aaaaaccaga agagtttaag atcgggaatt ccgaactcac gtatgctctg tgtgaggcgg	9180
taaggtccaa tggacgagcc ataccacttc tttatacttc atccattcag gctgaggtgg	9240
ataatgagta cggtttaagc aagcgagccg cagaagagca tctccaagtg ctaggtgagg	9300
atattggttg tctgtgtctac atatttcgcc ttctaatgt attcggtaaa tggtcgcgtc	9360
cgaattataa ttcagcggtt gcgacttttt gtcataatat tattcgagat attccgattc	9420
aaattaacaa ttcctcggca gagatcactc ttgtatacat agatgatgtg gttcgcacct	9480
tcatgaaagt catggatggg aagctatcca atgcagtttc actacaggtc gagccccagt	9540
atcagatttc tgttggtgag ctgcgagaac aattgtatga gtttcgtaat agtcgaaagt	9600
cactgactac cgcaaggggt ggctcgggat tgacgcgcgc cttgtactcg acttatctaa	9660
gtttcttgcc agaagatagt tttagttacg acgtgccaat gcattcggat ccgcgtggca	9720
cattcgtcga gatgctgaag accgcggact ctggccagtt ctogtttttt acggetcatc	9780

caggtgttac	caggggacgg	cattaccatc	actcgaaaac	cgaaaagttt	ctggttatca	9840
aagggatggc	acgtttcaag	tttagaaaac	tcctgaccgg	ggcattttac	gaaatttgca	9900
ctaattggtga	aaaggcagaa	attgtcgaaa	cagtacctgg	atggactcat	gacattacta	9960
atgtcggaaac	tgacgatatg	gtagtcatgt	tgtgggctaa	cgaagtattt	gatcgggaaa	10020
atccggatac	ctacgcttgt	tcagtagggc	aagggtgcgt	aggtatagtg	agataacaat	10080
gcagaagcta	aaagtcgtta	cggttggttg	aactcgctct	gagattattc	gcttgctctag	10140
ggctcatggcg	aagcttgatc	agtactgcga	tcattgtactt	gtccatactg	gacagaatta	10200
tgattacgaa	cttaatgaaa	tattttttca	ggacctcggt	ataagaaagc	cggattattt	10260
tctaaacgcc	gccgggtctt	ccggggctga	aacgataggg	aatgtaataa	tcgcagtcga	10320
tcgtgttctg	ggcgaaatag	atcccgatgc	gctgctcggt	ctgggtgata	ccaatagttg	10380
tatggcggta	ctgcctgcaa	aacggcgtaa	gataccgacc	tttcatatgg	aagcaggcaa	10440
tcgctgtttc	gatatgcgtg	tgccctgaaga	gataaatcgg	cgcattgtcg	atcatacagc	10500
tgatgtaaat	ttgacctata	gtacaattgc	gcgtgattat	ctcttgctg	aaggactttc	10560
tccagacatg	gttatcaaga	ctggtagccc	tatgttcgaa	gttctcgagc	actatcgtga	10620
cgggatcgag	tcctccgata	ttcttgaaag	gctcgggttg	aaaacagagc	ggttctttgt	10680
cgtgagtgcg	caccgagagg	aaaacataga	ttcgataaag	aatttcttga	agttggtttc	10740
tatgctcaac	gctgtggcag	aaaagtactc	gctgcccgtc	atcgtatcaa	ctcaccctag	10800
aacaaaaaag	agaattgagg	cgacggaggc	aaagtttcac	gagggtatta	aactgctgaa	10860
acccctcggc	tttaaggatt	acaataaact	gcaaattaca	gccaaaggcag	ttatttctga	10920
cagtgggacc	atcagtgagg	agtcttcaat	actgaatttt	cccgccttga	atattcgtga	10980
ggctcatgaa	cgcccagaag	gcatggaaga	ggctgtgggt	atgatggctg	gactggattc	11040
ggatcgagta	ctacaagcac	tcgaggtgtt	ggagggacag	aggcgcgacg	cagagcgcat	11100
gttacgcttg	gtcgctgact	atagcatgcc	caacgtttct	gaaaagattg	ttcgcatagt	11160
tcatagctat	cgggactatg	tcattgcgaac	tgtctggaaa	aaatattaac	ttgaggcgtg	11220
gagttgatgg	caaggatatt	tgtggtttct	gagtatgtcg	gtgccaatca	gaactccacg	11280
ggatactatt	gggagaagat	aataggaaaag	atgcagcggg	agtttggtgg	gctaaccgta	11340
attttccgcg	tgaccgcagg	tgaaaccccg	cctgtgggtt	caccttccgt	tgagcaagaa	11400
tgctttaagt	ttccgaggag	caataagaat	aggctccttt	ctagaggatt	ggcgcagatt	11460

tttcaggcgt	ttctgttctc	agtaaaattg	acttctcgtg	ccagacgagg	agatgtggta	11520
ttgagtggaa	ccaaccctgc	tcttctactg	atgacgtttc	ccttgctaag	gtatgccctc	11580
ggtttcaagt	gggtgctgct	ggtgcatgat	gtgtttcccg	agaacttggg	gccggcgggc	11640
gttctgaaga	aagatagtat	tgcctaccgg	cttctacgtc	gtctcttttc	tttcatttac	11700
tcacccgctg	atcgtctagt	cgtaataggg	cgcgatatgg	aagctcttat	gaaagagaag	11760
gtgaatgacc	cgcgatcttt	ggtctttatt	tcgaattggg	cctgtgagaa	agaggttttc	11820
ccagtaccga	gagaggatgc	tcctttttatc	aatattcctg	aatggaaagg	taaaagggtt	11880
ttccaatttt	ttggtaatgt	cggtcgatta	caaggatatag	aaaacatact	ttctgctatt	11940
cagttgggta	aaaacgagaa	ggcggccttt	gcttttattg	gagatggtgc	cttggctgcac	12000
agtgtaaaaa	aacacgcgct	ggaagatcag	tgtgctcggg	tgaggatatt	tggaaggctg	12060
ccattagccg	aaaagaattt	tggtttggct	gcctgtgacg	ttgccttagt	taccttagaa	12120
gaaggaatgt	tcgggcttgg	ggttcccagc	aaggcatatt	tctccatggc	agcagacaaa	12180
ccgattctag	ctgtcatgga	aaaaggggct	gaaatctccc	gtataataga	tgagaccgga	12240
atcgggttga	actgtccgcc	gaatgatccg	gttgctttgg	caagattgat	cgatgagatt	12300
tgtgaactcg	acttgtctag	tttaggcgga	gtcccgcgga	gtgtccttca	gcaaaattat	12360
tctgaatata	tttcattgga	aaaattcgcg	gcctgtgttc	gaccgcttct	gtctgagtcg	12420
aaaatatgat	gaaggtgctg	gtaaccgggg	ctagcggttt	tgtcgggagt	gcgctttgca	12480
ggtcgcttgc	tgccgcccc	tttcagggtg	tcggacaagt	acgatccctg	tacaatcccg	12540
ttacgggggt	tgagtatggt	cgagcggagc	tgaaagagag	cactaagctt	gatgctgcgc	12600
tgcgggggtg	tgaatgtgta	gttcatctag	ctggacgagc	ccatatcttt	ggaaggcagc	12660
gtgattcact	agatattttt	cggaagggtga	atcgcgatgc	tactctggcg	cttgctcggc	12720
aggcgatcga	agcatctgta	aagcgtttca	tttttggttag	ttctattggg	gtaaatggcg	12780
ctttaaccaa	agaaaagccc	ttcgatgaga	actccaagcc	ggctcctcat	gcagaatatg	12840
cgatttcaaa	gtttgaggct	gaagtagcgc	ttcgggagct	tttcaagcat	tcctcaacag	12900
aacttgttat	cgtcaggcct	ccactcgttt	acgactggaa	agctcctgga	aatttctcgc	12960
gattgttgaa	gctggttgct	tcgggacttc	ctcttccatt	tggttgcata	gataaccgac	13020
gaagttttgt	ttctctggat	aatttagttg	actttctagc	ttgctgtatg	acgcaccctt	13080
ctgctgccgg	cgaactgttt	ttggtatccg	atggtcagga	gatttctacc	aagcaactgg	13140
tgactgcgct	tgctgcggga	atggggcgtc	gccccatcat	gtggcctggt	cctaggttta	13200
ttctgagggt	tcttaaatta	gtaggaaagg	gtgggttata	cactcagtta	tgctgctcac	13260
tagaggtcga	ctcgtcgaaa	ggcaggcttt	tgcttggttg	ggaaccccg	aagagcacc	13320
tttccgcggt	ggaagatggt	ggtagaatat	atgtcaaacg	tactgaatga	ttatctgcag	13380

gcgcttttgct	actagcatgg	cgtaccacgc	agaacaatcg	aatagaaccc	tgttgaaggg	13440
gtgagagtat	ttttggggat	aaatttataa	atggaagaat	ggatatttggt	actcgctgca	13500
gctgggggttt	cgggactgct	tacaggcctc	ttgcgtcggt	atgccttagc	gaggagctta	13560
cttgacaccc	ctaactctcg	aagttcccat	gtcgttccca	ctccacgcgg	aggaggggtc	13620
gccattgtag	ttactttttg	tctcatgctg	cctatttggg	ctgtactggg	aaatatctca	13680
tgggccgtgt	cctgggcttt	acttctcgct	ggcggcgggg	ttgccattat	tggattcatg	13740
gatgatcacg	gtcatatcgc	cgcacgctgg	cgtctgctgg	gacattttag	tgcagccttg	13800
gtctcattgt	actttttgaa	tggcatacca	ccatttcaga	ttgttggtgt	cagttgggac	13860
ctgggggtgt	tcgaggact	tctctttgct	ttctatctcg	tgtggttgct	gaatctctat	13920
aacttcatgg	atgggatcga	tggacttgct	agccttcagg	ccatttttgt	ctgtgttggt	13980
ggggcattat	tatactggct	gaatggccaa	ctgacgcagg	ctttgctccc	cttatcgcta	14040
gcttttgccg	tttttgatt	cttgttctgg	aattttccac	ccccaaaaat	tttcatggga	14100
gatgcgggta	gtggtcttct	ggggattggt	ttaggaattc	tttccattca	tgccatgtgg	14160
atgaatacga	attttttctg	ggcatggttg	gtcctgttag	gcgttttcat	cgtcgatgcg	14220
acctataccc	tgattcgtcg	cttgctgaga	ggggacaagg	tgtatgaggc	tcatcgaagc	14280
catgcctatc	aatacgcaag	ccgatactat	ggaaagcatg	ctcctgttac	gattggcgtc	14340
acggcattga	acgtcatctg	gctcctccct	atagccttgt	tggtcgggag	tgggtctcta	14400
gagcctttga	tgggcatcgt	catagcctac	gtccctctcg	tttttctggc	agtgaggttc	14460
aaggcgggta	agctagagtc	gtccgctcag	gcctaaagga	gtaggggaat	gctagatcgt	14520
ttaagagtaa	agttgttatc	catgcctcgt	cgctggaaac	gtttgcttca	agtggctacg	14580
gatatacctc	tggatatggct	gtctctgtgg	ctcgcttttg	tggtcctgtc	aggcacagac	14640
gatatgatcg	acgtgttcgg	cgagcatgca	tggcttttca	tactgcgcc	ggtcatcgcc	14700
attccactat	tcatcgcctt	cggcatgtat	cgcgcggtga	tgcgctatct	cggtaacgac	14760
gcattgatcg	ccatcgccaa	ggcggtgacc	atctcggtc	tggtgctgtc	gctggtggtg	14820
tactggtatc	gtggcgcgcc	ggcgccggtg	ccgcgttccc	tggtgttcaa	ctactggtgg	14880
ttgagcatgc	tgctgatcgg	cggcttgctg	ctggccatgc	gccagtattt	catgggcgac	14940
tggtagtctg	ctgtgcagtc	ggtaccattt	ctcaaccgcc	aggatggcct	gccaggggtg	15000
gttatctatg	gggcgggggc	ggccggcaac	cagttggttg	cggcggttgcg	tctcggtcgg	15060
gcgatgcgtc	cgggtggcgtt	catcgatgac	gacaagcaga	tgcceaaccg	ggtcattgcc	15120
ggtctgcggg	tctataaccg	caagcatatc	cgccagatga	tcgacgagac	gggcgcgcag	15180
gaggttctcc	tggcgattcc	ttccgccact	cgggcccggc	gccgagagat	tctcgagtcc	15240

ctggagccgt tcccgtctga cgtgcgcagc atgcccggt tcatggacct ggccagcggc	15300
cgggtcaagg tggatgacct gcaggaggtg gacatcgctg acctgctggg gcgcgacagc	15360
gtcgcaccgc gcaaggagct gctggaacgg tgcattccgc gtcagggtgt gatggtgacc	15420
ggggcgggcg gctctatcgg ttccggaactc tgtcggcaga tcatgagttg ttgcgcctagc	15480
gtgctgatcc tgttcgaaca cagcgaatac aacctctata gcatccatca ggaactggag	15540
cgtcggatca agcgcgagtc gctttcgggtg aacctgttgc cgatcctcgg ttccggtgcgc	15600
aatcccagac gcctggtgga cgtgatgcgt acctggaagg tcaataccgt ctaccatgcg	15660
gcggcctaca agcatgtgcc gatcgtcgag cacaacatcg ccgagggcgt tctcaacaac	15720
gtgataggca ctttgcagtc ggtgcaggcc gcggtgcagg tcggcgtgca gaacttcgtg	15780
ctgatttcca ccgacaaggc ggtgcggccg accaatgtga tgggcagcac caagcgcctg	15840
gcggaaatgg tccttcaggc gctcagcaac gaatcggcgc cgggtgctgtt cggcgaccgg	15900
aaggacgtgc atcacgtcaa caagaccgt ttaccatgg tccgcttcgg caacgtcctc	15960
ggttcgtccg gttcgggtcat tccgctgttc cgcgagcaga tcaagcgcgg cggcccgggtg	16020
acggtcacc acccgagcat caccggttac ttcatgacca ttcccaggc ggcgagttg	16080
gtcatccagg ccggttcgat ggggcaggcc ggagatgtat tcgtgctgga catggggccg	16140
ccggtgaaga tcctggagct cgcggagaag atgatccacc tgtccggcct gagcgtgcgt	16200
tccgagcgtt cggcccatgg tgacatcgcc atcgagttca gtggcctgcg tcctggcgag	16260
aagctctacg aagagctgct gatcggtgac aacgtgaatc ccaccgacca tccgatgac	16320
atcggggcca acgaggaaca cctgagctgg gaggccttca aggtcgtgct ggagcagttg	16380
ctggccgccc tggagaagga cgactactcg cgggttcgcc agttgctgcg ggaaaccgtc	16440
agcggctatg cgcctgacgg tgaaatcgtc gactggatct atcgccagag gcggcgagaa	16500
ccctgagtca tcgttctccg gaaaaggccg cctagcggcc tttttgttt tctccgtacg	16560
atgtttccgg tgccggacca ggaagcgact gctttgctgg ggctgtcgat ccagggtgcgt	16620
tccacggcga taaggtggtt tcgtggatgg gcaacatgtg	16660

<210> 5

<211> 972

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 5

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcg	60
gcggatagcg atattaacat taaaaccggc accaccgata ttggcagcaa caccaccgtg	120
aaaaccggcg atctggtgac ctatgataaa gaaaacggca tgctgaaaaa agtgttttat	180
agcttcatcg atgataaaaa ccataacaaa aaactgctgg tgatccgtac caaaggcacc	240
attgcggggc agtatcgtgt gtatagcgaa gaaggcgca acaaagcgg cctggcgtgg	300
ccgagcgcg ttaaagtgca gctgcagctg ccggataacg aagtggcgca gattagcgat	360
tattatccgc gtaacagcat tgataccaaa gaatatatga gcaccctgac ctatggcttt	420
aacggcaacg tgaccggcga tgataccggc aaaattggtg gactgattgg cgcgaacgtg	480
agcattggcc ataccctgaa atatgtgcag ccggatttta aaaccattct ggaaagcccg	540
accgataaaa aagtgggctg gaaagtgatt tttaacaaca tgggtgaacca gaactggggc	600
ccgtatgatc gtgatagctg gaacccggtg tatggcaacc agctgtttat gaaaaccctg	660
aacggcagca tgaaagcggc ggataacttt ctggatccga acaaagcgag cagcctgctg	720
agcagcggct ttagcccgga ttttgcgacc gtgattacca tggatcgtaa agcgagcaaa	780
cagcagacca acattgatgt gatttatgaa cgtgtgctg atgattatca gctgcattgg	840
accagcacca actggaaagg caccaacacc aaagataaat ggattgatcg tagcagcgaa	900
cgttataaaa ttgattggga aaaagaagaa atgaccaacg gcagccatca tcatcatcat	960
cattaggtcg ac	972

<210> 6

<211> 990

<212> ADN

<213> Sequência Artificial

<220>

<221> font

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 6

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcg	60
gcggatagcg atattaacat taaaaccggc accaccgata ttggcagcaa caccaccgtg	120
aaaaccggcg atctggtgac ctatgataaa gaaaacggca tgctgaaaaa agtgttttat	180
agcttcatcg atgataaaaa ccataacaaa aaactgctgg tgatccgtac caaaggcacc	240
attgcggggc agtatcgtgt gtatagcgaa gaaggcgca acaaaagcgg cctggcgctg	300
ccgagcgcggt ttaaagtgca gctgcagctg ccggataacg aagtggcgca gattagcgat	360
tattatccgc gtaacagcat tgataccaaa gaatatatga gcaccctgac ctatggcttt	420
aacggcaacg tgaccggcga tgataccggc aaaattggtg gactgattgg cgcgaaacgtg	480
agcattggcc ataccctgaa atatgtgcag ccggatttta aaaccattct ggaaagcccg	540
accgataaaa aagtgggctg gaaagtgatt tttacaaca tggatgaacca gaactggggc	600
ccgtatgatc gtgatagctg gaaccgggtg tatggcaacc agctgtttat gaaaaccgt	660
aacggcaaag atcaaaatag aactaaaatg aaagcggcgg ataactttct ggatccgaac	720
aaagcgagca gcctgctgag cagcggcttt agcccggtt ttgcgaccgt gattaccatg	780
gatcgtaaag cgagcaaaca gcagaccaac attgatgtga tttatgaacg tgtgctgat	840
gattatcagc tgcattggac cagcaccaac tggaaaggca ccaacaccaa agataaatgg	900
attgatcgta gcagcgaacg ttataaaatt gattgggaaa aagaagaaat gaccaacggc	960
agccatcatc atcatcatca ttaagtgcac	990

<210> 7

<211> 990

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 7

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcg	60
gcggatagcg atattaacat taaaaccggc accaccgata ttggcagcaa caccaccgtg	120
aaaaccggcg atctggtgac ctatgataaa gaaaacggca tgctgaaaaa agtggtttat	180
agcttcatcg atgataaaaa ccataacaaa aaactgctgg tgatccgtac caaaggcacc	240
attgcggggc agtatcgtgt gtatagcgaa gaaggcgca acaaaagcgg cctggcggtg	300
ccgagcgcg ttaaagtgca gctgcagctg ccggataacg aagtggcgca gattagcgat	360
tattatccgc gtaacagcat tgataccaaa gaatatatga gcaccctgac ctatggcttt	420
aacggcaacg tgaccggcga tgataccggc aaaattggtg gactgattgg cgcgaacgtg	480
agcattggcc ataccctgaa atatgtgcag ccggatttta aaaccattct ggaaagccc	540
accgataaaa aagtgggctg gaaagtgatt tttaacaaca tgggtgaacca gaactggggc	600
ccgtatgatc gtgatagctg gaaccgggtg tatggcaacc agctgtttat gaaaaccgt	660
aacggcagca tgaaagcggc ggataacttt ctggatccga acaaagcgag cagcctgctg	720
agcagcggct ttagcccgga ttttgcgacc gtgattacca tggatcgtaa agcgaaagat	780
caaaatagaa ctaaaaaaca gcagaccaac attgatgtga tttatgaacg tgtgcgtgat	840
gattatcagc tgcattggac cagcaccaac tggaaaggca ccaacaccaa agataaatgg	900
attgatcgta gcagcgaacg ttataaaatt gattgggaaa aagaagaaat gaccaacggc	960
agccatcatc atcatcatca ttaagtcgac	990

<210> 8

<211> 990

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 8

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcg	60
gcggatagcg atattaacat taaaaccggc accaccgata ttggcagcaa caccaccgtg	120
aaaaccggcg atctggtgac ctatgataaa gaaaaccggc tgctgaaaaa agtgttttat	180
agcttcatcg atgataaaaa ccataacaaa aaactgctgg tgatccgtac caaaggcacc	240
attgcggggc agtatcgtgt gtatagcgaa gaaggcgca acaaaagcgg cctggcggtg	300
ccgagcgcggt ttaaagtga gctgcagctg ccggataacg aagtggcgca gattagcgat	360
tattatccgc gtaacagcat tgatacaaaa gaatatatga gcaccctgac ctatggcttt	420
aacggcaacg tgaccggcga tgataccggc aaaattggtg gactgattgg cgcgaaacgtg	480
agcattggcc ataccctgaa atatgtgcag ccggatttta aaaccattct ggaaagcccg	540
accgataaaa aagtgggctg gaaagtgatt tttacaaca tggatgaacca gaactggggc	600
ccgtatgatc gtgatagctg gaaccgggtg tatggcaacc agctgtttat gaaaaccgt	660
aacggcagca tgaaagcggc ggataacttt ctggatccga acaaagcgag cagcctgctg	720
agcagcggt ttagcccgga ttttgcgacc gtgattacca tggatcgtaa agcgagcaaa	780
cagcagacca acattgatgt gatttatgaa cgtgtgctg atgattatca gctgcattgg	840
accagcacca actggaaagg caccaacacc aaagataaag atcaaaatag aactaaatgg	900
attgatcgta gcagcgaacg ttataaaatt gattgggaaa aagaagaaat gaccaacggc	960
agccatcatc atcatcatca ttaagtcgac	990

<210> 9

<211> 1659

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 9

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcc	60
gcgagcgaaa acagcgtgac ccagagcgat agcgcgagca acgaaagcaa aagcaacgat	120
agcagcagcg tgagcgcggc gccgaaaacc gatgatacca acgtgagcga taccaaaacc	180
agcagcaaca ccaacaacgg cgaaaccagc gtggcgcgaga acccggcgca gcaggaaacc	240
acccagagca gcagcaccaa cgcgaccacc gaagaaaccc cggcgaccgg tgaagccacc	300
accaccacca ccaaccaggc caacaccccg gcgaccaccc agagcagcaa caccaacgcg	360
gaagaactgg tgaaccagac cagcaacgaa accaccttta acgataccaa caccgtgagc	420
agcgtgaaca gcccgcagaa cagcaccaac gcggaaaacg tgagcaccac ccaggatacc	480
agcaccgaag cgaccccgag caacaacgaa agcgcgcgcg agagcaccga tgcgagcaac	540
aaagatgtgg tgaatcaggc cgttaatacc agcgcgcgcg gtatgcgtgc ctttagcctg	600
gcggccgtgg ccgccgatgc tccagcagca ggtaccgata ttaccaacca gctgaccaac	660
gtgaccgtgg gcattgatag cggcaccacc gtgtatccgc atcaggcagg ttatgtgaaa	720
ctgaactatg gcttttagcgt gccgaacagc gcggtgaaag gcgatacctt taaaattacc	780
gtgccgaaag aactgaacct gaacggcgtg accagcaccg cgaaagtgc gccgattatg	840
gcaggtgatc aggtgctggc gaacggcgtg attgatagcg atggcaacgt gatttatacc	900
tttaccgatt atgtgaacac caaagatgat gtgaaagcga ccctgaccat gccagcatat	960
attgatccgg aaaacgtgaa gaaaaccggc aacgtgaccc tggcgaccgg cattggcagc	1020
accaccgcga acaaaaccgt tctggtggat tatgaaaaat acggcaaatt ctacaacctg	1080
agcatcaaag gcaccatcga tcagatcgat aaaaccaaca acacctatcg tcagaccatt	1140
tatgtgaacc cgagcggcga taacgtgatt gcgcgggtgc tgaccggcaa cctgaaaccg	1200
aacaccgata gcaacgcgct gattgatcag cagaacacca gcattaaagt gtataaagtg	1260
gataacgcgg cggatctgag cgaaagctat tttgtgaacc cggaaaactt tgaagatgtg	1320

accaacagcg tgaacattac ctttccgaac ccgaaccagt ataaagtgga atttaacacc	1380
ccggatgatc agattaccac cccgtatatt gtggtggtga acggccatat tgatccgaac	1440
agcaaaggcg atctggcgct gcgtagcacc ctgtatggct ataacagcaa cattatttgg	1500
cgtagcatga gctgggataa cgaagtggcg tttaacaacg gcagcggcag cggcgatggc	1560
attgataaac cgggtgtgcc ggaacagccg gatgaaccgg gcgaaattga accgattccg	1620
gaagatggca gccatcatca tcatcatcat taggtcgac	1659

<210> 10

<211> 1677

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 10

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcc	60
gcgagcgaaa acagcgtgac ccagagcgat agcgcgagca acgaaagcaa aagcaacgat	120
agcagcagcg tgagcgcggc gccgaaaacc gatgatacca acgtgagcga taccaaaacc	180
agcagcaaca ccaacaacgg cgaaaccagc gtggcgcgaga acccggcgca gcaggaaacc	240
acccagagca gcagcaccaa cgcgaccacc gaagaaaccc cgggtgaccgg tgaagccacc	300
accaccacca ccaaccaggc caacaccccc gcgaccaccc agagcagcaa caccaacggc	360
gaagaactgg tgaaccagac cagcaacgaa accaccttta acgataccaa caccgtgagc	420
agcgtgaaca gcccgcgaaa cagcaccaac gcggaaaacg tgagcaccac ccaggatacc	480
agcaccgaag cgaccccgag caacaacgaa agcgcgcccgc agagcaccga tgcgagcaac	540

aaagatgtgg tgaatcaggc cgttaatacc agcgcgcgcg gtatgcgtgc ctttagcctg	600
gcggccgtgg ccgccgatgc tccagcagca ggtaccgata ttaccaacca gctgaccaac	660
gtgaccgtgg gcattgatag cggcaccacc gtgtatccgc atcaggcagg ttatgtgaaa	720
ctgaactatg gcttttagcgt gccgaacagc gcggtgaaag gcgatacctt taaaattacc	780
gtgccgaaag aactgaacct gaacggcgtg accagcaaag atcaaaatag aactaaagcg	840
aaagtgcgcg cgattatggc aggtgatcag gtgctggcga acggcgtgat tgatagcgat	900
ggcaacgtga ttatatacctt taccgattat gtgaacacca aagatgatgt gaaagcgacc	960
ctgaccatgc cagcatatat tgatccggaa aacgtgaaga aaaccggcaa cgtgaccctg	1020
gcgaccggca ttggcagcac caccgcgaac aaaaccgttc tgggtggatta tgaaaaatac	1080
ggcaaattct acaacctgag catcaaaggc accatcgatc agatcgataa aaccaacaac	1140
acctatcgtc agaccattta tgtgaaccgc agcggcgata acgtgattgc gccggtgctg	1200
accggcaacc tgaaaccgaa caccgatagc aacgcgctga ttgatcagca gaacaccagc	1260
attaaagtgt ataaagtgga taacgcggcg gatctgagcg aaagctattt tgtgaaccgc	1320
gaaaactttg aagatgtgac caacagcgtg aacattacct ttccgaacct gaaccagtat	1380
aaagtggaat ttaacacccc ggatgatcag attaccaccc cgtatattgt ggtggtgaac	1440
ggccatattg atccgaacag caaaggcgat ctggcgctgc gtagcaccct gtatggctat	1500
aacagcaaca ttatttggcg tagcatgagc tgggataacg aagtggcggt taacaacggc	1560
agcggcagcg gcgatggcat tgataaaccg gtggtgccgg aacagccgga tgaaccgggc	1620
gaaattgaac cgattccgga agatggcagc catcatcatc atcatcatta agtcgac	1677

<210> 11

<211> 1677

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 11

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcc	60
gcgagcgaaa acagcgtgac ccagagcgat agcgcgagca acgaaagcaa aagcaacgat	120
agcagcagcg tgagcgcggc gccgaaaacc gatgatacca acgtgagcga taccaaaacc	180
agcagcaaca ccaacaacgg cgaaaccagc gtggcgcgaga acccggcgca gcaggaaacc	240
acccagagca gcagcaccaa cgcgaccacc gaagaaaccc cggtgaccgg tgaagccacc	300
accaccacca ccaaccaggc caacaccccg gcgaccaccc agagcagcaa caccaacgcg	360
gaagaactgg tgaaccagac cagcaacgaa accaccttta acgataccaa caccgtgagc	420
agcgtgaaca gcccgcagaa cagcaccaac gcggaaaacg tgagcaccac ccaggatacc	480
agcaccgaag cgaccccgag caacaacgaa agcgcgccgc agagcaccga tgcgagcaac	540
aaagatgtgg tgaatcaggc cgttaatacc agcgcgccgc gtatgcgtgc ctttagcctg	600
gcggccgtgg ccgccgatgc tccagcagca ggtaccgata ttaccaacca gctgaccaac	660
gtgaccgtgg gcattgatag cggcaccacc gtgtatccgc atcaggcagg ttatgtgaaa	720
ctgaactatg gcttttagcgt gccgaacagc gcggtgaaag gcgatacctt taaaattacc	780
gtgccgaaag aactgaacct gaacggcgtg accagcaccg cgaaagtgcc gccgattatg	840
gcagggtgatc aggtgctggc gaacggcgtg attgatagcg atggcaacgt gatttatacc	900
tttaccgatt atgtgaacac caaagataaa gatcaaaata gaactaaagt gaaagcgacc	960
ctgaccatgc cagcatatat tgatccggaa aacgtgaaga aaaccggcaa cgtgaccctg	1020
gcgaccggca ttggcagcac caccgcgaac aaaaccgttc tgggtggatta tgaaaaatac	1080
ggcaaattct acaacctgag catcaaaggc accatcgatc agatcgataa aaccaacaac	1140
acctatcgtc agaccattta tgtgaacccg agcggcgata acgtgattgc gccggtgctg	1200
accggcaacc tgaaaccgaa caccgatagc aacgcgctga ttgatcagca gaacaccagc	1260
attaaagtgt ataaagtgga taacgcggcg gatctgagcg aaagctatth tgtgaacccg	1320
gaaaactttg aagatgtgac caacagcgtg aacattacct ttccgaaccc gaaccagtat	1380
aaagtggaat ttaacacccc ggatgatcag attaccaccc cgtatattgt ggtggtgaac	1440
ggccatattg atccgaacag caaaggcgat ctggcgctgc gtagcaccct gtatggctat	1500
aacagcaaca ttatttggcg tagcatgagc tgggataacg aagtggcgtt taacaacggc	1560
agcggcagcg gcgatggcat tgataaaccg gtggtgccgg aacagccgga tgaaccgggc	1620
gaaattgaac cgattccgga agatggcagc catcatcatc atcatcatta agtcgac	1677

<210> 12

<211> 1677

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 12

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcc	60
gcgagcgaaa acagcgtgac ccagagcgat agcgcgagca acgaaagcaa aagcaacgat	120
agcagcagcg tgagcgcggc gccgaaaacc gatgatacca acgtgagcga taccaaaacc	180
agcagcaaca ccaacaacgg cgaaaccagc gtggcgcgaga acccggcgca gcaggaaacc	240
acccagagca gcagcaccaa cgcgaccacc gaagaaaccc cggtgaccgg tgaagccacc	300
accaccacca ccaaccaggc caacaccccc gcgaccaccc agagcagcaa caccaacgcg	360
gaagaactgg tgaaccagac cagcaacgaa accaccttta acgataccaa caccgtgagc	420
agcgtgaaca gcccgcagaa cagcaccaac gcggaaaacg tgagcaccac ccaggatacc	480
agcaccgaag cgaccccgag caacaacgaa agcgcgccgc agagcaccga tgcgagcaac	540
aaagatgtgg tgaatcaggc cgттаатacc agcgcgccgc gtatgcgtgc ctttagcctg	600
gcggccgtgg ccgccgatgc tccagcagca ggtaccgata ttaccaacca gctgaccaac	660
gtgaccgtgg gcattgatag cggcaccacc gtgtatccgc atcaggcagg ttatgtgaaa	720
ctgaactatg gcttttagcgt gccgaacagc gcggtgaaag gcgatacctt taaaattacc	780
gtgccgaaag aactgaacct gaacggcgtg accagcaccg cgaaagtgcc gccgattatg	840
gcaggtgatc aggtgctggc gaacggcgtg attgatagcg atggcaacgt gatttatacc	900
tttaccgatt atgtgaacac caaagataaa gatcaaaata gaactaaagt gaaagcgacc	960
ctgaccatgc cagcatatat tgatccggaa aacgtgaaga aaaccggcaa cgtgaccctg	1020
gcgaccggca ttggcagcac caccgcgaac aaaaccgttc tgggtggatta tgaaaaatac	1080

ggcaaattct acaacctgag catcaaaggc accatcgatc agatcgataa aaccaacaac	1140
acctatcgtc agaccattta tgtgaacccg agcggcgata acgtgattgc gccggtgctg	1200
accggcaacc tgaaaccgaa caccgatagc aacgcgctga ttgatcagca gaacaccagc	1260
attaaagtgt ataaagtgga taacgcggcg gatctgagcg aaagctattt tgtgaacccg	1320
gaaaactttg aagatgtgac caacagcgtg aacattacct ttccgaacc gaaccagtat	1380
aaagtggaat ttaacacccc ggatgatcag attaccaccc cgtatattgt ggtggtgaac	1440
ggccatattg atccgaacag caaaggcgat ctggcgctgc gtagcaccct gtatggctat	1500
aacagcaaca ttatttggcg tagcatgagc tgggataacg aagtggcggt taacaacggc	1560
agcggcagcg gcgatggcat tgataaaccc gtggtgccgg aacagccgga tgaaccgggc	1620
gaaattgaac cgattccgga agatggcagc catcatcatc atcatcatta agtcgac	1677

<210> 13

<211> 643

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polipéptido sintético"

<400> 13

Met Lys Lys Ile Trp Leu Ala Leu Ala Gly Leu Val Leu Ala Phe Ser
1 5 10 15

Ala Ser Ala Ala Glu Glu Ala Phe Asp Leu Trp Asn Glu Cys Ala Lys
20 25 30

Ala Cys Val Leu Asp Leu Lys Asp Gly Val Arg Ser Ser Arg Met Ser
35 40 45

Val Asp Pro Ala Ile Ala Asp Thr Asn Gly Gln Gly Val Leu His Tyr
50 55 60

Ser Met Val Leu Glu Gly Gly Asn Asp Ala Leu Lys Leu Ala Ile Asp
65 70 75 80

Asn Ala Leu Ser Ile Thr Ser Asp Gly Leu Thr Ile Arg Leu Glu Gly
85 90 95

Gly Val Glu Pro Asn Lys Pro Val Arg Tyr Ser Tyr Thr Arg Gln Ala
100 105 110

Arg Gly Ser Trp Ser Leu Asn Trp Leu Val Pro Ile Gly His Glu Lys
115 120 125

Pro Ser Asn Ile Lys Val Phe Ile His Glu Leu Asn Ala Gly Asn Gln
130 135 140

Leu Ser His Met Ser Pro Ile Tyr Thr Ile Glu Met Gly Asp Glu Leu
145 150 155 160

Leu Ala Lys Leu Ala Arg Asp Ala Thr Phe Phe Val Arg Ala His Glu
165 170 175

Ser Asn Glu Met Gln Pro Thr Leu Ala Ile Ser His Ala Gly Val Ser
180 185 190

Val Val Met Ala Gln Ala Gln Pro Arg Arg Glu Lys Arg Trp Ser Glu
195 200 205

Trp Ala Ser Gly Lys Val Leu Cys Leu Leu Asp Pro Leu Asp Gly Val
 210 215 220

Tyr Asn Tyr Leu Ala Gln Gln Arg Cys Asn Leu Asp Asp Thr Trp Glu
 225 230 235 240

Gly Lys Ile Tyr Arg Val Leu Ala Gly Asn Pro Ala Lys His Asp Leu
 245 250 255

Asp Ile Lys Asp Asn Asn Asn Ser Thr Pro Thr Val Ile Ser His Arg
 260 265 270

Leu His Phe Pro Glu Gly Gly Ser Leu Ala Ala Leu Thr Ala His Gln
 275 280 285

Ala Cys His Leu Pro Leu Glu Ala Phe Thr Arg His Arg Gln Pro Arg
 290 295 300

Gly Trp Glu Gln Leu Glu Gln Cys Gly Tyr Pro Val Gln Arg Leu Val
 305 310 315 320

Ala Leu Tyr Leu Ala Ala Arg Leu Ser Trp Asn Gln Val Asp Gln Val
 325 330 335

Ile Arg Asn Ala Leu Ala Ser Pro Gly Ser Gly Gly Asp Leu Gly Glu
 340 345 350

Ala Ile Arg Glu Gln Pro Glu Gln Ala Arg Leu Ala Leu Thr Leu Ala
 355 360 365

Ala Ala Glu Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn Asp Glu
 370 375 380

Ala Gly Ala Ala Ser Ala Asp Val Val Ser Leu Thr Cys Pro Val Ala
 385 390 395 400

Lys Asp Gln Asn Arg Thr Lys Gly Glu Cys Ala Gly Pro Ala Asp Ser
 405 410 415

Gly Asp Ala Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala Glu Phe Leu
 420 425 430

Gly Asp Gly Gly Asp Val Ser Phe Ser Thr Arg Gly Thr Gln Asn Trp
 435 440 445

Thr Val Glu Arg Leu Leu Gln Ala His Arg Gln Leu Glu Glu Arg Gly
 450 455 460

Tyr Val Phe Val Gly Tyr His Gly Thr Phe Leu Glu Ala Ala Gln Ser
465 470 475 480

Ile Val Phe Gly Gly Val Arg Ala Arg Ser Gln Asp Leu Asp Ala Ile
485 490 495

Trp Arg Gly Phe Tyr Ile Ala Gly Asp Pro Ala Leu Ala Tyr Gly Tyr
500 505 510

Ala Gln Asp Gln Glu Pro Asp Ala Arg Gly Arg Ile Arg Asn Gly Ala
515 520 525

Leu Leu Arg Val Tyr Val Pro Arg Trp Ser Leu Pro Gly Phe Tyr Arg
530 535 540

Thr Gly Leu Thr Leu Ala Ala Pro Glu Ala Ala Gly Glu Val Glu Arg
545 550 555 560

Leu Ile Gly His Pro Leu Pro Leu Arg Leu Asp Ala Ile Thr Gly Pro
565 570 575

Glu Glu Glu Gly Gly Arg Val Thr Ile Leu Gly Trp Pro Leu Ala Glu
580 585 590

Arg Thr Val Val Ile Pro Ser Ala Ile Pro Thr Asp Pro Arg Asn Val
595 600 605

Gly Gly Asp Leu Asp Pro Ser Ser Ile Pro Asp Lys Glu Gln Ala Ile
610 615 620

Ser Ala Leu Pro Asp Tyr Ala Ser Gln Pro Gly Lys Pro Pro Arg Glu
625 630 635 640

Asp Leu Lys

<210> 14

<211> 624

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polipéptido sintético"

<400> 14

Ala	Glu	Glu	Ala	Phe	Asp	Leu	Trp	Asn	Glu	Cys	Ala	Lys	Ala	Cys	Val
1				5					10					15	
Leu	Asp	Leu	Lys	Asp	Gly	Val	Arg	Ser	Ser	Arg	Met	Ser	Val	Asp	Pro
		20						25					30		
Ala	Ile	Ala	Asp	Thr	Asn	Gly	Gln	Gly	Val	Leu	His	Tyr	Ser	Met	Val
		35					40					45			
Leu	Glu	Gly	Gly	Asn	Asp	Ala	Leu	Lys	Leu	Ala	Ile	Asp	Asn	Ala	Leu
	50					55					60				
Ser	Ile	Thr	Ser	Asp	Gly	Leu	Thr	Ile	Arg	Leu	Glu	Gly	Gly	Val	Glu
65					70					75					80
Pro	Asn	Lys	Pro	Val	Arg	Tyr	Ser	Tyr	Thr	Arg	Gln	Ala	Arg	Gly	Ser
				85					90					95	
Trp	Ser	Leu	Asn	Trp	Leu	Val	Pro	Ile	Gly	His	Glu	Lys	Pro	Ser	Asn
			100					105						110	
Ile	Lys	Val	Phe	Ile	His	Glu	Leu	Asn	Ala	Gly	Asn	Gln	Leu	Ser	His
		115					120					125			
Met	Ser	Pro	Ile	Tyr	Thr	Ile	Glu	Met	Gly	Asp	Glu	Leu	Leu	Ala	Lys
	130						135				140				
Leu	Ala	Arg	Asp	Ala	Thr	Phe	Phe	Val	Arg	Ala	His	Glu	Ser	Asn	Glu
145					150					155					160
Met	Gln	Pro	Thr	Leu	Ala	Ile	Ser	His	Ala	Gly	Val	Ser	Val	Val	Met
				165					170					175	
Ala	Gln	Ala	Gln	Pro	Arg	Arg	Glu	Lys	Arg	Trp	Ser	Glu	Trp	Ala	Ser
			180					185					190		

Gly Lys Val Leu Cys Leu Leu Asp Pro Leu Asp Gly Val Tyr Asn Tyr
 195 200 205

Leu Ala Gln Gln Arg Cys Asn Leu Asp Asp Thr Trp Glu Gly Lys Ile
 210 215 220

Tyr Arg Val Leu Ala Gly Asn Pro Ala Lys His Asp Leu Asp Ile Lys
 225 230 235 240

Asp Asn Asn Asn Ser Thr Pro Thr Val Ile Ser His Arg Leu His Phe
 245 250 255

Pro Glu Gly Gly Ser Leu Ala Ala Leu Thr Ala His Gln Ala Cys His
 260 265 270

Leu Pro Leu Glu Ala Phe Thr Arg His Arg Gln Pro Arg Gly Trp Glu
 275 280 285

Gln Leu Glu Gln Cys Gly Tyr Pro Val Gln Arg Leu Val Ala Leu Tyr
 290 295 300

Leu Ala Ala Arg Leu Ser Trp Asn Gln Val Asp Gln Val Ile Arg Asn
 305 310 315 320

Ala Leu Ala Ser Pro Gly Ser Gly Gly Asp Leu Gly Glu Ala Ile Arg
 325 330 335

Glu Gln Pro Glu Gln Ala Arg Leu Ala Leu Thr Leu Ala Ala Ala Glu
 340 345 350

Ser Glu Arg Phe Val Arg Gln Gly Thr Gly Asn Asp Glu Ala Gly Ala
 355 360 365

Ala Ser Ala Asp Val Val Ser Leu Thr Cys Pro Val Ala Lys Asp Gln
 370 375 380

Asn Arg Thr Lys Gly Glu Cys Ala Gly Pro Ala Asp Ser Gly Asp Ala
 385 390 395 400

Leu Leu Glu Arg Asn Tyr Pro Thr Gly Ala Glu Phe Leu Gly Asp Gly
 405 410 415

Gly Asp Val Ser Phe Ser Thr Arg Gly Thr Gln Asn Trp Thr Val Glu
420 425 430

Arg Leu Leu Gln Ala His Arg Gln Leu Glu Glu Arg Gly Tyr Val Phe
435 440 445

Val Gly Tyr His Gly Thr Phe Leu Glu Ala Ala Gln Ser Ile Val Phe
450 455 460

Gly Gly Val Arg Ala Arg Ser Gln Asp Leu Asp Ala Ile Trp Arg Gly
465 470 475 480

Phe Tyr Ile Ala Gly Asp Pro Ala Leu Ala Tyr Gly Tyr Ala Gln Asp
485 490 495

Gln Glu Pro Asp Ala Arg Gly Arg Ile Arg Asn Gly Ala Leu Leu Arg
500 505 510

Val Tyr Val Pro Arg Trp Ser Leu Pro Gly Phe Tyr Arg Thr Gly Leu
515 520 525

Thr Leu Ala Ala Pro Glu Ala Ala Gly Glu Val Glu Arg Leu Ile Gly
530 535 540

His Pro Leu Pro Leu Arg Leu Asp Ala Ile Thr Gly Pro Glu Glu Glu
545 550 555 560

Gly Gly Arg Val Thr Ile Leu Gly Trp Pro Leu Ala Glu Arg Thr Val
565 570 575

Val Ile Pro Ser Ala Ile Pro Thr Asp Pro Arg Asn Val Gly Gly Asp
580 585 590

Leu Asp Pro Ser Ser Ile Pro Asp Lys Glu Gln Ala Ile Ser Ala Leu
595 600 605

Pro Asp Tyr Ala Ser Gln Pro Gly Lys Pro Pro Arg Glu Asp Leu Lys
610 615 620

<210> 15

<211> 1148

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 15

taatgaaata cctgctgccg accgctgctg ctggtctgct gctcctcgct gccagccgg	60
cgatggccat gcatatgagc aaagaagaag caccaaaaat acaaatgccg cctcaacctg	120
taacaacccat gagtgctaaa tctgaagatt taccacttag ttttacttac cctgctaaac	180
ttgtcagtga ttatgatgtc attataaaac ctcaagttag cggcgtaata gtaaataaac	240
tttttaaagc tggagataag gtaaaaaaag gacaaacatt atttattata gaacaagata	300
aatttaaagc tagtgttgat tcagcttacg gacaagcttt aatggctaag gcaactttcg	360
aaaatgcaag caaggatttt aatcgttcta aagctctttt tagcaaaagt gcaatctctc	420
aaaaagaata cgactcttct cttgctacat ttaacaattc aaaagctagt ctagcaagtg	480
ctagagcaca gcttgcaaat gcaagaattg atctagatca taccgagata aaagctcctt	540
ttgatggtac tataggagat gctttagtta atataggaga ttatgtaagt gcttcaacaa	600
ctgaactagt tagagttaca aatttaaata ctatttacgc agatttcttt atttcagata	660
cagataaaact aaatttagtc cgcaatactc aaagtggaaa atgggattta gacagcattc	720
atgcaaattt aaatcttaat ggagaaaccg ttcaaggcaa actttatttt attgattcgg	780
ttatagatgc taatagtgga acagtaaaag ccaaagccgt atttgataac aataactcaa	840
cacttttacc gggtgctttt gcaacaatta cttcagaagg ttttatacaa aaaaatggct	900
ttaaagtgcc tcaaataggt gttaaacaag atcaaaatga tgtttatgtt cttcttgttt	960
aaaatggaaa agtagaaaaa tcttctgtac atataagcta ccaaaacaat gaatacgcca	1020
ttattgacaa aggattgcaa aatggcgata aaatcatttt agataacttt aaaaaattc	1080
aagttggtag cgaagttaaa gaaattggag cacaactcga gcaccaccac caccaccact	1140
gagtcgac	1148

<210> 16

<211> 990

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 16

catatgaaga aaatctggct ggcgctggcg ggcctggtgc tggcgtttag cgctagcgcg	60
gcggatagcg atattaacat taaaaccggc accaccgata ttggcagcaa caccaccgtg	120
aaaaccggcg atctggtgac ctatgataaa gaaaacggca tgctgaaaaa agtggtttat	180
agcttcatcg atgataaaaa ccataacaaa aaactgctgg tgatccgtac caaaggcacc	240
attgcggggc agtatcgtgt gtatagcgaa gaaggcgca acaaaagcgg cctggcggtg	300
ccgagcgcggt ttaaagtgca gctgcagctg ccggataacg aagtggcgca gattagcgat	360
tattatccgc gtaacagcat tgataccaaa gaatatatga gcaccctgac ctatggcttt	420
aacggcaacg tgaccggcga tgataccggc aaagatcaaa atagaactaa aattggtgga	480
ctgattggcg cgaacgtgag cattggccat accctgaaat atgtgcagcc ggattttaaa	540
accattctgg aaagcccgac cgataaaaaa gtgggctgga aagtgatttt taacaacatg	600
gtgaaccaga actggggccc gtatgatcgt gatagctgga acccgggtga tggcaaccag	660
ctgtttatga aaaccgtaa cggcagcatg aaagcggcgg ataactttct ggatccgaac	720
aaagcgagca gcctgctgag cagcggcttt agcccggtt ttgcgaccgt gattaccatg	780
gatcgtaaag cgagcaaaac gcagaccaac attgatgtga tttatgaacg tgtgcgtgat	840
gattatcagc tgcattggac cagcaccaac tggaaaggca ccaacaccaa agataaatgg	900
attgatcgta gcagcgaacg ttataaaatt gattgggaaa aagaagaaat gaccaacggc	960
agccatcatc atcatcatca ttaagtcgac	990

<210> 17

<211> 19442

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 17

gaattccctg aggcaattct tctttgatga cggctgatgg tgaggttgac ctggtgaagc	60
tggtcaagga gctttgggtt aacaagggtc tgattcttct gactactctt cttgcattaa	120
tcgggtcttt tacctatgog tatctgagta agcctgtata tgaatatagg gttgcagtag	180
tgccctctgc tcttgggtct atcgaagggt tcaatgttgg tagaaggag aatggcctag	240
atgcatatac tgttagaagt atctatgcga tcttttcgcg caatctgctt tcggatgaga	300
ataaaaaaga gttcttctat aagatatacc ttccccaggt gggtgaggga gcggaaagcg	360
aagatgagca ggaggagttt tataagaagt tctccaaaga ggtaaagatt gatcctgcta	420
acaagccaga tgcagaacct tatacggtaa ttgtggaggg cacgaagcga gaggttcttg	480
ctacatgggc acaagctttc gtgcgttttg ctgcggatcg ggccgtgcat gaggttattg	540
atagtgcagg tagagatttc caggtaagaa atgctgcaat gcagagccgc ataaccgtgc	600
tgcagaatat ggcgaagggc cgccgtgatg atagaattgc acgtttgaag gaggcattgc	660
tgattgcgga gtcgctcaag atagatggcc cgccattaat agaaggggcg tccgagcaac	720
aactctcctc gatcatggat ggtgacttga tgtacatgcg aggagctaag gcgctgcgcg	780
ctgaaatcaa caatcttgag tcgcgtagtg tagatgctcc attcattcct gagttgagaa	840
ctctccaaga gaaactatct tggaactcca gtttgtctgt ggattctgat gcggtggctg	900
tctacaagga agacgaggga ctctcttttt caaatcaacc cattaagccg aagaagattc	960
ttatagttac tataggtact ttggcaggat tgataattgg aattctactc gcagtgctcg	1020
ctggttttat aaggaagctt cgtagcgatg gctctcttcg ctaagttttg atctaaaccc	1080

tgatgccact tactggcatc agggcttact tgttgtagtt gttgaattgt attagggggt	1140
atcgagtaga atgtcaataa ggcggggcgt tttctactcc ggtatatcca tgggcagcaa	1200
ctatcttttg cctcttgctg cgatttccttt tctcacgaga acattgtcaa gogaagcgtt	1260
tggccaattg gtgattgcc aggcgcgtggc tgtcattcta tgtcaactgg tagactttgg	1320
atattattctg gcaggatcaa gaaaggctgc cattatcgat aacaaagttg aactgtctag	1380
tttcttttct gttgtacaga gtgctagatt cttattgttg ctgctttcac ttttagtgct	1440
ggccattttg gctgtatctt ctatttttacc aatcccccttg cttgtattgg ttgcggctgc	1500
tcttcggga gtagttggaa attatcttca agcagtatgg ttcttttcagg gaagagcgt	1560
gtttggatgg ttggcgctta ccaatttttt gtctaaggta ttttatttcc tattggctgt	1620
tttttttgtc aogaaggatt ctgacctgt gctggcttcg ttgggggttg gtttttcta	1680
tgtcataggt ggaagtgtc tctgttgtat tttattttct atgggaatac ggtggcgccc	1740
ggttctcgag aaagacagaa ttctcgatat attgcgtgac ggtgctcgat cttttcttcc	1800
tctggctttt cttagcttgc acatgcaagt gctcgttgcg gcggttggtg ttggttggtg	1860
agcctccgcg gcgggagtg tttctactgc ggataaatcc cttcgcgga tcgcggtgc	1920
tacttcaccc atagctagcg ctctatttcc gacttttagc aggatgtatg cgagtgccga	1980
cccggcagtc ggcagtttaa gaaggaaagc gctaggtctg atgttaactaa tagctattcc	2040
tagttgttta tttcttttct tattttctga atacatttca tatctcctat tcccgaaca	2100
gtccagaggt ctaactgttg taataagaat gttttcgata gtgccagtgt ttgcttgtat	2160
tgggtgtctg tatggagggt tgactcttgt tccttctggg tatgatggtg tatatttgcg	2220
agcaattttt tttgcggaat tgggcggggg attaacattt atcctcttgg cgctttgggg	2280
ggatgagctt tttggagcgt ggacgcgtgt cgttacagag gtctcttttg ggatgggaat	2340
gtttttcctg gccacggtta agttgagaga gaaaagggga ctttgatctt aagatgagga	2400
tagcgattga aaagataatt ggtttgctga aaaaccagtc ctctaaagaa tcgaatgta	2460
agattcatcg cttggcgtat attacaaact caaaatttga tggcaataac tatatagata	2520
gatggtgtaa aatcaggaat tctcacattg gtgaatacag ttatattgga tttggtagtg	2580
attttaataa tgtagaagta ggaagatatt gttcgatata ttcggatgta aaaattgggt	2640
taggaaaaca tcctacacac ttttttagct catcacgat tttttattct aataataatc	2700

catttaacat aaagcaaaag tttatagact ttaatgacca accaagccgt acaacaatta	2760
aaaatgatgt gtggatttgt gcaaatgtaa ttattatgga tggattaaca ataaatactg	2820
gtgcagtcac agcagccggc tcagttgtta ctaaaaatgt aggagcatat gaggttggtg	2880
gtgggggttc tgcaaaagtg attaagaagc gatttgacaa taaaacaatt gaaaaacttt	2940
tggaaagcaa gtggtgggag aaaacgcctg acaactaaa aggattttcg gttgaatatt	3000
taaataaaaa ggatactagt tatccgtatg atgtgcgga ttatgcgtaa caggaggtgt	3060
acaatgagca ggaggtgtac aatgagaatt ttaaatattg tatcgagtaa tattgttcaa	3120
gacccaaggg tacttaaaaca aatagaaaca attaaaggcg ttacgaatga ttataaaatt	3180
gttggaatga ataattcaca agctactaat aggcgattgg aaaatttaga ttgtaattat	3240
cgtttggttag gtagcaaggt agatcccaa aatattcttt cttaaattaat taagcgtata	3300
agatttgcaa caggtgttat ccgagaaatt aaagctttta aacctgacgt gattcatgca	3360
aatgatttcg acgtattatt aatggtctat ttaagcaatt ataaaaaagc taatattgtt	3420
tatgatgogc atgaaatata tgcgaaaaat gcctttatta ataaagttcc acttatttca	3480
aagtttgtag aaagtataga aaaacacata gtaaaacatc gtgttaatgc cttcgtaaca	3540
gtaagtcacg cagcaaaaga atattatcaa tctaaaggat ataagaagga agcgaatgtt	3600
attacgaatg cacctatttt aaatgatagc agagaattta aagaaatcga aaacttttaa	3660
gaaatcgat atcaaggtca aattgtaatg gacagaggat atgaagagtt tattattgct	3720
tcatcagctt ttaaacaaaa tgctccttca ttcataattc gagggtttgg tccgcatgaa	3780
gaagtataa aagaactgat tagttataac tcggaaaata ttaggttgga taaaccagtt	3840
gaagtaaaag aattgggtga taagttagca gaaagtaatg ttggtgttat cttgacgaaa	3900
cctgtatcta ttaattttga atatacagta tctaataaaa tttttgaatg tatacatgct	3960
ggtttaccag taattttatc tctgtcaaaa gagcatattt atctcaatga aaaatataaa	4020
tttggcattg ttttaaagga agttacgccg ttagaaattg aaaaggcggg tagaaaatta	4080
agagataatc acgatttggt taatcattta cgtcaaaatg caattaaggc gtctaaaatt	4140
ttgaattggc aaatagaaag tgaacgatta gtagaattat ataaatttga acaaaaactc	4200
atctcagaag aggatctgta aggacccggg gatcctctag caggaggaac tatgaaattt	4260
tttgtacttt gtgcaattat cagcatgaac atatttatag taatctctac atttactaaa	4320

gaagtattag ggttccctat agagccggtg tattactcaa ccatggttgg tatagcatta	4380
attactacgg tgtttgctat ttataagata attgtcacgc aagaaattcc gcgagggtta	4440
atattattaa ttgctatatg tttgctttat ctagcttttt attatttttc accagataag	4500
gaagagaaac tagctaaaaa taatattcta ttctttttta catgggcagt tccagcggca	4560
attagtggta tttatattaa atatataaac aaggctacgg tagaaagatt ttttaaatta	4620
gtatttttca tattttctat ttcatttatt tttgtaattt taataccaaa acttacaggt	4680
gagataccta gctatatcaa ttttggaact atgaactatc aaaacgcttc gtacctttca	4740
gcatttactg ccggattagg catttatttc attatgaaag gttcagtga acataagtgg	4800
atatatgttc tatttacaat aattgatatc cctattgtgt ttataccagg agggcggtga	4860
ggtgctattt tattaattct ttacggctta tttgcattta tacttattac gtttaaaaga	4920
ggaataccta ttgcagtaaa aagcattatg tataatttttg cattaagcat atctagtgt	4980
ttgatttact ttctttttac aaaagggtcg aatactagaa cattttcata tctacaaggt	5040
ggaacactta atttagaagg tacttctgga agaggacoga tttatgaaaa aggtatttac	5100
tttattcaac aaagtccgtt attaggctat gggccattta actattataa actaatcgga	5160
aataataccac ataacatcat tattgagttg attctatcat ttggcttatt agggtttttt	5220
atcataatga tttgcatttt gctactagtt tataaaatga ttaggaacta tgatccaaac	5280
actatagatt tactcgttat gtttatagca atctatccaa tcacattatt aatgtttagt	5340
tcaaattatt tagttgtaag tgaatttttg tttgtgttgt tctattttat taaaaagga	5400
cggcgtcatc atggtgatta taaagatgat gatgataaat aagtttaaac aggaggcatc	5460
atggttaaga aagtttttat tatggatagc gtaaagacaa taattggtac gttgcttata	5520
gctttaggat tacaattttt agcttatcca attattaatc aacgagtagg taatgaagcg	5580
tttggttcta ttttaacgat ttatacaata ataacaatca cgagtgttgt attaggcaat	5640
acgcttaaca atatacgatt aattaatatg aatctataca aatccaatca ttactactgg	5700
aaatttgtgt cgatactttt aatttcaatt ctgattgaga gtatagcttt aattattgta	5760
tttctttact tttttaattt gaacaccatc gatattatct ttttaattct acttaattat	5820
ttaatgtgtt taaggattta tctgaatgta ttttttagga tgactttaaa atataatcag	5880
attttgtata ttgctcttat tcaattttta ggtttgctga taggactatt tctatattat	5940
ttaatccaaa actggattgt ttgttttatt accagtgaat tgtttgcaac gatatataca	6000

ttgggttaaat tacgggggatt aactataggc gagtatcaaa gtgaagataa taatgtgggc	6060
aaagattatg tgatgctact gagtacaat agccttaata atttgaatct ctacttagat	6120
agattaatct tattaccaat tataggtgga acagctgtaa ctatatcatt tctttcaaca	6180
tttattggga aaatgttagc tacatctctg tatccgatta ataatgtagt actttcatat	6240
atttctgtaa atgaaagcga caatataaag aagcaatatt tgaaaactaa tctatttgct	6300
atagctgcac tatgttttagt catgattata tgttatccaa ttacattaat tattgtctct	6360
ttactgtata acattgattc aagtttatat tcgaagttta ttatttttagg taatataggt	6420
gttttattca atgcagtgag tattatgatc caaactttta atacaaaaca cgcacataa	6480
acattacaag cgaattatat gacgcttcac acgattacat ttatatccat aactatttta	6540
atgacaattg cgtttggtct aaatggattc ttttgacaa cgctgttcag caacattatt	6600
aagtatgtga ttttaaataat tataggttta aagtctaaat tcattaataa aaaggacgct	6660
gatagttatc cgtatgatgt gccggattat gcgtaagtgc accatatgaa tatcctcctt	6720
agttcctatt ccgaagttcc tattctctag aaagtatagg aacttcggcg cgcctacctg	6780
tgacggaaga tcacttcgca gaataaataa atcctggtgt cctctgtgat accgggaagc	6840
cctgggccaa cttttggcga aaatgagacg ttgatcgga cgtaagaggt tccaactttc	6900
accataatga aataagatca ctaccgggcg tattttttga gttgtcgaga ttttcaggag	6960
ctaaggaagc taaaatggag aaaaaaatca ctggatatac caccgttgat atatcccaat	7020
ggcatcgtaa agaacatttt gaggcatttc agtcagttgc tcaatgtacc tataaccaga	7080
cogttcagct ggatattacg gcctttttta agaccgtaa gaaaaataag cacaagtttt	7140
atccggcctt tattcacatt cttgcccgcc tgatgaatgc tcatccggaa ttacgtatgg	7200
caatgaaaga cggtgagctg gtgatatggg atagtgttca cccttgttac accgttttcc	7260
atgagcaaac tgaaacgttt tcatcgctct ggagtgaata ccacgacgat ttccggcagt	7320
ttctacacat atattcgcaa gatgtggcgt gttacggtga aaacctggcc tatttcctta	7380
aagggtttat tgagaatatg tttttcgtct cagccaatcc ctgggtgagt ttcaccagtt	7440
ttgatttaaa cgtggccaat atggacaact tcttcgcccc cgttttcacc atgggcaaat	7500
attatacgca aggcgacaag gtgctgatgc cgctggcgat tcaggttcat catgccgttt	7560
gtgatggctt ccatgtcggc agatgcttaa tgaatacaac agtaactgca tgagtggcag	7620

ggcggggcgt aaggcgcgcc atttaaataga agttcctatt ccgaagttcc tattctctag	7680
aaagtatagg aacttcgaag cagctccagc ctacacgtaa agagggtggg tatggataag	7740
aactctgttc tgttaattac cgggtggaact ggttcctttg gaaatgccgt tttgaagcgt	7800
tttctggata cagatattgg tgaaatacgt gtattcagtc gggatgagaa gaagcaagat	7860
gatatgcgta agtgctatgc tcaccccaag ttgaggtttt atattggaga tgtccgcgac	7920
tatcaaagca ctctgaatgc tactcgcggt gtggactaca tttccatgc ggctgcattg	7980
aaacaggttc cctcttgtga gttctaaccg atggaggctg tgaagaccaa tgtcatcgga	8040
acggaaaatg tcctcgaatc tgctatccag aatggcgtca aaaaagtcgt ttgtctgagt	8100
acggataagg cagtttacc aatcaatgcc atgggtattt caaaggccat gatggaaaag	8160
gtcatggctg ccaagtctcg aaacctggaa cgcaactcta ccgtaatttg tggcaccgt	8220
tatggtaacg tcatggcttc aaggggctcg gtcattcctc tctttatcga gcaaatgcga	8280
tcaggccagc ctctcacaat cactgatcca aacatgacac gttcatgat gacgcttacg	8340
gatgccgtag accttgtgct ttatgctttc gagcatggta ccaatggaga tctttttgta	8400
cagaaagcac cggctgcaac catcgagggtg ctggctcatg cgcttactca attgcttggc	8460
aagaatgggt atcctatcaa tgtaataggc acgcgtcatg gagagaagct ttatgaggcg	8520
cttcttagtc gagaggaaat ggctgtgct gaagacatgg gtgactatta tcgtatcccc	8580
ccagatttgc gtgatctgaa ctatagtaag tttgtggagc aaggcgagga aaaaattact	8640
catacagagg actataattc tcataatacc aaacgtctgg atatcgaagg aatgaaaaag	8700
ctgttgctga agttggattt cattcgtgct attcagcgtg gcgagagtgc cagtccagag	8760
gaataaacga tgaaagttct tgtaactggc gcgaatggat ttgttggaag gaatctgtgc	8820
gctcatcttg cagagcgggg tggatocgag gtggtgccat tcacccgcga gagtagtgtt	8880
ggtaatttgc ctgagctaata tcgttccgtc gattttattt ttcattctgc cggggccaat	8940
cgtccggaaa aaccagaaga gtttaagatc gggaattccg aactcacgta tgctctgtgt	9000
gaggcggtaa ggtccaatgg acgagccata ccacttcttt atacttcac cattcaggct	9060
gagggtgata atgagtacgg ttttaagcaag cgagccgcag aagagcatct ccaagtgcta	9120
ggtgaggata ttggttgtcc tgtctacata ttctgccttc ctaatgtatt cggtaaattgg	9180
tcgcgtccga attataattc agcggttgcg actttttgtc ataataattat tcgagatatt	9240
ccgattcaaa ttaacaattc ctggcagag atcactcttg tatacataga tgatgtggtt	9300
cgcaccttca tgaaagtcac ggatgggaag ctatccaatg cagtttcact acaggctcag	9360

ccccagtatc agattttctgt tgggtgagctc gcagaacaat tgtatgagtt tcgtaatagt	9420
cgaaagtcac tgactaccgc aagggttggc tcgggattga cgcgcgcctt gtactcgact	9480
tatctaagtt tcttgccaga agatagtttt agttacgacg tgccaatgca ttcggatccg	9540
cgtggcacat tcgtcgagat gctgaagacc gcggactctg gccagttctc gttttttacg	9600
gctcatccag gtgttaccag gggcgggcat taccatcact cgaaaaccga aaagtttctg	9660
gttatcaaag ggatggcacg tttcaagttt agaaacatcc tgaccggggc atttttacgaa	9720
atttgacta atggtgaaaa ggcagaaatt gtcgaaacag tacctggatg gactcatgac	9780
attactaatg tcggaactga cgatatggta gtcatgttgt gggctaacga agtatttgat	9840
cgggaaaatc cggataccta cgcttgttca gtaggcgaag gtgcgtaagg tatagtgaga	9900
taacaatgca gaagctaaaa gtcgttacgg ttgttggaac tcgtcctgag attattcgct	9960
tgtctagggt catggcgaag cttgatcagt actgcgatca tgtacttgtc catactggac	10020
agaattatga ttacgaactt aatgaaatat tttttcagga cctcgggtata agaaagccgg	10080
attattttct aaacgcgcgc gggctcttcg gggctgaaac gatagggaat gtaataatcg	10140
cagtogatcg tgttctgggc gaaatagatc ccgatgcgct gctcgtgctg ggtgatacca	10200
atagttgtat ggcgggtactg cctgcaaaac ggcgtaagat accgaccttt catatggaag	10260
caggcaatcg ctgtttcgat atgcgtgtgc ctgaagagat aaatcggcgc attgtcgatc	10320
atacagctga tgtaaatttg acctatagta caattgcgcg tgattatctc ttgcgtgaag	10380
gactttctcc agacatggtt atcaagactg gtagccctat gttcgaagtt ctcgagcact	10440
atcgtgacgg gatcgagtcc tccgatattc ttgaaaggct cgggttgaaa acagagcgg	10500
tctttgtcgt gagtgcgcac cgagaggaaa acatagattc ggataagaat ttcttgaagt	10560
tggttttctat gctcaacgct gtggcagaaa agtactcgct gcccgtcacg gtatcaactc	10620
accctagaac aaaaaagaga attgaggcga cggaggcaaa gtttcacgag ggtattaaac	10680
tgctgaaacc cctcggcttt aaggattaca ataaactgca aattacagcc aaggcagtta	10740
tttctgacag tgggaccatc agtgaggagt cttcaatact gaattttccc gctttgaata	10800
ttcgtgaggc tcatgaacgc ccagaaggca tggaagaggc tgtggtgatg atggtcggac	10860
tggattcgga tcgagtacta caagcactcg aggtgttgga gggacagagg cgcgacgcag	10920
agcgcagtgt acgcttggtc gctgactata gcatgcccaa cgtttctgaa aagattgttc	10980
gcatagttca tagctatcgg gactatgtca tgcgaactgt ctggaaaaaa tattaacttg	11040

aggcgtggag	ttgatggcaa	ggatatttgt	ggtttctgag	tatgtcgggtg	ccaatcagaa	11100
ctccacggga	tactattggg	agaagataat	aggaaagatg	cagcgggagt	ttggtgggct	11160
aaccgtaatt	ttcccgtga	ccgcaggtga	aaccccgctt	gtggtttcac	cttccgttga	11220
gcaagaatgc	tttaagtttc	cgaggagcaa	taagaatagg	ctcctttcta	gaggattggc	11280
gcagatTTTT	caggcgTTTT	tgttctcagt	aaaattgact	tctcgtgcc	gacgaggaga	11340
tgtggtattg	agtggaacca	accctgctct	tctactgatg	acgtttccct	tgctaaggta	11400
tgccctcggg	ttcaagtggg	tgctgctggg	gcatgatgtg	tttcccgaga	acttggtgcc	11460
ggcgggctgt	ctgaagaaag	atagtattgc	ctaccggctt	ctacgtcgtc	tcttttcttt	11520
catttactca	tccgctgatc	gtctagtcgt	aatagggcgc	gatatggaag	ctcttatgaa	11580
agagaagggtg	aatgacccgc	gatctttggg	ctttatttcg	aattgggcct	gtgagaaaga	11640
ggttttccca	gtaccgagag	aggatgctcc	ttttatcaat	attcctgaat	ggaaaggtaa	11700
aagggttttc	caattttttg	gtaatgtcgg	tcgattacaa	ggtatagaaa	acatactttc	11760
tgctattcag	ttggttaaaa	acgagaaggc	ggcttttgct	tttattggag	atggtgcctt	11820
ggtcgacagt	gtaaaaaaac	acgcgctgga	agatcagtgt	gctcggttga	ggtattttgg	11880
aaggctgcc	ttagccgaaa	agaattttgg	tttggtgcc	tgtgacgttg	ccttagttac	11940
cttagaagaa	ggaatgttcg	ggcttggggt	tcccagcaag	gcatatttct	ccatggcagc	12000
agacaaaccg	attctagctg	tcatggaaaa	aggggctgaa	atctcccgt	taatagatga	12060
gaccggaatc	ggttggaact	gtccgccgaa	tgatccgggt	gctttggcaa	gattgatcga	12120
tgagatttgt	gaactcgact	tgtctagttt	aggcggagtc	ccgcggagtg	tccttcagca	12180
aaattattct	gaatatattt	cattggaaaa	attcgtgcc	tgtgttcgac	cgcttctgtc	12240
tgagtgcgaa	atatgatgaa	ggtgctggta	accggggcta	gcggttttgt	cgggagtgcg	12300
ctttgcagg	cgcttgctgc	cgcccccttt	caggttgtcg	gacaagtacg	atccctgtac	12360
aatcccgcta	cggggggtga	gtatgttcga	gcggagctga	aagagagcac	taagcttgat	12420
gctgcgctgc	gggggtgtga	atgtgtagtt	catctagctg	gacgagccca	tatctttgga	12480
aggcagcgtg	attcactaga	tatttttctg	aagggtgaatc	gcgatgctac	tctggcgtt	12540
gctcggcagg	cgatcgaagc	atctgtaaa	cgtttcattt	ttgttagttc	tattggtgta	12600
aatggcgctt	taaccaaaga	aaagcccttc	gatgagaact	ccaagccggc	tcctcatgca	12660
gaatatgcga	tttcaaagtt	tgaggctgaa	gtagcgcttc	gggagctttt	caagcattcc	12720
tcaacagaa	ttgttatcgt	caggccctca	ctcgtttacg	actggaaagc	tcctggaaat	12780
ttctcgcgat	tgttgaagct	ggttgcttcg	ggacttcctc	ttccatttgg	ttgcatagat	12840
aaccgacgaa	gttttgTTTT	tctggataat	ttagttgact	ttctagcttg	ctgtatgacg	12900

cacccttctg	ctgccggcga	actgtttttg	gtatccgatg	gtcaggagat	ttctaccaag	12960
caactggtga	ctgcgcttgc	tgcggaatg	gggcgtcgcc	ccatcatgtg	gcctgttcct	13020
aggtttattc	tgaggtttct	taaattagta	ggaaaggggtg	ggttatacac	tcagttatgc	13080
tgctcactag	aggtcgactc	gtcgaaaggc	aggcttttgc	ttggttggga	accccgcaag	13140
agcacccttt	ccgcgttgga	agatgttggt	agaatatatg	tcaaacgtac	tgaatgatta	13200
tctgcaggcg	ctttgctact	agcatggcgt	accacgcaga	acaatcgaat	agaaccctgt	13260
tgaaggggtg	agagtatttt	tggggataaa	tttataaatg	gaagaatggt	atttgttact	13320
cgctgcagct	ggggtttcgg	gactgcttac	aggcctcttg	cgctggttatg	ccttagcgag	13380
gagcttactt	gacaccccta	actctcgaag	ttcccatgtc	gttcccactc	cacgcggagg	13440
aggggtcgcc	attgtagtta	ctttttgtct	catgctgcct	atttgggctg	tactgggaaa	13500
tatctcatgg	gccgtgtcct	gggctttact	tctcgctggc	ggcgggggtg	ccattattgg	13560
attcatggat	gatcacggtc	atatcgccgc	acgctggcgt	ctgctgggac	attttagtgc	13620
agccttggtc	tattgtact	ttttgaatgg	cataccacca	tttcagattg	ttggtgtcag	13680
ttgggacctg	gggtggttcg	gaggacttct	ctttgctttc	tatctcgtgt	ggttgctgaa	13740
tctctataac	ttcatggatg	ggatcgatgg	acttgctagc	cttcaggcca	ttttgtctg	13800
tgttggtggg	gcattattat	actggctgaa	tggccaactg	acgcaggctt	tgctcccctt	13860
atcgctagct	tttgccgttt	ttggattctt	gttctggaat	tttccacccc	caaaaatttt	13920
catgggagat	gcgggtagtg	gtcttctggg	gattgtttta	ggaattcttt	ccattcatgc	13980
catgtggatg	aatacgaatt	ttttctgggc	atggttggtc	ctgttagggc	ttttcatcgt	14040
cgatgcgacc	tataccctga	ttcgtcgctt	gctgagaggg	gacaagggtg	atgaggctca	14100
tcgaagccat	gcctatcaat	acgcaagccg	atactatgga	aagcatgctc	ctgttacgat	14160
tggcgtcacg	gcattgaacg	tcatctggct	cctccctata	gccttggttg	tcgggagtgg	14220
gtctctagag	cctttgatgg	gcatcgtcat	agcctacgtc	cctctcgttt	ttctggcagt	14280
gaggttcaag	gcgggtaagc	tagagtcgtc	cgctcaggcc	taaaggagta	ggggaatgct	14340
agatcgttta	agagtaaagt	tgttatccat	gcctcgtcgc	tggaaacggt	tgcttcaagt	14400
ggctacggat	atccttctgg	tatggctgtc	tctgtggctc	gcttttgtgg	tccgtctagg	14460
cacagacgat	atgatcgacg	tggtcggcga	gcatgcatgg	cttttcatca	ctgcgccggt	14520
catcgccatt	ccactattca	ttcgcttcgg	catgtatcgc	gcggtgatgc	gctatctcgg	14580
taacgacgca	ttgatcgcca	tcgccaaggc	ggtgaccatc	tcggctctgg	tgctgtcgct	14640
ggtggtgtac	tggatatcgtg	gcgcgccggc	gccggtgccg	cgttccctgg	tgttcaacta	14700
ctggtggttg	agcatgctgc	tgatcggcgg	cttgctgtctg	gccatgcgcc	agtatttcat	14760
gggcgactgg	tactctgctg	tgcagtcggt	accatttctc	aaccgccagg	atggcctgcc	14820

caggggtggtt atctatgggg cggggggcggc cggcaaccag ttggttgagg cgttgcgctct	14880
cggtcggggcg atgcgtccgg tggcggttcat cgatgacgac aagcagatcg ccaaccgggt	14940
cattgccgggt ctgcgggtct ataccgccaa gcatatccgc cagatgatcg acgagacggg	15000
cgcgcaggag gttctcctgg cgattccttc cgccactcgg gcccgggcgcc gagagattct	15060
cgagtccctg gagccgttcc cgctgcacgt gcgcagcatg cccggcttca tggacctggc	15120
cagcgggccgg gtcaaggtgg atgacctgca ggaggtggac atcgctgacc tgctggggcg	15180
cgacagcgtc gcaccgcgca aggagctgct ggaacgggtg atccgcggtc aggtgggtgat	15240
ggtgaccggg gcggggcggt ctatcggttc ggaactctgt cggcagatca tgagttgttc	15300
gcctagcggtg ctgatcctgt tcgaacacag cgaatacaac ctctatagca tccatcagga	15360
actggagcgt cggatcaagc gcgagtcgct ttcgggtgaac ctggtgccga tcctcggttc	15420
ggtgcgcaat cccgagcgcc tgggtggacgt gatgcgtacc tggaaaggta ataccgtcta	15480
ccatgccccg gcctacaagc atgtgccgat cgtcgagcac aacatcgccg agggcggttct	15540
caacaacgtg ataggcacct tgcattcggt gcaggcccggt gtgcagggtc gcgtgcagaa	15600
cttcgtgctg atttccaccg acaaggcggt gcggccgacc aatgtgatgg gcagcaccaa	15660
gcgcctggcg gaaatggtcc ttcaggcgct cagcaacgaa tcggcgcccg tgctgttcgg	15720
cgaccggaag gacgtgcatc acgtcaacaa gaccggttcc accatgggtc gcttcggcaa	15780
cgtcctcggt tcgtccggtt cggtcattcc gctgttccgc gagcagatca agcgcgggcg	15840
cccggtgacg gtcaccacc cgagcatcac ccgttacttc atgaccattc ccgaggcggc	15900
gcagttggtc atccaggccg gttcgatggg gcaggcgga gatgtattcg tgctggacat	15960
ggggccggcg gtgaagatcc tggagctcgc cgagaagatg atccacctgt ccggcctgag	16020
cgtgcgttcc gagcgttcgc cccatggtga catcgccatc gagttcagtg gcctgcgtcc	16080
tggcgagaag ctctacgaag agctgctgat cggtgacaac gtgaatccca ccgaccatcc	16140
gatgatcatg cgggccaacg aggaacacct gagctgggag gccttcaagg tcgtgctgga	16200
gcagttgctg gccgccgtgg agaaggacga ctactcgcggt gttcgccagt tgctgcggga	16260
aaccgtcagc ggctatgcgc ctgacgggtga aatcgtcgac tggatctatc gccagaggcg	16320
gcgagaacct tgagtcatcg ttctccggaa aaggccgcct agcggccttt tttgttttct	16380
ccgtacgatg tttccggtgc cggaccagga agcgactgct ttgctggggc tgtcgatcca	16440
ggtgcgttcc acggcgataa ggtggtttcg tggatgggca acatgtcggc aaggtaaagt	16500
cagccgcatt gttgaattca tcgaaaaacc ggatcagcca caaacgctgg aatcagacat	16560
catggccgtg ggccgttatg tgctttctgc cgatatttgg ccggaacttg aacgcactca	16620
gccaggtgca tggggacgta ttcagctgac tgatgccatt gccgaactgg cgaaaaaaca	16680

gtctgttgac gccatgctga tgactggtga cagctacgac tgtggtaaaa aaatgggtta	16740
tatgcaggcg tttgtgaagt atggactacg caacctgaaa gaaggagcga agttccgcaa	16800
aggtattgag aaattgctta gcgagtaagt ttaaaaaata gacgccctta tagggcgtaa	16860
taacaaataa cggtagtcaa cattcgacgc ggtgatgcag atatgcccgg aatgctgata	16920
ccgttttttc attctaaaaa actcatcatt tcattgagtt aactacaaaa tttagcactg	16980
ttttttataa tgtttcttct tgtttctggc atcaattggg aagataatta gtgtttgagt	17040
ttagaggctt tgcggcagag aagcggagct taacacgtct gtgagagtac gcagtgcact	17100
ggtagctgta aagccagtgg cggtagcgtg tttaaataaa tacattagta atactacata	17160
ttacatcatt gtaggctatt taagcgctac atgataagcg acagcgctag caatcaaacc	17220
ttttaaagtt acttctcagg aatagtaaaa ggaggacagc tatgttgaaa aaagagtatt	17280
taaaaaaccc ttatttagtt ttgtttgcga tgattatatt agcttatgtt tttagtgtat	17340
tttgcagggt ttattgggtt tgggtggcaa gtgagtttaa tgagtatttt ttcaataacc	17400
agttaatgat catttcaaatt gatggctatg cttttgctga gggcgcaaga gatatgatag	17460
caggttttca tcagcctaatt gatttgagtt attatggatc ttctttatcc gcgcttactt	17520
attggcttta taaaatcaca cctttttctt ttgaaagtat cattttatat atgagtactt	17580
ttttatcttc tttggtggtg attcctaacta ttttgctagc taacgaatac aaacgtcctt	17640
taatgggctt tgtagctgct ctttttagcaa gtatagcaaa cagttattat aatcgcaacta	17700
tgagtgggta ttatgatacg gatatgctgg taattgtttt gcctatgttt attttatttt	17760
ttatggtaag aatgatttta aaaaaagact ttttttcatt gattgccttg ccgttattta	17820
taggaattta tctttggtgg tatccttcaa gttatacttt aaatgtagct ttaattggac	17880
tttttttaat ttatacactt atttttcata gaaaagaaaa gattttttat atagctgtga	17940
ttttgtcttc tcttactctt tcaaataatag catggtttta tcaaagtgcc attatagtaa	18000
tactttttgc tttattcgcc ttagagcaaa aacgcttaaa ttttatgatt ataggaattt	18060
taggtagtgc aactttgata tttttgattt taagtgggtg ggttgatcct ataacttatc	18120
agcttaaatt ttatatTTTT agaagtgatg aaagtgcgaa tttaacgcag ggcttttatgt	18180
attttaatgt caatcaaacc atacaagaag ttgaaaatgt agatcttagc gaatttatgc	18240
gaagaattag tggtagtgaa attgtttttt tgttttcttt gtttggtttt gtatggcttt	18300
tgagaaaaca taaaagtatg attatggctt tacctatatt ggtgcttggg tttttagcct	18360
taaaaggggg gcttagattt accatttatt ctgtacctgt aatggcctta ggatttggtt	18420
ttttattgag cgagtttaag gctataatgg ttaaaaaata tagccaatta acttcaaattg	18480
tttgtattgt ttttgcaact attttgactt tagctccagt atttatccat atttacaact	18540
ataaagcgcc aacagttttt tctcaaaatg aagcatcatt attaaatcaa ttaaaaaata	18600

tagccaatag agaagattat gtggtaactt ggtgggatta tggttatcct gtgcgttatt	18660
atagcgatgt gaaaacttta gtagatgggtg gaaagcattt aggtaaggat aatTTTTTcc	18720
cttcttttgc ttttaagcaaa gatgaacaag ctgcagctaa tatggcaaga cttagtgtag	18780
aatatacaga aaaaagcttt tatgctccgc aaaatgatat tttaaaaaca gacattttgc	18840
aagccatgat gaaagattat aatcaaagca atgtggattt gtttctagct tcattatcaa	18900
aacctgattt taaaatcgat acgcaaaaaa ctcgatgat tttatctttat atgcccgtc	18960
gaatgtcttt gatTTTTTct acggtggcta gttttctttt tattaattta gatacaggag	19020
ttttggataa accttttacc tttagcacag cttatccact tgatgtttaa aatggagaaa	19080
tttatcttag caacggagtg gttttaagcg atgatttttag aagtttttaa ataggtgata	19140
atgtggtttc tgtaaatagt atcgtagaga ttaattctat taaacaaggt gaatacaaaa	19200
tcactccaat tgatgataag gtcagtttt atatttttta tttaaaggat agtgctattc	19260
cttacgcaca atttatttta atggataaaa ccatgtttta tagtgcttat gtgcaaatgt	19320
tttttttagg aaattatgat aagaatttat ttgacttggg gattaattct agagatgcta	19380
aggtttttta acttaaaatt taccatacag atgttccaga ttacgcttaa acatgtgaat	19440
tc	19442

<210> 18

<211> 19616

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 18

gaattccctg aggcaattct tctttgatga cggctgatgg tgaggttgac ctggtgaagc	60
tggtcaagga gctttgggtt aacaagggtc tgattcttct gactactctt cttgcattaa	120
tcgggtcttt tacctatgcg tatctgagta agcctgtata tgaatatagg gttgcagtag	180
tgccctctgc tcttgggtct atcgaagggt tcaatgttgg tagaagggag aatggcctag	240
atgcatatac tgtagaagt atctatgoga tcttttcgcg caatctgctt tcggatgaga	300
ataaaaaaga gttcttctat aagatatacc ttccccaggt gggtgaggga gcggaaagcg	360
aagatgagca ggaggagttt tataagaagt tctccaaaga ggtaaagatt gatcctgcta	420
acaagccaga tgcagaccgt tatacggtaa ttgtggaggg cacgaagcga gaggttcttg	480
ctacatgggc acaagctttc gtgcgttttg ctgcggatcg ggccgtgcat gaggttattg	540
atagtgcagg tagagatttc caggtaaaga atgctgcaat gcagagccgc ataaccgtgc	600
tgcagaatat ggogaagggc cgccgtgatg atagaattgc acgtttgaag gaggcattgc	660
tgattgcgga gtgcctcaag atagatggcc cgccattaat agaaggggcg tccgagcaac	720
aactctctc gatcatggat ggtgacttga tgtacatgcg aggagctaag gcgctgcgcg	780
ctgaaatcaa caatcttgag tcgcgtagt tagatgctcc attcattcct gagttgagaa	840
ctctccaaga gaaactatct tggaaactcca gtttgtctgt ggattctgat gcggtggctg	900
tctacaagga agacgagga ctctcttttt caaatcaacc cattaagccg aagaagattc	960
ttatagttac tataggtact ttggcaggat tgataattgg aattctactc gcagtgcctg	1020
ctggttttat aaggaagctt cgtagcgatg gctctcttcg ctaagttttg atctaaaccc	1080
tgatgccact tactggcatc agggcttact tgtttagtgt gttgaattgt attaggggt	1140
atcgagtaga atgtcaataa ggcggggcgt tttctactcc ggtatatcca tgggcagcaa	1200
ctatcttttg cctcttgctg cgattccttt tctcacgaga acattgtcaa gcgaagcgtt	1260
tggccaattg gtgattgcc aggccgtggc tgtcattcta tgtcaactgg tagacttttg	1320
atttattctg gcaggatcaa gaaaggctgc cattatcgat aacaaagttg aactgtctag	1380
tttcttttct gttgtacaga gtgctagatt cttattgttg ctgctttcac ttttagtgct	1440
ggccattttg gctgtatctt ctattttacc aatccccttg cttgtattgg ttgoggctgc	1500
tcttcoggca gtagttggaa attatcttca agcagtatgg ttctttcagg gaagagcgct	1560
gtttggatgg ttggcgctta ccaatttttt gtctaaggta ttttatttcc tattggctgt	1620
tttttttgtc acgaaggatt ctgacctgt gctggcttcg ttggggtttg gtttttcta	1680

tgtcataggt ggaagtgtc tctgtgtat tttatcttct atgggaatac ggtggcgccc	1740
ggttctcgag aaagacagaa ttctcgatat attgcgtgac ggtgctcgat cttttctttc	1800
tctggctttt cttagcttgc acatgcaagt gctcgttgcg gcggttggtg ttgttggtgg	1860
agcctccgag gccggagtgc tttctactgc ggataaatc cttcgcgga togcggtgc	1920
tacttcaccc atagctagcg ctctatttcc gacttttagc aggatgtatg cgagtgcga	1980
cccggcagtc ggcagtttaa gaaggaaagc gctaggtctg atgttactaa tagctattcc	2040
tagttgttta tttcttttct tattttctga atacatttca tatctcctat tcccggaaca	2100
gtccagaggt ctaactgttg taataagaat gtttctgata gtgccagtgt ttgcttgtat	2160
tgggtgttctg tatggagggt tgactcttgc tcttctggg tatgatggtg tatatttgcg	2220
agcaatcttt tttgcggaat tgggcggggt attaacattt atcctcttgg cgctttgggg	2280
ggatgagctt tttggagcgt ggacgctggt cgttacagag gtctctttgg ggatgggaat	2340
gtttttcctg gccacggtta agttgagaga gaaaagggga ctttgatctt aaggcgatcg	2400
ctaggaggac agctatgcgt attgcgattc tgggcgcgac caacattaaa catatgagcc	2460
tgctgagcca ttatctgaac catattgatc tgaacattaa cgaagtggat attatttata	2520
ccgataaata tgatattgaa gaacatattc agggcatcaa caactactac aaatacaaag	2580
tggatatcaa agaagattgg accttcacaa agaaagcgat tgctgtattat cgttttctgc	2640
cgtatgcgat gaaaattctg aaagaaaacc gttatgattt tgtgattgtg tggggcagct	2700
acaccggcca tctgttcaaa agctttctgg aaaaacatta caaaaacaaa ttcactctga	2760
acatccgtga ttactttttc gaaaacaaca aactgattaa atatcgatg aagaaaatcg	2820
tggatgagag ccgtgtgacc accctgagca gcgaaggctt tctgaaattc ctgccgaaaa	2880
gcgaaaaata ccgtatcatc tacagctaca acatgagcat catccgtgaa agcaacgtga	2940
ccgatggctt taaaaaacgt tggccgatta acattggctt tattggcaac gtgcgtttta	3000
acgaaattaa ccagaaactg attaaagaac tggcgaacga tagccgtttt catatgcagt	3060
atcttggcac cggcagcgaa aaactggaag tgtttgcgag tgaaaacttt attacaaca	3120
ttacctttag cggcggtttt gatctgaaag aaaccccgaa atatctgaac gaaattgata	3180
ttctgaacaa cctgtttggc aaccagaaca ttgcgctgga taccgcgctg agcattcgta	3240

tgtattatgc gctgtttctg aacaaaccga ttattaccac cgatgatacc tttaccgcga	3300
ccgaagcgaa caaatTTggc ctgggcttta gcattaaccc ggaaaacctg aaaggcattg	3360
gcgatgaact gatggattgg tataacaacc tggatgtgat ggatattaac cataaacgtg	3420
aagcgtatcg taacgatgtg attgaaaaca acaaacagtt ttatcaggaa attggccgta	3480
tttttaacga agaacagaaa ctgattagcg aagaagatct gtaacgttta aacaggagga	3540
cagctatgaa caaaatttat aacgtgacca gctatgtgat tgcgattctg atgtttccgt	3600
gcctgatgct gggcgataaa ccgctgctgt ttctggcgcc gattagctat ggcgtgggca	3660
aactgttcat cagcttcagc aacaaccoga acttcaaatt cagcaaaatc gtgtacgatg	3720
tgctgggctt tctgcgtctg gtgtttattc cggcgatgat tgtgtttttc caggatagca	3780
ccattgataa cctgccgctg ggccaggcgt attttaacca ggcggtgatt tatatgagcg	3840
tggaatttat tattggcagc ctgtttattc tgattctgag caaactgttc aagcatgaag	3900
ttgtgagccg taacagcttt accctgagcg gcagcagcat ttattatatt gtgtttggcc	3960
tggtgatttg cggcattttt gtggcgtttc cggaagtgcg taaaaacatt agctttctga	4020
ttattaaaac cgatgcgatg ggccgtggca ccgaagcgac cagcggcctg aacgtgctgt	4080
ttgtgatgct gtttcagctg gcgctggcgc tgctgtttct gatcatcgcg tacgcgagct	4140
acaaaaaata caaagaaaac ccgaaaatca tctacgtggt gctgccgctg gcgattggca	4200
ttctgaacat tagcctgatt gtgggcgaac gtogtagcta tcagctgtat accatggtgg	4260
cggctgctgac cgttgtgagc atcctgttta gcaaacataa acgtcgtatc aacatcatca	4320
tcatcagcgt gggcatcttc gtgctggcgc tgatgaccct gtataaagaa ctgtatgtgt	4380
ttaactatag cagctatagc gaagcgctga acagcaccag cgtgagcaac ctgaaaattg	4440
tggataccct gcagagctat ttttatggcc cgagcaacat tgcggcgagc attgattatc	4500
tgaactatta taacggcagc tttaaacagt atctgtttga taacaccogt gcggtgtttg	4560
gctttaactt tttcctggat aaaaaacagc tgattaccag ccagctgttt aaccagctga	4620
tttatggcag caaacagctg accggccatc tgattagcag cgcgggctat ggcattatTT	4680
atTTtggccc gctgtttttc tacctgaacc tgattgCGaa catctttttc gcgtttctga	4740
gcgaatacat catccgtaaa agccatagcc tggaagtgat cttcatcggc acctacatct	4800
acatgcgtct gattaccagc attttttagcc atccgacccc gctgattacc ctgattagca	4860

tgattctggt ggtgtatgtg attgcatca tcccgggcat catcatcaag aaattcacca	4920
aaaaagtggg catcgaagat tacaaagatg atgatgataa ataacgttta aacaggagga	4980
cagctatgat tgtgaaaacc tttatgaaaa gcaaaatttt tcgtctgatg aacaccccg	5040
tgctgctgtt ttataaaaaa gaatatctga cgggctatta ttttgaaaac aaagtggcgg	5100
gctggctgtg ggcgtggaaa gcggtgccat tcaagctgct gggcattaac accagcctgc	5160
cgtttccggc ggatattacc gtgcgtatgc ataaccgaa caacattgtg tttgataaaa	5220
acgatattca tttttttcag agcccgggca cctattttta caactttagc gcggtgattt	5280
atattggcgg tggcgtgtat attgcgccga acgtgggcat tattaccgag aaccataaca	5340
ttaaaaacct gaaaagccat gcgcggggcg aagatgtgaa aattggcaac tatagctgga	5400
ttggcatgaa cagcgtgatt ctgccgggcg tggaactggg cgaacatacc attgtgggag	5460
cgggcagcgt ggtgaccaa agctttccgg aaggcaacgt ggtgattggc ggcaaccgg	5520
cgaaaattat taagaaaatc agctatccgt atgatgtgcc ggattatgag taattaatta	5580
accaggtgca cgaagaaaat tatgagatta aataaattta ttggcgattc gtttttaatg	5640
attttaagca gtggcatcgc tcaagtcata ttaatcatca ctacccaat tattacaaga	5700
ctatattcac ctacagaatt tgggtgagttt acaatttttt caaatatcgc aatgatttta	5760
ataccaataa taaatgcaag atacgatttg ttgattgtga ataccaaaaa tgaccgtagt	5820
gctaataatac tttcacaaat cagttttttg atatcattgc ttattttatt aatactgata	5880
ccaatatttg cgattagtgc atgtttatac ccaaacttta tattagattt tattttcatt	5940
attattatgt tgtttttggt aagtttaaca aacattttta caaattatct aaataaggaa	6000
agaaagtata aagtgttaag tttgattaat gtgttttagag ctggatcaat ggctttactt	6060
caaatcattt tcggactttt agcattagga agtttaggat taattattgg tttttcatta	6120
tcctatatcg caggcattac actaggatat aaaacgttta aaaagcactt taatattgtg	6180
agagataaag aagaaactaa agcattattt ttagaaaata aaaatcagtt agtttattca	6240
acaccatcaa tattattaaa tagtttgtct ttctcgggtg ttgtgttctt tataggtatt	6300
ttgtatacca atacagaagt gggatattat ggtatggcca taagagtact aggcatacca	6360
gtgacaatta tttcattagg gttatcaaaa atatttatgc aacaagccaa tgactattat	6420
attgaacatg gtaacttccg aaatttatta cttaaattta gttccatact ggttatagtt	6480
tctataattc tttatgtgcc actttatttg ttcagtgaag aattagtcaa tatattatta	6540
ggacatagct ggggtgacgc aattacagtt ataaaaattg ttatccatt atttggtata	6600

aggctgattg tatcaacggt atcactttct gtgattgtat tacaaaaaca acagttagaa	6660
ttaatactac aagcgttatt tttaatagggt actactgcaa catTTgttat atcaaaaatg	6720
cttaatttaa cttttttaaa ctttgtatct attaatacaa ttgttttaat cgtatcgtac	6780
atgatatttt tcatagcact ctattatttt gctaaaaata aacagttcaa aaattctagt	6840
tatccgtatg atgtgccgga ttatgcgtaa gtgcaccata tgaatatcct ccttagttcc	6900
tattccgaag ttctattct ctagaaagta taggaacttc ggcgcgccta cctgtgacgg	6960
aagatcactt cgcagaataa ataaatcctg gtgtccctgt tgataccggg aagccctggg	7020
ccaacttttg gcgaaaatga gacgttgatc ggcacgtaag aggttccaac tttcaccata	7080
atgaaataag atcactaccg ggcgtatttt ttgagttgtc gagattttca ggagctaagg	7140
aagctaaaat ggagaaaaaa atcactggat ataccaccgt tgatatatcc caatggcatc	7200
gtaaagaaca ttttgaggca tttcagtcag ttgctcaatg tacctataac cagaccgttc	7260
agctggatat tacggccttt ttaaagaccg taaagaaaaa taagcacaag ttttatccgg	7320
cctttattca cattcttgcc cgcctgatga atgctcatcc ggaattacgt atggcaatga	7380
aagacggtga gctggtgata tgggatagtg ttcacccttg ttacaccgtt ttccatgagc	7440
aaactgaaac gttttcatcg ctctggagtg aataccacga cgatttcogg cagtttctac	7500
acatatattc gcaagatgtg gcgtgttacg gtgaaaacct ggcctatttc cctaaagggg	7560
ttattgagaa tatgtttttc gtctcagcca atccctgggt gagtttcacc agttttgatt	7620
taaacgtggc caatatggac aacttcttcg ccccggtttt caccatgggc aaatattata	7680
cgcaaggcga caaggtgctg atgcgcgtgg cgattcaggt tcatcatgcc gtttgtgatg	7740
gcttccatgt cggcagatgc ttaatgaata caacagtact gcgatgagtg gcagggcggg	7800
gcgtaaggcg cgccatttaa atgaagttcc tattccgaag ttctattct ctagaaagta	7860
taggaacttc gaagcagctc cagcctacac gttaaagaggt ggtgtatgga taagaactct	7920
gttctgttaa ttaccggtgg aactggttcc tttggaaatg ccgttttgaa gcgttttctg	7980
gatacagata ttggtgaaat acgtgtattc agtcgggatg agaagaagca agatgatatg	8040
cgtaagtgct atgctacccc caagttgagg ttttatattg gagatgtccg cgactatcaa	8100
agcactctga atgctactcg cgggtgtggac tacattttcc atgcggctgc attgaaacag	8160

gttcctctct	gtgagttcta	cccgatggag	gctgtgaaga	ccaatgtcat	cggaacggaa	8220
aatgtcctcg	aatctgctat	ccagaatggc	gtcaaaaaag	tcgtttgtct	gagtacggat	8280
aaggcagttt	acccaatcaa	tgccatgggt	atttcaaagg	ccatgatgga	aaaggtcattg	8340
gtcgccaagt	ctcgaaacct	ggaacgcact	cctaccgtaa	tttgtggcac	ccgttatggg	8400
aacgtcatgg	cttcaagggg	ctcggtcatt	cctctcttta	tcgagcaa	gcgatcaggc	8460
cagcctctca	caatcactga	tccaaacatg	acacgtttca	tgatgacgct	tacggatgcc	8520
gtagaccttg	tgctttatgc	tttcgagcat	ggtaccaatg	gagatctttt	tgtacagaaa	8580
gcaccggctg	caaccatcga	ggtgctggct	catgcgctta	ctcaattgct	tggcaagaat	8640
ggttatccta	tcaatgtaat	aggtacgcgt	catggagaga	agctttatga	ggcgcttctt	8700
agtcgagagg	aaatggcctg	tgctgaagac	atgggtgact	attatcgtat	ccccccagat	8760
ttgcgtgatc	tgaactatag	taagtttg	gagcaaggcg	aggaaaaat	tactcataca	8820
gaggactata	attctcataa	taccaaacgt	ctggatatcg	aaggaatgaa	aaagctgttg	8880
ctgaagttgg	atttcattcg	tgctattcag	cgtggcgaga	gtgccagtcc	agaggaataa	8940
acgatgaaag	ttcttgtaac	tggcgcgaat	ggatttggtg	gaaggaatct	gtgcgctcat	9000
cttgcagagc	ggggtgggat	cgaggtggg	ccattcacc	gcgagagtag	tgttggaat	9060
ttgcctgagc	taattcggtc	cgtcgatttt	atttttcac	ttgccgggg	caatcgccg	9120
gaaaaaccag	aagagtttaa	gatcggaat	tccgaactca	cgtatgctct	gtgtgaggcg	9180
gtaaggtcca	atggacgagc	cataccactt	ctttatactt	catccattca	ggctgagggtg	9240
gataatgagt	acggtttaag	caagcgagcc	gcagaagagc	atctccaagt	gctagggtgag	9300
gatattgggt	gtcctgtcta	catatttcgc	cttcctaagt	tattcggtaa	atggctcggt	9360
ccgaattata	attcagcggg	tgcgactttt	tgtcataata	ttattcgaga	tattccgatt	9420
caaattaaca	attcctcggc	agagatcact	cttgatataca	tagatgatgt	ggttcgcacc	9480
ttcatgaaag	tcatggatgg	gaagctatcc	aatgcagttt	cactacaggt	cgagccccag	9540
tatcagattt	ctggttggtga	gctcgagaa	caattgtatg	agtttcgtaa	tagtcgaaag	9600
tcactgacta	ccgcaagggg	tggctcggga	ttgacgcgcg	ccttgctactc	gacttatcta	9660
agtttcttgc	cagaagatag	ttttagttac	gacgtgccaa	tgcatcggga	tccgcgtggc	9720
acattcgctc	agatgctgaa	gaccgcggac	tctggccagt	tctcgttttt	tacggctcat	9780
ccagggtgta	ccagggggcg	gcattaccat	cactcgaaaa	ccgaaaagtt	tctggttatc	9840
aaagggatgg	cacgtttcaa	gtttagaaac	atcctgaccg	gggcatttta	cgaaatttgc	9900
actaatggtg	aaaaggcaga	aattgtcgaa	acagtacctg	gatggactca	tgacattact	9960
aatgtcggaa	ctgacgatat	ggtagtcatg	ttgtgggcta	acgaagtatt	tgatcgggaa	10020
aatccggata	cctacgcttg	ttcagtaggc	gaagggtcgt	aaggatatgt	gagataacaa	10080

tgcagaagct	aaaagtcggt	acgggttggtg	gaactcgtcc	tgagattatt	cgcttgctcta	10140
gggtcatggc	gaagcttgat	cagtactgcg	atcatgtact	tgtccatact	ggacagaatt	10200
atgattacga	acttaatgaa	atattttttc	aggacctcgg	tataagaaag	ccggattatt	10260
ttctaaacgc	cgccgggtct	tccggggctg	aaacgatagg	gaatgtaata	atcgcagtcg	10320
atcgtgttct	gggcgaaata	gatcccgatg	cgctgctcgt	gctgggtgat	accaatagtt	10380
gtatggcggg	actgcctgca	aaacggcgta	agataccgac	ctttcatatg	gaagcaggca	10440
atcgctgttt	cgatatgcgt	gtgcctgaag	agataaatcg	gcgcattgtc	gatcatacag	10500
ctgatgtaaa	tttgacctat	agtacaattg	cgcgtgatta	tctcttgctg	gaaggacttt	10560
ctccagacat	ggttatcaag	actggtagcc	ctatgttcga	agttctcgag	cactatcgtg	10620
acgggatcga	gtcctccgat	attcttgaaa	ggctcgggtt	gaaaacagag	cggttctttg	10680
tcgtgagtgc	gcaccgagag	gaaaacatag	attcggataa	gaatttcttg	aagttggttt	10740
ctatgctcaa	cgctgtggca	gaaaagtact	cgctgcccg	catcgtatca	actcacccta	10800
gaacaaaaaa	gagaattgag	gcgacggagg	caaagtttca	cgagggtatt	aaactgctga	10860
aaccctcgg	ctttaaggat	tacaataaac	tgcaaattac	agccaaggca	gttatttctg	10920
acagtgggac	catcagtgag	gagtcttcaa	tactgaattt	tcccgccttg	aatattcgtg	10980
aggctcatga	acgcccagaa	ggcatggaag	aggctgtggt	gatgatggtc	ggactggatt	11040
cggatcgagt	actacaagca	ctcgagggtg	tggagggaca	gaggcgcgac	gcagagcgca	11100
tgttacgctt	ggtcgctgac	tatagcatgc	ccaacgtttc	tgaaaagatt	gttcgcatag	11160
ttcatagcta	tcgggactat	gtcatgcgaa	ctgtctggaa	aaaatattaa	cttgaggcgt	11220
ggagttagtg	gcaaggatat	ttgtggtttc	tgagtatgtc	ggtgccaatc	agaactccac	11280
gggatactat	tgggagaaga	taataggaaa	gatgcagcgg	gagtttggtg	ggctaaccgt	11340
aattttcccg	ctgaccgcag	gtgaaacccc	gcctgtggtt	tcaccttccg	ttgagcaaga	11400
atgctttaag	tttccgagga	gcaataagaa	taggtctcct	tctagaggat	tggcgcagat	11460
ttttcaggcg	tttctgttct	cagtaaaatt	gacttctcgt	gccagacgag	gagatgtggt	11520
attgagtgga	accaaccctg	ctcttctact	gatgacgttt	cccttgctaa	ggtatgccct	11580
cggtttcaag	tgggtgctgc	tgggtgcatga	tgtgtttccc	gagaacttgg	tgccggcggg	11640
cgttctgaag	aaagatagta	ttgcctaccg	gcttctacgt	cgtctctttt	ctttcattta	11700
ctcatccgct	gatcgtctag	tcgtaatagg	gcgcgatatg	gaagctctta	tgaaagagaa	11760
ggtgaatgac	ccgcgatctt	tgggtctttat	ttcgaattgg	gcctgtgaga	aagaggtttt	11820
cccagtaccg	agagaggatg	ctccttttat	caatattcct	gaatggaaaag	gtaaaagggt	11880
tttccaattt	tttggtaatg	tcggtcgatt	acaaggtata	gaaaacatac	tttctgctat	11940

cctggggtgg	ttcggaggac	ttctctttgc	tttctatctc	gtgtggttgc	tgaatctcta	13920
taacttcatg	gatgggatcg	atggacttgc	tagccttcag	gccatTTTTg	tctgtgttgg	13980
tggggcatta	ttatactggc	tgaatggcca	actgacgcag	gctttgctcc	ccttatcgct	14040
agcttttgcc	gtttttggat	tcttgttctg	gaattttcca	cccccaaaa	ttttcatggg	14100
agatgcgggt	agtgtcttc	tggggattgt	tttaggaatt	ctttccattc	atgccatgtg	14160
gatgaatacg	aattttttct	gggcatggtt	ggtcctgtta	ggcgttttca	tcgtcgatgc	14220
gacctatacc	ctgattcgtc	gcttgctgag	aggggacaag	gtgtatgagg	ctcatcgaag	14280
ccatgcctat	caatacgcaa	gccgatacta	tggaaagcat	gctcctgtta	cgattggcgt	14340
cacggcattg	aacgtcatct	ggctcctccc	tatagccttg	ttggtcggga	gtgggtctct	14400
agagcctttg	atgggcatcg	tcatagccta	cgtccctctc	gtttttctgg	cagtgaggtt	14460
caaggcgggt	aagctagagt	cgtccgctca	ggcctaaag	agtaggggaa	tgctagatcg	14520
tttaagagta	aagtgttat	ccatgcctcg	tcgctggaaa	cgtttgcttc	aagtggctac	14580
ggatatcctt	ctggtatggc	tgtctctgtg	gctcgctttt	gtggtccgtc	taggcacaga	14640
cgatatgatc	gacgtgttcg	gcgagcatgc	atggcttttc	atcactgcgc	cggtcatcgc	14700
cattccacta	ttcattcgct	tcggcatgta	tcgcgcggtg	atgcgctatc	tcggtaacga	14760
cgcattgatc	gccatcgcca	aggcggtgac	catctcggct	ctggtgctgt	cgctgggtgt	14820
gtactggtat	cgtggcgcgc	cggcgccggt	gccgcgttcc	ctggtgttca	actactggtg	14880
gttgagcatg	ctgctgatcg	gcggcttgcg	tctggccatg	cgccagtatt	tcatgggcga	14940
ctggtactct	gctgtgcagt	cggtaaccatt	tctcaaccgc	caggatggcc	tgcccaggggt	15000
ggttatctat	ggggcggggg	cggccggcaa	ccagttggtt	gcggcgttgc	gtctcggtcg	15060
ggcgatgcgt	ccggtggcgt	tcatcgatga	cgacaagcag	atcgccaacc	gggtcattgc	15120
cggctctgcg	gtctataccg	ccaagcatat	ccgccagatg	atcgacgaga	cgggcgcgca	15180
ggaggttctc	ctggcgattc	cttcgccac	tcgggcccgg	cgcgagaga	ttctcgagtc	15240
cctggagccg	ttcccgtgc	acgtgcgcag	catgcccggc	ttcatggacc	tggccagcgg	15300
ccgggtcaag	gtggatgacc	tgcaggaggt	ggacatcgct	gacctgctgg	ggcgcgacag	15360
cgtcgcaccg	cgcaaggagc	tgttggaacg	gtgcatccgc	ggtcaggtgg	tgatggtgac	15420
cggggcgggc	ggctctatcg	gttcggaact	ctgtcggcag	atcatgagtt	gttcgcctag	15480
cgtgctgatc	ctgttcgaac	acagcgaata	caacctctat	agcatccatc	aggaactgga	15540
gcgtcggatc	aagcgcgagt	cgttttcggt	gaacctgttg	ccgacccctc	gttcggtgcg	15600
caatcccag	cgcctggtgg	acgtgatgcg	tacctggaag	gtcaataacc	tctaccatgc	15660
ggcggcctac	aagcatgtgc	cgatcgtcga	gcacaacatc	gccgagggcg	ttctcaacaa	15720

ttcatcagcc	taatgatttg	agttattatg	gatcttcttt	atccgcgctt	acttattggc	17700
tttataaaat	cacacctttt	tcttttgaaa	gtatcatttt	atatatgagt	acttttttat	17760
cttctttggt	ggtgattcct	actattttgc	tagctaacga	atacaaacgt	cctttaatgg	17820
gctttgtagc	tgctctttta	gcaagtatag	caaacagtta	ttataatcgc	actatgagtg	17880
ggtattatga	tacggatatg	ctggtaattg	ttttgcctat	gtttatttta	ttttttatgg	17940
taagaatgat	tttaaaaaaa	gacttttttt	cattgattgc	cttgccgtta	tttataggaa	18000
tttatctttg	gtggtatcct	tcaagttata	ctttaaatgt	agctttaatt	ggactttttt	18060
taattttatac	acttattttt	catagaaaag	aaaagatttt	ttatatagct	gtgattttgt	18120
cttctcttac	tctttcaaata	atagcatggt	tttatcaaag	tgccattata	gtaatacttt	18180
ttgcttttatt	cgccttagag	caaaaacgct	taaattttat	gattatagga	attttaggtta	18240
gtgcaacttt	gatatttttg	attttaagt	gtgggggtga	tcctataact	tatcagctta	18300
aattttatat	ttttagaagt	gatgaaagt	cgaatttaac	gcagggcttt	atgtatttta	18360
atgtcaatca	aaccatacaa	gaagttgaaa	atgtagatct	tagcgaattt	atgcgaagaa	18420
ttagtggttag	tgaaattggt	tttttggttt	ctttggttgg	ttttgtatgg	cttttgagaa	18480
aacataaaaag	tatgattatg	gctttaccta	tattggtgct	tggtttttta	gccttaaaaag	18540
gggggcttag	atttaccatt	tattctgtac	ctgtaatggc	cttaggattt	ggttttttat	18600
tgagcgagtt	taaggctata	atggttaaaa	aatatagcca	attaacttca	aatgtttgta	18660
ttgtttttgc	aactattttg	acttttagctc	cagtatttat	ccatatttac	aactataaaag	18720
cgccaacagt	tttttctcaa	aatgaagcat	cattattaaa	tcaattaaaa	aatatagcca	18780
atagagaaga	ttatgtggtta	acttggtggg	attatggtta	tcctgtgcgt	tattatagcg	18840
atgtgaaaac	tttagtagat	ggtggaaagc	atttaggtta	ggataatttt	ttcccttctt	18900
ttgctttaag	caaagatgaa	caagctgcag	ctaatatggc	aagacttagt	gtagaatata	18960
cagaaaaaag	cttttatgct	cgcgaaaatg	atattttaaa	aacagacatt	ttgcaagcca	19020
tgatgaaaga	ttataatcaa	agcaatgtgg	atttgtttct	agcttcatta	tcaaaacctg	19080
atttttaaata	cgatacgcca	aaaactcgtg	atatttatct	ttatatgcc	gctagaatgt	19140
ctttgatttt	ttctacggtg	gctagttttt	cttttattaa	tttagatata	ggagttttgg	19200
ataaaccttt	taccttttagc	acagcttatc	cacttgatgt	taaaaatgga	gaaatttatc	19260
ttagcaacgg	agtggtttta	agcgatgatt	ttagaagttt	taaaataggt	gataatgtgg	19320
tttctgtaaa	tagtatcgta	gagattaatt	ctattaaaca	aggtgaatac	aaaatcactc	19380
caattgatga	taaggctcag	ttttatatatt	tttattttaa	ggatagtgt	attccttacg	19440
cacaatttat	tttaatggat	aaaaccatgt	ttaatagtgc	ttatgtgcaa	atgttttttt	19500

taggaaatta tgataagaat ttatttgact tgggtgattaa ttctagagat gctaaggttt 19560
 ttaaacttaa aatttaocca tacgatgttc cagattacgc ttaaacatgt gaattc 19616

<210> 19

<211> 20597

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 19

gaattccctg aggcaattct tctttgatga cggctgatgg tgaggttgac ctggtgaagc	60
tgggtcaagga gctttgggtt aacaagggtc tgattcttct gactactctt cttgcattaa	120
tcgggtcttt tacctatgcg tatctgagta agcctgtata tgaatatagg gttgcagtag	180
tgcctcctgc tcttgggtct atogaagggt tcaatgttgg tagaaggag aatggcctag	240
atgcatatac tgttagaagt atctatgcga tcttttcgcg caatctgctt tcggatgaga	300
ataaaaaaga gttcttctat aagatatacc ttccccaggt ggggtgagga gcggaagcg	360
aagatgagca ggaggagttt tataagaagt tctccaaaga ggtaaagatt gatcctgcta	420
acaagccaga tgcagaccgt tatacggtaa ttgtggaggg cacgaagcga gaggttcttg	480
ctacatgggc acaagctttc gtgcgttttg ctgcggatcg ggccgtgcat gaggttattg	540
atagtgcagg tagagatttc caggtaagaa atgctgcaat gcagagccgc ataaccgtgc	600
tgcagaatat ggcgaagggc cgccgtgatg atagaattgc acgtttgaag gaggcattgc	660
tgattgcgga gtcgctcaag atagatggcc cgccattaat agaaggggcg tccgagcaac	720
aactctcctc gatcatggat ggtgacttga tgtacatgcg aggagctaag gcgctgcgcg	780
ctgaaatcaa caatcttgag tcgcgtagtg tagatgctcc attcattcct gagttgagaa	840
ctctccaaga gaaactatct tggaactcca gtttgtctgt ggattctgat gcggtggctg	900

tctacaagga agacgagga ctctcttttt caaatcaacc cattaagccg aagaagattc	960
ttatagttac tataggtact ttggcaggat tgataattgg aattctactc gcagtgcctg	1020
ctggttttat aaggaagctt cgtagcgatg gctctcttcg ctaagttttg atctaaaccc	1080
tgatgccact tactggcatc agggcttact tgttgtagtt gttgaattgt attagggggt	1140
atcgagtaga atgtcaataa ggccggggcgt tttctactcc ggtatatcca tgggcagcaa	1200
ctatcttttg cctcttgctg cgattccttt tctcacgaga acattgtcaa gcgaagcgtt	1260
tggccaattg gtgattgcc aggccgtggc tgtcattcta tgtcaactgg tagactttgg	1320
atttattctg gcaggatcaa gaaaggctgc cattatcgat aacaaagttg aactgtctag	1380
tttcttttct gttgtacaga gtgctagatt cttattgttg ctgctttcac ttttagtgct	1440
ggccattttg gctgtatctt ctattttacc aatccccttg cttgtattgg ttgcggctgc	1500
tcttcggca gtagttggaa attatcttca agcagtatgg ttctttcagg gaagagcgct	1560
gtttggatgg ttggcgctta ccaatttttt gtctaaggta ttttatttcc tattggctgt	1620
tttttttgc acgaaggatt ctgacctgt gctggcttcg ttggggtttg gtttttcta	1680
tgtcataggt ggaagtgtc tctgttgat tttattttct atgggaatac ggtggcgccc	1740
ggttctcgag aaagacagaa ttctogatat attgcgtgac ggtgctcgat cttttctttc	1800
tctggctttt cttagcttgc acatgcaagt gctcgttgcg gcggttggtg ttgttggtgg	1860
agcctccgag gccggagtgc tttctactgc ggataaattc cttcgcgagg tgcggctgc	1920
tacttcaccc atagctagcg ctctatttcc gacttttagc aggatgtatg cgagtgcga	1980
cccggcagtc gccagtttaa gaaggaaagc gctaggtctg atgttactaa tagctattcc	2040
tagttgttta tttcttttct tattttctga atacatttca tatctctat tcccggaaca	2100
gtccagaggt ctaactgttg taataagaat gttttcgata gtgccagtgt ttgcttgtat	2160
tgggtgttctg tatggagggt tgactcttgt tctttctggg tatgatggtg tatatttgcg	2220
agcaattttt ttgcggaat tgggcgggggt attaacattt atcctcttgg cgctttgggg	2280
ggatgagctt ttgggagcgt ggacgctggc cgttacagag gtctctttgg ggatgggaat	2340
gtttttctctg gccacgggta agttgagaga gaaaagggga ctttgatctt aagatgagag	2400
tagaaaataa taatgtttct gggcaaaccc atgaccggga acagattgat ttgattgatt	2460
tactagtga gttgtggcgt ggcaagatga caatcatcat ttccgtcatt gtggctattg	2520

ccctagctat tggatatttg gcagtagcga aggagaaatg gacgtcaaca gcaattatca	2580
ctcagccccga tgtggggcaa attgctgggt ataacaatgc catgaatggt atctatggtc	2640
aggctgcacc gaaagtatcg gatttgcagg agacgttaat tggtcgcttc agttctgcct	2700
tctctgcatt agcagaaacg ctggataatc aggaagaacc agaaaaactt accatogaac	2760
cttctgttaa gaaccagcaa ttaccattga ctgtttctta tgttgggcaa actgcagagg	2820
gcgcacaaat gaagttggcc caatacattc agcaagttga cgataaagt aatcaagagt	2880
tagaaaagga tctcaaggac aacattgctc tgggacggaa aaacttgag gactctttaa	2940
gaacgcagga agtgggttgcg caggagcaga aagatctgcg tatccgtcag attcaggaag	3000
cggtgcagta tgcgaatcag gcgcaggtga caaaaccgca gattcaacag actggcgaag	3060
atatcacaca agatacgttg ttccttttgg ggagcgaagc gctggagtcg atgattaagc	3120
atgaggcgac ccgtccgttg gtgttctcac caaactacta tcagactcgt caaacctgc	3180
ttgatatcga aagcttaaag gttgatgac ttgatattca tgcttaccgc tatgtaatga	3240
aaccgacgtt acctattcgt cgtgatagcc cgaaaaaggc aattaccttg attctggcgg	3300
tgctgctggg tggcatggtt ggcgcgggga ttgtgctggg gcgtaatgct ctacgcaatt	3360
acaacgcgaa gtaagcgatc gctaggagga cagctatgcg tattgcgatt ctgggcgga	3420
ccaacattaa acatatgagc ctgctgagcc attatctgaa ccatattgat ctgaacatta	3480
acgaagtga tattatttat accgataaat atgatattga agaacatatt cagggcatca	3540
acaactacta caaatacaaa gtggatatca aagaagattg gaccttcac aagaaagcga	3600
ttgcgtatta tcgttttcgt ccgtatgcga tgaaaattct gaaagaaaac cgttatgatt	3660
ttgtgattgt gtggggcagc tacaccggcc atctgttcaa aagctttctg gaaaaacatt	3720
acaaaaacaa attcatcctg aacatccgtg attacttttt cgaaaacaac aaactgatta	3780
aatatcgtat gaagaaaatc gtggatgcga gccgtgtgac caccctgagc agcgaaggct	3840
ttctgaaatt cctgccgaaa agcgaaaaat accgtatcat ctacagctac aacatgagca	3900
tcatccgtga aagcaacgtg accgatggct ttaaaaaacg ttggccgatt aacattggct	3960
ttattggcaa cgtgcgtttt aacgaaatta accagaaact gattaaagaa ctggcgaacg	4020
atagccgttt tcatatgcag tattttggca ccggcagcga aaaactggaa gtgtttgcgc	4080
gtgaaaaactt tattaacaac attaccttta gcggcggctt tgatctgaaa gaaaccccca	4140
aatatctgaa cgaaattgat attctgaaca acctgttttg caaccagaac attgcgctgg	4200

ataccgcgct gagcattcgt atgtattatg cgctgtttct gaacaaaccg attattacca	4260
ccgatgatac ctttaccgcg accgaagcga acaaatttgg cctgggcttt agcattaacc	4320
cggaaaacct gaaaggcatt ggogatgaac tgatggattg gtataacaac ctggatgtga	4380
tggatattaa ccataaacgt gaagcgtatc gtaacgatgt gattgaaaac aacaaacagt	4440
tttatcagga aattggccgt atttttaacg aagaacagaa actgattagc gaagaagatc	4500
tgtaacgttt aaacaggagg acagctatga acaaaattta taacgtgacc agctatgtga	4560
ttgcgattct gatgtttccg tgccatgatc tgggcgataa accgctgctg tttctggcgc	4620
cgattagcta tggcgtgggc aaactgttca tcagcttcag caacaaccgc aacttcaaat	4680
tcagcaaaat cgtgtacgat gtgctgggct ttctgcgtct ggtgtttatt ccggcgatga	4740
ttgtgttttt ccaggatagc accattgata acctgccgct gggccaggcg tattttaacc	4800
aggcggatgat ttatatgagc gtggaattta ttattggcag cctgtttatt ctgattctga	4860
gcaaaactgtt caagcatgaa gttgtgagcc gtaacagctt taccctgagc ggcagcagca	4920
tttattatat tgtgtttggc ctggtgatct gcggcatttt tgtggcgttt ccggaagtgc	4980
gtaaaaacat tagctttctg attattaaaa ccgatgcgat gggccgtggc accgaagcga	5040
ccagcggcct gaacgtgctg tttgtgatgc tgtttcagct ggcgctggcg ctgctgttct	5100
tgatcatcgc gtacgcgagc tacaaaaaat acaaagaaaa cccgaaaatc atctacgtgg	5160
tgctgccgct ggcgattggc attctgaaca ttagcctgat tgtgggcgaa cgtcgtagct	5220
atcagctgta taccatggtg gcggtgctga ccgttgtgag catcctgttt agcaaacata	5280
aacgtcgtat caacatcatc atcatcagcg tgggcattct cgtgctggcg ctgatgacct	5340
tgtataaaga actgtatgtg ttttaactata gcagctatag cgaagcgtg aacagcacca	5400
gcgtgagcaa cctgaaaatt gtggataccc tgcagagcta tttttatggc ccgagcaaca	5460
ttgcggcgag cattgattat ctgaactatt ataacggcag ctttaaacag tatctgtttg	5520
ataacaccgc tgcggtgttt ggctttaact ttttcctgga taaaaaacag ctgattacca	5580
gccagctgtt taaccagctg atttatggca gcaaacagct gaccggccat ctgattagca	5640
gcgcgggcta tggcattatt tattttggcc cgctgttttt ctacctgaac ctgattgoga	5700
acatcttttt cgcgtttctg agcgaatata tcatccgtaa aagccatagc ctggaagtga	5760
tcttcatcgg cacctacatc tacatgcgtc tgattaccag cttttttagc catccgacct	5820
cgctgattac cctgattagc atgattctgg tgggtgatgt gattgcgac atcccgggca	5880

tcatcatcaa gaaattcacc aaaaaagtgg gcatcgaaga ttacaaagat gatgatgata	5940
aataacgttt aaacaggagg acagctatga ttgtgaaaac ctttatgaaa agcaaaat	6000
ttcgtctgat gaacaccccg ctgctgctgt tttataaaaa agaatatctg accggctatt	6060
atTTTgaaaa caaagtggcg ggctggctgt gggcgtggaa agcggtgcca ttcaagctgc	6120
tgggcattaa caccagcctg ccgtttccgg cggatattac cgtgcgtatg cataaccga	6180
acaacattgt gtttgataaa aacgatattc atatTTTtca gagcccgggc acctatTTTa	6240
acaactttag cgcgggtgatt tatattggcc gtggcgtgta tattgcgccg aacgtgggca	6300
ttattaccgc gaaccataac attaaaaacc tgaaaagcca tgcgccgggc gaagatgtga	6360
aaattggcaa ctatagctgg attggcatga acagcgtgat tctgccgggc gtggaactgg	6420
gcgaacatac cattgtgggc gcgggcagcg tggtgaccaa aagctttccg gaaggcaacg	6480
tggtgattgg cggcaacccg gcgaaaatta ttaagaaat cagctatccg tatgatgtgc	6540
cggattatgc gtaattaatt aaccaggtgc acgaagaaaa ttatgagatt aaataaat	6600
attggcgatt cgtttttaat gattttaagc agtggcatcg ctcaagtcatt attaatcatc	6660
actaccccaa ttattacaag actatatcca cctacagaat ttgggtgagtt tacaatTTTt	6720
tcaaatatcg caatgatttt aataccaata ataaatgcaa gatacgattt gttgattgtg	6780
aataccaaaa atgaccgtag tgctaataata ctttcacaaa tcagtTTTTt gatatcattg	6840
cttattttat taatactgat accaatattt gcgattagtg catgtttata cccaaacttt	6900
atattagatt ttattttcat tattattatg ttgtttttgg taagtTTTaa aaacattTTTt	6960
acaaattatc taaataagga aagaaagtat aaagtgttaa gtttgattaa tgtgtttaga	7020
gctggatcaa tggctttact tcaaatcatt ttccgacttt tagcattagg aagtttagga	7080
ttaattattg gtttttcatt atcctatatc gcaggcatta cactaggata taaaacgttt	7140
aaaaagcact ttaatatgtg gagagataaa gaagaaacta aagcattatt tttagaaaat	7200
aaaaatcagt tagttttattc aacaccatca atattattaa atagtttgtc tttctcggtt	7260
gttgtgttct ttataggtat tttgtatacc aatacagaag tgggtatttt tggtatggcc	7320
ataagagtac taggcatacc agtgacaatt atttcattag ggttatcaaa aatatttatg	7380
caacaagcca atgactatta tattgaacat ggtaacttc gaaatttatt acttaaattt	7440
agttccatac tggttatagt ttctataatt ctttatgtgc cactttattt gttcagtga	7500

gaattagtca	atatattatt	aggacatagc	tgggttgacg	caattacagt	tataaaaatt	7560
gttatcccat	tatttgttat	aaggctgatt	gtatcaacgg	tatcactttc	tgtgattgta	7620
ttacaaaaac	aacagttaga	attaatacta	caagcgttat	ttttaatagg	tactactgca	7680
acatttgtta	tatcaaaaat	gcttaattta	acttttttaa	actttgtatc	tattaataca	7740
attgttttaa	tcgtatcgta	catgatattt	ttcatagcac	tctattattt	tgctaaaaat	7800
aaacagttca	aaaattctag	ttatccgtat	gatgtgcogg	attatgcgta	agtgcaccat	7860
atgaatatcc	tccttagttc	ctattccgaa	gttcctattc	tctagaaagt	ataggaactt	7920
cggcgcgcct	acctgtgacg	gaagatcact	tcgcagaata	aataaatcct	ggtgtccctg	7980
ttgataccgg	gaagccctgg	gccaaacttt	ggcgaaaatg	agacgttgat	cggcacgtaa	8040
gaggttccaa	ctttcaoccat	aatgaaataa	gatcactacc	gggcgtattt	tttgagttgt	8100
cgagattttc	aggagctaag	gaagctaaaa	tggagaaaaa	aatcactgga	tataccacog	8160
ttgatataatc	ccaatggcat	cgtaaagaac	atthttgaggc	atthtcagtca	gttgctcaat	8220
gtacctataa	ccagaccggt	cagctggata	ttacggcctt	tttaaagacc	gtaaagaaaa	8280
ataagcacia	gttttatccg	gcctttattc	acattcttgc	ccgcctgatg	aatgctcatc	8340
cggaattacg	tatggcaatg	aaagacgggtg	agctgggtgat	atgggatagt	gttcaccctt	8400
gttacaccgt	tttccatgag	caaactgaaa	cgthtttcatc	gctctggagt	gaataccacg	8460
acgatttccg	gcagttttcta	cacatatatt	cgcaagatgt	ggcgtgttac	ggtgaaaacc	8520
tggcctattt	ccctaaaggg	tttattgaga	atatgttttt	cgtctcagcc	aatccctggg	8580
tgagtttcac	cagttttgat	ttaaacgtgg	ccaatatgga	caacttcttc	gccccgtht	8640
tcaccatggg	caaataattat	acgcaaggcg	acaagggtgct	gatgccgctg	gcgattcagg	8700
ttcatcatgc	cgthttgtgat	ggcttccatg	tcggcagatg	cttaatgaat	acaacagttac	8760
tgcgatgagt	ggcagggcgg	ggcgtaaaggc	gcgccattta	aatgaagttc	ctattccgaa	8820
gttcctattc	tctagaaagt	ataggaactt	cgaagcagct	ccagcctaca	cgtaaagagg	8880
tggtgtatgg	ataagaactc	tgthctgtta	attaccgggtg	gaactggthc	ctthtgaaat	8940
gccgthttga	agcgtthttct	ggatacagat	attgggtgaaa	tacgtgtatt	cagtcgggat	9000

gagaagaagc aagatgatat gcgtaagtgc tatgctcacc ccaagttgag gttttatatt	9060
ggagatgtcc gcgactatca aagcactctg aatgctactc gcggtgtgga ctacattttc	9120
catgcggctg cattgaaaca ggttcctctt tgtgagttct acccgatgga ggctgtgaag	9180
accaatgtca tcggaacgga aaatgtcctc gaatctgcta tccagaatgg cgtcaaaaaa	9240
gtcgtttgtc tgagtacgga taaggcagtt tacccaatca atgccatggg tattttcaaag	9300
gccatgatgg aaaaggctcat ggtcgccaag tctcgaaacc tggaacgcac tcctaccgta	9360
atttgtggca cccgttatgg taacgtcatg gcttcaaggg gctcggtcac tcctctcttt	9420
atcgagcaaa tgcgatcagg ccagcctctc acaatcactg atccaaacat gacacgtttc	9480
atgatgacgc ttacggatgc cgtagacctt gtgctttatg ctttcgagca tggtagcaat	9540
ggagatcttt ttgtacagaa agcaccgggt gcaaccatcg aggtgctggc tcatgcgctt	9600
actcaattgc ttggcaagaa tggttatcct atcaatgtaa taggtacgcy tcatggagag	9660
aagctttatg aggcgcttct tagtcgagag gaaatggcct gtgctgaaga catgggtgac	9720
tattatcgta tccccccaga tttgcgtgat ctgaactata gtaagtttgt ggagcaaggc	9780
gaggaaaaaa ttactcatac agaggactat aattctcata ataccaaacg tctggatatac	9840
gaaggaatga aaaagctgtt gctgaagttg gatttcattc gtgctattca gcgtggcgag	9900
agtgccagtc cagaggaata aacgatgaaa gttcttgtaa ctggcgcgaa tggatttggt	9960
ggaaggaatc tgtgcgctca tcttcgagag cgggggtggt tcgaggtggg gccattcacc	10020
cgcgagagta gtgttggtaa tttgcctgag ctaattcgtt ccgtcgattt tatttttcat	10080
cttgccgggg tcaatcgtcc ggaaaaacca gaagagttaa agatcgggaa ttccgaactc	10140
acgtatgctc tgtgtgaggg ggtaagggtcc aatggacgag ccataccact tctttatact	10200
tcatccattc aggctgaggt ggataatgag tacggtttaa gcaagcgagc cgcagaagag	10260
catctccaag tgctaggtga ggatattggg tgtcctgtct acatatttcg ccttcctaata	10320
gtattcggta aatggtcgcg tccgaattat aattcagcgg ttgcgacttt ttgtcataat	10380
attattcgag atattccgat tcaaattaac aattcctcgg cagagatcac tcttgatac	10440
atagatgatg tggttcgcac ctcatgaaa gtcattgatg ggaagctatc caatgcagtt	10500
tcactacagg tcgagcccca gtatcagatt tctgttggtg agctcgcaga acaattgtat	10560
gagtttcgta atagtcgaaa gtcactgact accgcaaggg ttggctcggg attgacgcgc	10620
gccttgact cgacttatct aagtttcttg ccagaagata gttttagtta cgacgtgcc	10680
atgcattcgg atccgcgtgg cacattcgtc gagatgctga agaccgcgga ctctggccag	10740
ttctcgtttt ttacggctca tccaggtgtt accaggggcy ggcattacca tcactcgaaa	10800
accgaaaagt ttctggttat caaagggatg gcacgtttca agtttagaaa catcctgacc	10860

ggggcatttt	acgaaatttg	cactaatggt	gaaaaggcag	aaattgtcga	aacagtacct	10920
ggatggactc	atgacattac	taatgtcgga	actgacgata	tggtagtcac	gttgtgggct	10980
aacgaagtat	ttgatcggga	aaatccggat	acctacgctt	gttcagtagg	cgaagggtgcg	11040
taaggtatag	tgagataaca	atgcagaagc	taaaagtcgt	tacggttgtt	ggaactcgtc	11100
ctgagattat	tcgcttgtct	agggtcatgg	cgaagcttga	tcagtactgc	gatcatgtac	11160
ttgtccatac	tggacagaat	tatgattacg	aacttaatga	aatatTTTTT	caggacctcg	11220
gtataagaaa	gccggattat	tttctaaacg	ccgccgggtc	ttccggggct	gaaacgatag	11280
ggaatgtaat	aatcgcagtc	gatcgtgttc	tgggcgaaat	agatcccgat	gcgctgctcg	11340
tgctgggtga	taccaatagt	tgtatggcgg	tactgcctgc	aaaacggcgt	aagataccga	11400
cctttcatat	ggaagcaggc	aatcgctgtt	tcgatatgcg	tgtgcctgaa	gagataaatc	11460
ggcgcattgt	cgatcataca	gctgatgtaa	atttgacctc	tagtacaatt	gcgcgtgatt	11520
atctcttgcg	tgaaggactt	tctccagaca	tggttatcaa	gactggtagc	cctatgttgc	11580
aagttctcga	gcactatcgt	gacgggatcg	agtcctccga	tattcttgaa	aggctcgggt	11640
tgaaaacaga	gcggttcttt	gtcgtgagtg	cgcaccgaga	ggaaaacata	gattcggata	11700
agaatttctt	gaagttgggt	tctatgctca	acgctgtggc	agaaaagtac	tcgctgcccc	11760
tcacgcgtatc	aactcaccct	agaacaaaaa	agagaattga	ggcgacggag	gcaaagtttc	11820
acgaggggat	taaactgctg	aaacccctcg	gctttaagga	ttacaataaa	ctgcaaatta	11880
cagccaaggc	agttatttct	gacagtggga	ccatcagtga	ggagtcttca	atactgaatt	11940
ttcccgcttt	gaatattcgt	gaggctcatg	aacgcccaga	aggcatggaa	gaggctgtgg	12000
tgatgatggg	cggactggat	tcggatcgag	tactacaagc	actcgagggt	ttggagggac	12060
agaggcgcg	cgcagagcgc	atgttacgct	tggtcgctga	ctatagcatg	cccaacgttt	12120
ctgaaaagat	tgttcgcata	gttcatagct	atcgggacta	tgtcatgcga	actgtctgga	12180
aaaaatatta	acttgaggcg	tggagttgat	ggcaaggata	tttgtggttt	ctgagtatgt	12240
cgggtccaat	cagaactcca	cgggatacta	ttgggagaag	ataataggaa	agatgcagcg	12300
ggagtttggt	gggctaaccg	taattttccc	gctgaccgca	ggtgaaaccc	cgctgtgggt	12360
ttcaccttcc	gttgagcaag	aatgctttaa	gtttccgagg	agcaataaga	ataggctcct	12420
ttctagagga	ttggcgcaga	tttttcaggc	gtttctgttc	tcagtaaaat	tgacttctcg	12480
tgccagacga	ggagatgtgg	tattgagtgg	aaccaaccct	gctcttctac	tgatgacggt	12540
tcccttgcta	aggtatgcc	tcggtttcaa	gtgggtgctg	ctgggtgcatg	atgtgtttcc	12600
cgagaacttg	gtgcgggcgg	gcgttctgaa	gaaagatagt	attgcctacc	ggcttctacg	12660
tcgtctcttt	tctttcattt	actcatccgc	tgatcgtcta	gtcgtaatag	ggcgcgatat	12720
ggaagctctt	atgaaagaga	aggtgaatga	cccgcgatct	ttggtcttta	tttcgaattg	12780

ggcctgtgag	aaagaggttt	tcccagtacc	gagagaggat	gctcctttta	tcaatattcc	12840
tgaatggaaa	ggtaaaagg	ttttccaatt	ttttggtaat	gtcggtcgat	tacaaggtat	12900
agaaaacata	ctttctgcta	ttcagttggt	taaaaacgag	aaggcggctt	ttgcttttat	12960
tggagatggt	gccttggtcg	acagtgtaaa	aaaacacgcg	ctggaagatc	agtgtgctcg	13020
gttgaggtat	tttggaaggc	tgccattagc	cgaaaagaat	tttggtttgg	ctgcctgtga	13080
cgttgcccta	gttaccttag	aagaaggaat	gttcgggctt	ggggttccca	gcaaggcata	13140
tttctccatg	gcagcagaca	aaccgattct	agctgtcatg	gaaaaagggg	ctgaaatctc	13200
cogtataata	gatgagaccg	gaatcggttg	gaactgtccg	ccgaatgatc	cggttgcttt	13260
ggcaagattg	atcgatgaga	tttgtgaact	cgacttgtct	agtttaggcg	gagtcccgcg	13320
gagtgtcctt	cagcaaaatt	attctgaata	tatttcattg	gaaaaattcg	ctgcctgtgt	13380
tcgaccgctt	ctgtctgagt	cgaaaatatg	atgaagggtc	tggtaacggg	ggctagcggg	13440
tttgtcggga	gtgcgctttg	caggtcgctt	gctgcgcgcc	cctttcaggt	tgtcggacaa	13500
gtacgatccc	tgtacaatcc	cgttacgggg	gttgagtatg	ttcgagcgga	gctgaaagag	13560
agcactaagc	ttgatgctgc	gctgcggggg	gttgaatgtg	tagttcatct	agctggacga	13620
gcccatactt	ttggaaggca	gcgtgattca	ctagatattt	ttcggaaggt	gaatcgcgat	13680
gctactctgg	cgcttgctcg	gcaggcgatc	gaagcatctg	taaagcgttt	cattttttgtt	13740
agttctattg	gtgtaaattg	cgctttaacc	aaagaaaagc	ccttcgatga	gaactccaag	13800
ccggctcctc	atgcagaata	tgcgatttca	aagtttgagg	ctgaagtagc	gcttcgggag	13860
cTTTTcaagc	attcctcaac	agaacttggt	atcgtcaggc	ctccactcgt	ttacgactgg	13920
aaagctcctg	gaaatttctc	gcgattgttg	aagctgggtg	cctcgggact	tcctcttcca	13980
tttggttgca	tagataaccg	acgaagtttt	gtttctctgg	ataatttagt	tgactttcta	14040
gcttgctgta	tgacgcaccc	ttctgctgcc	ggcgaactgt	ttttggtatc	cgatgggtcag	14100
gagatttcta	ccaagcaact	ggtgactgcg	cttgctgcgg	gaatggggcg	tcgccccatc	14160
atgtggcctg	ttcctagggt	tattctgagg	tttcttaaat	tagtaggaaa	gggtgggtta	14220
tacactcagt	tatgctgctc	actagaggtc	gactcgtcga	aaggcaggct	tttgcttggt	14280
tgggaacccc	gcaagagcac	cctttccgcg	ttggaagatg	ttggtagaat	atatgtcaaa	14340
cgtactgaat	gattatctgc	aggcgctttg	ctactagcat	ggcgtaccac	gcagaacaat	14400
cgaatagaac	cctgttgaa	gggtgagagt	atTTTTgggg	ataaatttat	aaatggaaga	14460
atggtatTTg	ttactcgctg	cagctggggg	ttcgggactg	cttacaggcc	tcttgcgctg	14520
ttatgcctta	gcgaggagct	tacttgacac	ccctaactct	cgaagttccc	atgtcgttcc	14580
cactccacgc	ggaggagggg	tcgccattgt	agttactttt	tgtctcatgc	tgccattttg	14640

ggctgtactg	ggaaatatct	catggggccgt	gtcctgggct	ttactttctcg	ctggcggcg	14700
ggttgccatt	attggattca	tggatgatca	cggtcatac	gccgcacgct	ggcgtctgct	14760
gggacatttt	agtgcagcct	tggctctcatt	gtactttttg	aatggcatac	caccatttca	14820
gattgttggt	gtcagttggg	acctggggtg	gttcggagga	cttctctttg	ctttctatct	14880
cgtgtggttg	ctgaatctct	ataacttcat	ggatgggatc	gatggacttg	ctagccttca	14940
ggccattttt	gtctgtgttg	gtggggcatt	attatactgg	ctgaatggcc	aactgacgca	15000
ggctttgctc	cccttatcgc	tagcttttgc	cgtttttgga	ttcttgttct	ggaattttcc	15060
acccccaaaa	attttcatgg	gagatgcggg	tagtggctct	ctggggattg	tttttaggaat	15120
tctttccatt	catgccatgt	ggatgaatac	gaatttttcc	tgggcatggt	tggctctggt	15180
aggcgttttc	atcgctgatg	cgacctatac	cctgattcgt	cgcttgctga	gaggggacaa	15240
ggtgtatgag	gctcatcgaa	gccatgccta	tcaatacgca	agccgatact	atggaaagca	15300
tgctcctggt	acgattggcg	tcacggcatt	gaacgtcatc	tggctcctcc	ctatagcctt	15360
gttggtcggg	agtgggtctc	tagagccttt	gatgggcatac	gtcatagcct	acgtccctct	15420
cgtttttctg	gcagtgaggt	tcaaggcggg	taagctagag	tcgtccgctc	aggcctaaag	15480
gagtagggga	atgctagatc	gtttaagagt	aaagttgtta	tccatgcctc	gtcgtcggaa	15540
acgtttgctt	caagtggcta	cggatatcct	tctggtatgg	ctgtctctgt	ggctcgtctt	15600
tgtggtccgt	ctaggcacag	acgatatgat	cgacgtgttc	ggcgagcatg	catggccttt	15660
catcactgcg	ccggtcatcg	ccattccact	attcattcgc	ttcggcatgt	atcgcgcggt	15720
gatgcgctat	ctcggtaacg	acgcattgat	cgccatcgcc	aaggcggtga	ccatctcggc	15780
tctggtgctg	tcgctggtgg	tgtactggta	tcgtggcgcg	ccggcgccgg	tgccgcgttc	15840
cctggtgttc	aactactggg	ggttgagcat	gctgctgatc	ggcggccttg	gtctggccat	15900
gcgccagtat	ttcatgggcg	actggtactc	tgctgtgcag	tcggtaccat	ttctcaaccg	15960
ccaggatggc	ctgccagggg	tggttatcta	tggggcgggg	gcggccggca	accagttggt	16020
tgccggcgtt	cgtctcggtc	ggcgatgcg	tccggtggcg	ttcatcgatg	acgacaagca	16080
gatcgccaac	cgggtcattg	ccggtctgcg	ggtctatacc	gccaagcata	tccgccagat	16140
gatcgacgag	acgggcgcgc	aggaggttct	cctggcgatt	ccttccgcca	ctcgggcccg	16200
gcgccgagag	attctcgagt	ccctggagcc	gttcccgtg	cacgtgcgca	gcatgcccg	16260
cttcatggac	ctggccagcg	gccgggtcaa	ggtggatgac	ctgcaggagg	tggacatcgc	16320
tgacctgctg	gggcgcgaca	gcgtcgcacc	gcgcaaggag	ctgctggaac	ggtgcatccg	16380
cggtcaggtg	gtgatggtga	ccggggcggg	cggctctatc	ggttcggaac	tctgtcggca	16440
gatcatgagt	tgttcgcta	gcgtgctgat	cctgttcgaa	cacagcgaat	acaacctcta	16500
tagcatccat	caggaactgg	agcgtcggat	caagcgcgag	tcgctttcgg	tgaacctggt	16560

gccgatcctc	ggttcggtgc	gcaatcccga	gcgcttggtg	gacgtgatgc	gtacctggaa	16620
ggtcaatacc	gtctaccatg	cggcggccta	caagcatgtg	ccgatcgtcg	agcacaacat	16680
cgccgagggc	gttctcaaca	acgtgatagg	caccttgcat	gcggtgcagg	ccgcggtgca	16740
ggtcggcgtg	cagaacttcg	tgctgatttc	caccgacaag	gcggtgcggc	cgaccaatgt	16800
gatgggcagc	accaagcgcc	tggcggaaat	ggtccttcag	gcgctcagca	acgaatcggc	16860
gcggtgctg	ttcggcgacc	ggaaggacgt	gcatcacgtc	aacaagaccc	gtttcaccat	16920
ggtccgcttc	ggcaacgtcc	tcggttcgtc	cggttcggtc	attccgctgt	tcgcgcagca	16980
gatcaagcgc	ggcggcccgg	tgacggtcac	ccaccgcagc	atcacccggt	acttcatgac	17040
cattcccag	gcggcgcagt	tggtcatcca	ggccggttcg	atggggcagg	gcggagatgt	17100
attcgtgctg	gacatggggc	cgcgggtgaa	gatcctggag	ctcgccgaga	agatgatcca	17160
cctgtccggc	ctgagcgtgc	gttccgagcg	ttcgcccat	ggtgacatcg	ccatcgagtt	17220
cagtggcctg	cgtcctggcg	agaagctcta	cgaagagctg	ctgatcggtg	acaacgtgaa	17280
tcccaccgac	catccgatga	tcatgcgggc	caacgaggaa	cacctgagct	gggaggcctt	17340
caaggtcgtg	ctggagcagt	tgctggccgc	cgtggagaag	gacgactact	cgcgggttcg	17400
ccagttgctg	cgggaaaccg	tcagcggcta	tgcgctgac	ggtgaaatcg	tcgactggat	17460
ctatcgccag	aggcggcgag	aacctgagt	catcgttctc	cggaaaaggc	cgcctagcgg	17520
ccttttttgt	tttctccgta	cgatgtttcc	ggtgccggac	caggaagcga	ctgctttgct	17580
ggggctgtcg	atccaggtgc	gttccacggc	gataaggtgg	tttcgtggat	gggcaacatg	17640
tgcgaaggt	aaagtcagcc	gcattgttga	attcatcgaa	aaaccggatc	agccacaaac	17700
gctggaatca	gacatcatgg	ccgtggggccg	ttatgtgctt	tctgccgata	tttggccgga	17760
acttgaacgc	actcagccag	gtgcatgggg	acgtattcag	ctgactgatg	ccattgccga	17820
actggcgaaa	aaacagtctg	ttgacgccat	gctgatgact	ggtgacagct	acgactgtgg	17880
taaaaaaatg	ggttatatgc	aggcgtttgt	gaagtatgga	ctacgcaacc	tgaaagaagg	17940
agcgaagttc	cgcaaaggta	ttgagaaatt	gcttagcgag	taagttaa	aaatagacgc	18000
ccttataggg	cgtaataaca	aataacggta	gtcaacattc	gacgcgggtga	tgacagatatg	18060
cccggaatgc	tgataccgtt	ttttcattct	aaaaaactca	tcatttcatt	gagttaacta	18120
caaaatttag	cactgttttt	tataatgttt	cttcttgttt	ctggcatcaa	ttggtaagat	18180
aattagtgtt	tgagtttaga	ggctttgcgg	cagagaagcg	gagcttaaca	cgtctgtgag	18240
agtacgcagt	gcactggtag	ctgtaaagcc	agtggcggta	gcgtgtttaa	ataaatacat	18300
tagtaatact	acatattaca	tcattgtagg	ctatttaagc	gctacatgat	aagcgacagc	18360
gctagcaatc	aaatctttta	aagttacttc	tcaggaatag	taaaaggagg	acagctatgt	18420

tgaaaaaaga gtattttaaaa aacccttatt tagttttggt tgcatgatt atattagctt	18480
atgttttttag tgtattttgc aggttttatt gggtttggtg ggcaagtgag tttaatgagt	18540
attttttcaa taatcagtta atgatcattt caaatgatgg ctatgctttt gctgagggcg	18600
caagagatat gatagcaggt tttcatcagc ctaatgattt gagttattat ggatcttctt	18660
tatccgcgct tacttattgg ctttataaaa tcacacctt ttcttttgaa agtatcattt	18720
tatatatgag tactttttta tcttctttgg tggtgattcc tactattttg ctagctaacg	18780
aatacaaacg tcctttaatg ggctttgtag ctgctctttt agcaagtata gcaaacagtt	18840
attataatcg cactatgagt gggatttatg atacggatat gctggtaatt gttttgccta	18900
tgtttatttt attttttatg gtaagaatga ttttaaaaaa agactttttt tcattgattg	18960
ccttgccggt atttatagga atttatctt ggtggtatcc ttcaagttat actttaaatg	19020
tagctttaat tggacttttt ttaatttata cacttatttt tcatagaaaa gaaaagattt	19080
tttatatagc tgtgattttg tcttctctta ctctttcaaa tatagcatgg ttttatcaaa	19140
gtgccattat agtaatactt tttgctttat tcgccttaga gcaaaaacgc ttaaatttta	19200
tgattatagg aatttttagt agtgcaactt tgatattttt gattttaagt ggtggggttg	19260
atcctatact ttatcagctt aaattttata tttttagaag tgatgaaagt gcgaatttaa	19320
cgaggggctt tatgtatttt aatgtcaatc aaaccataca agaagttgaa aatgtagatc	19380
ttagcgaaat tatgcgaaga attagtggta gtgaaattgt ttttttgtt tctttgtttg	19440
gttttgtagt gcttttgaga aaacataaaa gtatgattat ggctttacct atattggtgc	19500
ttgggttttt agccttaaaa ggggggctta gatttaccat ttattctgta cctgtaatgg	19560
ccttaggatt tggtttttta ttgagcgagt ttaaggctat aatggttaaa aaatatagcc	19620
aattaacttc aaatgtttgt attgtttttg caactatttt gacttttagct ccagtattta	19680
tccatattta caactataaa gcgccaacag tttttctca aaatgaagca tcattattaa	19740
atcaattaaa aaatatagcc aatagagaag attatgtggt aacttggtgg gattatggtt	19800
atcctgtgcg ttattatagc gatgtgaaaa ctttagtaga tgggtgaaag catttaggta	19860
aggataattt tttcccttct tttgctttta gcaaagatga acaagctgca gctaatatgg	19920
caagacttag tgtagaatat acagaaaaaa gcttttatgc tccgcaaaat gatattttaa	19980
aaacagacat tttgcaagcc atgatgaaag attataatca aagcaatgtg gatttgtttc	20040
tagcttcatt atcaaacctt gattttaaaa tcgatacgcc aaaaactcgt gatattttatc	20100
tttatatgcc cgctagaatg tctttgattt tttctacggt ggctagtttt tcttttatta	20160
atttagatac aggagttttg gataaacctt ttaccttag cacagcttat ccacttgatg	20220
ttaaaaatgg agaaatttat cttagcaacg gagtggtttt aagcgatgat tttagaagtt	20280
ttaaaaatagg tgataatgtg gtttctgtaa atagtatcgt agagattaat tctattaaac	20340

aaggtgaata caaaatcact ccaattgatg ataaggctca gttttatatt ttttatttaa	20400
aggatagtgc tattccttac gcacaattta ttttaatgga taaaaccatg tttaatagt	20460
cttatgtgca aatgtttttt ttaggaaatt atgataagaa tttatttgac ttggtgatta	20520
attctagaga tgctaagggt tttaaactta aaatttacct atacgatgtt ccagattacg	20580
cttaaacaatg tgaattc	20597

<210> 20

<211> 8815

<212> ADN

<213> Sequência Artificial

<

220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 20

agtcgacctg caggatgaga gtagaaaata ataatgtttc tgggcacaaac catgacccgg	60
aacagattga tttgattgat ttactagtgc agttgtggcg tggcaagatg acaatcatca	120
tttccgtcat tgtggctatt gccctagcta ttggatattt ggcagtagcg aaggagaaat	180
ggacgtcaac agcaattatc actcagcccg atgtggggca aattgctggc tataacaatg	240
ccatgaatgt tatctatggt caggctgcac cgaaagtatc ggatttgcag gagacgttaa	300
ttggtcgctt cagttctgcc ttctctgcat tagcagaaac gctggataat caggaagaac	360
cagaaaaact taccatogaa ctttctgtta agaaccagca attaccattg actgtttctt	420
atgttgggca aactgcagag ggcgcacaaa tgaagttggc ccaatacatt cagcaagttg	480
acgataaagt gaatcaagag ttagaaaagg atctcaagga caacattgct ctgggacgga	540
aaaacttgca ggactcttta agaacgcagg aagtggttgc gcaggagcag aaagatctgc	600

gtatccgtca gattcaggaa gcgttgagc atgcgaatca ggcgcaggtg aaaaaaccgc	660
agattcaaca gactggcgaa gatatacac aagatacgtt gttccttttg gggagcgaag	720
cgctggagtc gatgattaag catgaggcga cccgtccgtt ggtgttctca ccaaactact	780
atcagactcg tcaaaacctg cttgatatac aaagcttaaa ggttgatgat cttgatattc	840
atgcttaccg ctatgtaatg aaaccgacgt tacctattcg tcgtgatagc ccgaaaaagg	900
caattacctt gattctggcg gtgctgctgg gtggcatggt tggcgcgagg attgtgctgg	960
ggcgtaatgc tctacgcaat tacaacgcga agtaacctgc aggcatacaa gcttctgttt	1020
tggcggatga gagaagaaat tcgtcgcccg ccataaactg ccaggcatca aattaagcag	1080
aaggccatcc tgacggatgg cctttttgcg tttctacaaa ctcttcctgt ctagcaggtg	1140
gcacttttcg gggaaatgtg cgcggaaccc ctatttggtt atttttctaa atacattcaa	1200
atatgtatcc gctcatgcta gaaatatttt atctgattaa taagatgac ttcttgagat	1260
cgttttggtc tgcgcgtaat ctcttgctct gaaaacgaaa aaaccgcctt gcagggcggt	1320
ttttogaagg ttctctgagc taccaactct ttgaaccgag gtaactggct tggaggagcg	1380
cagtcaccaa aacttgctct ttcatgttag ccttaaccgg cgcatactt caagactaac	1440
tcctctaaat caattaccag tggctgctgc cagtggctgt tttgcatgtc tttccgggtt	1500
ggactcaaga cgatagttac cggataaggc gcagcggctc gactgaacgg ggggttcgtg	1560
catacagtc agcttgagc gaactgccta cccggaactg agtgtcaggc gtggaatgag	1620
acaaacgcgg ccataacagc ggaatgacac cggtaaaccg aaaggcagga acaggagagc	1680
gcacgagggg gccgccaggg gaaacgcctg gtatctttat agtcctgtcg ggtttcgcca	1740
ccactgattt gagcgtcaga tttcgtgatg cttgtcaggg gggcggagcc tatggaaaaa	1800
cggctttgac gcggccctct cacttcctg ttaagtatct tcctggcatc ttccaggaaa	1860
ttccgcccc gttcgtaagc catttcctg cgcgcagtc gaacgacga gcgtagcgag	1920
tcagtgagcg aggaagcggg atatatcctg tatcacatat tctgctgacg caccgggtgca	1980
gccttttttc tcctgccaca tgaagcactt cactgacacc ctcatcagtg ccaacatagt	2040
aagccagtat aactccgct agcgtgatg tccggcgggt cttttgcctg tacgcaccac	2100
cccgtcagta gctgaacagg agggacagct gatagaaaca gaagccactg gagcacctca	2160
aaaacaccat catacactaa atcagtaagt tggcagcatc acccgacgca ctttgcgccg	2220

aataaagtgt aggcctggagc tgcctcgaag ttccctatact ttctagagaa taggaacttc	2280
ggaataggaa cttcaagatc ccctcacgct gccgcaagca ctcagggcgc aagggctgct	2340
aaaggaagcg gaacacgtag aaagccagtc cgcagaaacg gtgctgaccc cggatgaatg	2400
tcagctactg ggctatctgg acaagggaaa acgcaagcgc aaagagaaaag caggtagctt	2460
gcagtgggct tacatggcga tagctagact gggcggtttt atggacagca agogaaccgg	2520
aattgccagc tggggcgccc tctggtaagg ttgggaagcc ctgcaaagta aactggatgg	2580
ctttcttgcc gccaaaggatc tgatggcgca ggggatcaag atctgatcaa gagacaggat	2640
gaggatcgtt tcgcatgatt gaacaagatg gattgcacgc aggttctccg gccgcttggg	2700
tggagaggct attcggctat gactgggcac aacagacaat cggctgctct gatgcgcgcg	2760
tgttccggct gtcagcgagc gggcgcccgc ttctttttgt caagaccgac ctgtccgggtg	2820
ccctgaatga actgcaggac gaggcagcgc ggctatcgtg gctggccacg acgggcgttc	2880
cttgccgagc tgtgctcgac gttgtcactg aagcgggaag ggactggctg ctattgggcg	2940
aagtgccggg gcaggatctc ctgtcatctc accttgctcc tgccgagaaa gtatccatca	3000
tggctgatgc aatgcggcgg ctgcatacgc ttgatccggc tacctgccc a ttgaccacc	3060
aagcgaaaca tcgcatcgag cgagcacgta ctcgatgga agccggtctt gtcgatcagg	3120
atgatctgga cgaagagcat caggggctcg cgccagccga actgttcgcc aggcctcaagg	3180
cgcgcatgcc cgacggcgag gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata	3240
tcatggtgga aaatggccgc ttttctggat tcatcgactg tggccggctg ggtgtggcgg	3300
accgctatca ggacatagcg ttggctaccc gtgatattgc tgaagagctt ggcggcgaat	3360
gggctgaccg cttectcgtg ctttacggta tcgccgtcc cgattcgag cgcctcgct	3420
tctatgcct tcttgacgag ttcttctgag cgggactctg gggttcgaaa tgaccgacca	3480
agcgacgcc aacctgccat cagcagattt cgattccacc gccgccttct atgaaaggtt	3540
gggcttcgga atcgttttcc gggacgcgg ctggatgac ctccagcgcg gggatctcat	3600
gctggagtto ttgcgccacc ccagcttcaa aagcgcctctg aagttcctat actttctaga	3660
gaataggaac ttcggaatag gaactaagga ggatattcat atggtttttt taaggcagtt	3720
attggtgccc ttaaaccgct ggtgctacgc ctgaataagt gataataagc ggatgaatgg	3780

cagaaattcg aaagcaaatt cgacccggtc gtcggttcag ggcagggtcg ttaaatagcc	3840
gcttatgtct attgctggtt taccggttta ttgactaccg gaagcagtgt gaccgtgtgc	3900
ttctcaaattg cctgaggcca gtttgctcag gctctccccg tggaggtaat aattgacgat	3960
atgatcattt attctgcctc ccagagcctg ataaaaacgg ttagcgcttc gttaatacag	4020
atgtagggtg tccacagggt agccagcagc atcctgcgat gcagatccgg aacataatgg	4080
tgcaggggcg ttgtttcggc gtgggtatgg tggcaggccc cgtggccggg ggactgttgg	4140
gcgctgccgg cacctgtcct acgagttgca tgataaagaa gacagtcata agtgcggcga	4200
cgatagtcat gccccgcgcc caccggaagg agctaccgga cagcgggtgc gactgttgta	4260
actcagaata agaaatgagg ccgctcatgg cgttgactct cagtcatagt atcgtggtat	4320
caccggttgg ttccactctc tgttgccggc aacttcagca gcacgtaggg gacttccgcg	4380
tttccagact ttacgaaaca cggaaaccga agaccattca tgttggtgct caggtcgcag	4440
acgttttgca gcagcagtcg cttcacgttc gctcgcgtat cgggtgattca ttctgctaac	4500
cagtaaggca accccgccag cctagccggg tcctcaacga caggagcacg atcatgcgca	4560
cccgtggcca ggacccaacg ctgcccgaga tgcgccgcgt gcggtgctg gagatggcgg	4620
acgcgatgga tatgttctgc caagggttgg tttgcgcatt cacagttctc cgcaagaatt	4680
gattggctcc aattcttggg gtggtgaatc cgttagcgag gtgcgccgg cttccattca	4740
ggtcgagggtg gcccggctcc atgcaccgcg acgcaacgcg gggaggcaga caaggatatg	4800
ggcggcgctt acaatccatg ccaaccctt ccatgtgctc gccgaggcgg cataaatcgc	4860
cgtgacgatc agcgggtccag tgatcgaagt taggctggta agagccgcga gcgatccttg	4920
aagctgtccc tgatggtcgt catctacctg cctggacagc atggcctgca acgcgggcat	4980
cccgatgccg ccggaagcga gaagaatcat aatggggaag gccatccagc ctgcgctgc	5040
gaacgccagc aagacgtagc ccagcgcgtc ggccaattcg cgctaactta cattaattgc	5100
gttgcgctca ctgcccgtt tccagtcggg aaacctgtcg tgccagctgc attaatgaat	5160
cggccaacgc gcggggagag gcggtttgcg tattgggcgc cagggtggtt tttcttttca	5220
ccagtgcagc gggcaacagc tgattgcctt tcaccgcctg gccctgagag agttgcagca	5280
agcgggtccac gctggtttgc ccagcaggc gaaaatcctg tttgatggtg gttgacggcg	5340
ggatataaca tgagctgtct tcggtatcgt cgtatcccac taccgagata tccgcaccaa	5400

cgcgagccc	ggactcggta	atggcgcgca	ttgcgcccag	cgccatctga	tcgttggcaa	5460
ccagcatcgc	agtgggaacg	atgccctcat	tcagcatttg	catggtttgt	tgaaaaccgg	5520
acatggcact	ccagtcgcct	tcccgttccg	ctatcggctg	aatttgattg	cgagtgagat	5580
atttatgcca	gccagccaga	cgcagacgcg	ccgagacaga	acttaatggg	cccgctaaca	5640
gcgcgatttg	ctggtgaccc	aatgcgacca	gatgctccac	gccagtcgc	gtaccgtctt	5700
catgggagaa	aataatactg	ttgatgggtg	tctggtcaga	gacatcaaga	aataacgccg	5760
gaacattagt	gcaggcagct	tccacagcaa	tggcatcctg	gtcatccagc	ggatagttaa	5820
tgatcagccc	actgacgcgt	tgcgcgagaa	gattgtgcac	cgccgcttta	caggcttcga	5880
cgccgcttcg	ttctaccatc	gacaccacca	cgctggcacc	cagttgatcg	gcgcgagatt	5940
taatcgccgc	gacaatttgc	gacggcgcgt	gcagggccag	actggaggtg	gcaacgccaa	6000
tcagcaacga	ctgtttgccc	gccagttgtt	gtgccacgog	gttgggaatg	taattcagct	6060
ccgccatcgc	cgcttccact	ttttcccgcg	ttttcgcaga	aacgtggctg	gcctggttca	6120
ccacgcggga	aacggtctga	taagagacac	cggcatactc	tgcgacatcg	tataacgtta	6180
ctggtttcac	attcaccacc	ctgaattgac	tctcttcggg	gcgctatcat	gccataccgc	6240
gaaaggtttt	gcaccattcg	atggtgtcaa	cgtaaataga	tgcgcgttcg	ccttcgcgcg	6300
cgaattggcc	gccatgccgg	cgataatggc	ctgcttctcg	cggaaacgtt	tgggtggcggg	6360
accagtgacg	aaggcttgag	cgagggcgtg	caagattccg	aataccgcaa	gcgacaggcc	6420
gatcatcgtc	gcgctccagc	gaaagcggtc	ctcgccgaaa	atgaccacga	gcgctgccgg	6480
cacctgtcct	acgagttgca	tgataaagaa	gacagtcata	agtgcggcga	cgatagtcac	6540
gccccgcgcc	caccggaagg	agctgactgg	gttgaaggct	ctcaagggca	tcggcggagc	6600
ttatcgactg	cacggtgcac	caatgcttct	ggcgctcaggc	agccatcgga	agctgtggta	6660
tggctgtgca	ggtcgtaaat	cactgcataa	ttcgtgtcgc	tcaaggcgca	ctcccgttct	6720
ggataatgtt	ttttgcgcgc	acatcataac	ggttctggca	aatattctga	aatgagctgt	6780
tgacaattaa	tcacgcgctc	gtataatgtg	tggaattgtg	agcggataac	aatttcacac	6840
aggaaacaga	attcgagctc	atgaaaaaga	tttggctggc	cctggcagga	ctggttctgg	6900
ccttttcagc	aagtgcagct	gaagaagcct	ttgatctgtg	gaatgagtgt	gcaaaagcat	6960
gtgtactgga	tctgaaagat	ggtgtgagat	ccagcagaat	gtcagtggat	ccagccattg	7020
cagatacaaa	tggccagggt	gtactgcatt	actctatggc	tctggaaggc	ggtaatgatg	7080

ccttgaaaact	ggccattgat	aatgcactgt	ctatcaccag	tgatggtctg	acaatcagac	7140
tggagggagg	ggtggaaccc	aataagcctg	tcagatacag	ctatacaaga	caagccagag	7200
gttcttggag	cctgaactgg	ctggtgccta	ttgggcatga	aaaaccatct	aacattaaag	7260
tttttattca	tgaactgaat	gcaggcaatc	agctgtctca	tatgagccca	atttatacca	7320
ttgaaatggg	ggatgaactg	ctggctaaac	tggccagaga	tgctacattc	tttgtcagag	7380
cccatgaatc	aaatgagatg	cagcctaccc	tggccattag	ccatgctggt	gtgagtgttg	7440
tcattggcaca	aactcagccc	aggagagaga	aaaggtggtc	tgagtggacc	agtggcaaag	7500
tgctgtgcct	gctggatcct	ctggatggtg	tttataacta	tctggcccaa	cagaggtgta	7560
acctggatga	tacctgggaa	ggtaaaatct	atagagtgtc	ggcaggtaat	ccagcaaac	7620
atgacctgga	tatcaaggat	aataacaata	gcaccctac	tgtaatcagc	catagactgc	7680
atttcccaga	gggaggttca	ctggctgccc	tgactgctca	tcaggcctgt	catctgccac	7740
tggaaaacttt	caccagacac	aggcagccaa	gaggctggga	acagctggaa	caatgtggct	7800
atccagttca	gaggctggtt	gccctgtacc	tggcagcaag	actgagctgg	aatcaggtag	7860
atcaggttat	tagaaatgca	ctggccagcc	cagggtggtg	gggtgacctg	ggtgaggcaa	7920
ttagagaaca	gcctgagcag	gccagactgg	ccctgactct	ggcagcagct	gaaagtgaaa	7980
gatttgtgag	acaggggaca	ggcaatgatg	aagcaggtgc	agctaataca	gatgttgttt	8040
cactgacttg	tcctgttgct	aaagatcaga	acaggaccaa	aggtgaatgt	gctggaccag	8100
ctgattcagg	agatgcactg	ctggagagga	actatccaac	tgggtgcagaa	ttcctgggag	8160
atggtggtga	tgtttctttt	agcaccagag	gcacacagaa	ctggactgtg	gaaagactgc	8220
tgcaggcaca	tagacagctg	gaagaaagag	gctatgtatt	tggtggctac	catggtactt	8280
tcctggaagc	agcacagtcc	attgtctttg	gaggggttag	agccagaagc	caggatctgg	8340
atgctatttg	gagaggtttt	tatattgctg	gggatccagc	cctggcctat	ggatatgcac	8400
aagatcagga	acctgatgcc	agaggcagaa	tcagaaatgg	tgccctgctg	agggtttatg	8460
ttcctaggtc	tagcctgcca	ggattttata	gaacctctct	gacctggca	gccctgaag	8520
cagcaggtga	ggtggagaga	ctgattggtc	atcctctgcc	actgagactg	gatgccatta	8580
cagggccaga	agaagaaggt	ggcagagtga	caattctggg	ttggcccctg	gcagagagga	8640
cagtagttat	tccttcagca	atccctacag	atccaaggaa	tgtgggtggg	gacctggatc	8700
catcctcaat	tccagataag	gaacaggcaa	tttcagccct	gcctgattat	gctagtcagc	8760
caggtaaacc	acctagagaa	gatctgaaac	accaccacca	ccaccactga	tctag	8815

<210> 21

<211> 99

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 21

gaattcatga gagtagaaaa taataatggt tctgggcaaa accatgaccc ggaacagatt	60
gatttgattg atttactagt gcagttgtgg cgtggcaaga tgacaatcat catttccgtc	120
attgtggcta ttgccctagc tattggatat ttggcagtag cgaaggagaa atggacgtca	180
acagcaatta tcactcagcc cgatgtgggg caaattgctg gctataacaa tgccatgaat	240
gttatctatg gtcaggctgc accgaaagta tcggatttgc aggagacggt aattggctgc	300
ttcagttctg ccttctctgc attagcagaa acgctggata atcaggaaga accagaaaaa	360
cttaccatcg aaccttctgt taagaaccag caattacat tgactgtttc ttatgttggg	420
caaactgcag agggcgacaa aatgaagttg gcccaatata ttcagcaagt tgacgataaa	480
gtgaatcaag agttagaaaa ggatctcaag gacaacattg ctctgggacg gaaaaacttg	540
caggactctt taagaacgca ggaagtgggt gcgcaggagc agaaagatct gcgtatccgt	600
cagattcagg aagcgttgca gtatgcgaat caggcgagcag tgacaaaacc gcagattcaa	660
cagactggcg aagatatcac acaagatacg ttgttccttt tggggagcga agcgctggag	720
tcgatgatta agcatgaggc gacccgtccg ttggtgttct caccaaacta ctatcagact	780
cgtcaaaaacc tgcttgatat cgaaagctta aaggttgatg atcttgatat tcatgcttac	840
cgctatgtaa tgaaaccgac gttacattt cgtcgtgata gcccgaaaaa ggcaattacc	900
ttgattctgg cgggtgctgt ggggtggcatg gttggcgagg ggattgtgct ggggcgtaat	960
gctctacgca attacaacgc gaagtaagtc gac	993

<210> 22

<211> 13

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Péptido sintético"

<400> 22

Asp Asn Asn Asn Ser Thr Pro Thr Val Ile Ser His Arg
1 5 10

<210> 23

<211> 4

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Péptido sintético"

<400> 23

Asp Gln Asn Arg
1

<210> 24
 <211> 5
 <212> PRT
 <213> Sequência Artificial

 <220>
 <221> fonte
 <223> /nota="Descrição de Sequência Artificial: Sequência
 consenso sintética"

 <220>
 <221> VARIANTE
 <222> (1) .. (1)
 <223> /substituir="Glu"

 <220>
 <221> VARIANTE
 <222> (2) .. (2)
 <223> /substituir="Arg" ou "Asn" ou "Asp" ou "Cys" ou "Gln" ou
 "Glu" ou "Gly" ou "His" ou "Ile" ou "Leu" ou "Lys" ou "Met" ou
 "Phe" ou "Ser" ou "Thr" ou "Trp" ou "Tyr" ou "Val"

 <220>
 <221> MISC_CHARACTERISTICA
 <222> (1) .. (2)
 <223> /nota=" Os residuos dados na sequência não têm preferência
 em relação àqueles nas anotações para as referidas posições "

 <220>
 <221> MOD_RES
 <222> (4) .. (4)

<223> /substituir="Arg" ou "Asn" ou "Asp" ou "Cys" ou "Gln" ou "Glu" ou "Gly" ou "His" ou "Ile" ou "Leu" ou "Lys" ou "Met" ou "Phe" ou "Ser" ou "Thr" ou "Trp" ou "Tyr" ou "Val"

<220>

<221> VARIANTE

<222> (5) .. (5)

<223> /substituir="Thr"

<220>

<221> MISC_CHARACTERISTICAE

<222> (4) .. (5)

<223> /nota="Os residuos dados na sequência não têm preferência em relação àqueles nas anotações para as referidas posições"

<220>

<221> MISC_CHARACTERISTICAE

<222> (1) .. (5)

<223> /nota="Ver descrição como apresentada para a descrição detalhada de substituições e formas de realização preferidas"

<400> 24

Asp Ala Asn Ala Ser

1

5

<210> 25

<211> 5

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Péptido sintético"

<400> 25

Asp Asn Asn Asn Ser
1 5

<210> 26

<211> 5

<212> PRT

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Péptido sintético"

<400> 26

Asp Gln Asn Arg Thr
1 5

<210> 27

<211> 2160

<212> ADN

<213> Sequência Artificial

<220>

<221> fonte

<223> /nota="Descrição de Sequência Artificial: Polinucleótido Sintético"

<400> 27

gaattcatgt	tgaaaaaaga	gtattttaaa	aacccttatt	tagttttggt	tgcgatgatt	60
atattagctt	atgttttttag	tgtattttgc	aggttttatt	gggtttggtg	ggcaagtgag	120
tttaatgagt	attttttcaa	taatcagtta	atgatcattt	caaatgatgg	ctatgccttt	180
gctgagggcg	caagagatat	gatagcaggt	tttcatcagc	ctaatgattt	gagttattat	240
ggatcttctt	tatccgcgct	tacttattgg	ctttataaaa	tcacaccttt	ttcttttgaa	300
agtatcattt	tatatatgag	tactttttta	tcttctttgg	tggtgattcc	tactattttg	360
ctagctaacg	aatacaaacg	tcctttaatg	ggctttgtag	ctgctctttt	agcaagtata	420
gcaaacagtt	attataatcg	cactatgagt	gggtattatg	atacggatat	gctggtaatt	480
gttttgccct	tgtttatatt	attttttatg	gtaagaatga	ttttaaaaaa	agactttttt	540
tcattgattg	ccttgccggt	atztatagga	atztatcttt	ggtggtatcc	ttcaagttat	600
actttaaatg	tagctttaat	tggacttttt	ttaatttata	cacttatttt	tcatagaaaa	660
gaaaagattt	tttatatagc	tgtgattttg	tcttctctta	ctctttcaaa	tatagcatgg	720
ttttatcaaa	gtgccattat	agtaataact	tttgctttat	tcgccttaga	gcaaaaacgc	780
ttaaatttta	tgattatagg	aatttttaggt	agtgcactt	tgatattttt	gatttttaagt	840
ggtgggggtg	atcctatact	ttatcagcct	aaattttata	tttttagaag	tgatgaaagt	900
gcgaatttaa	cgcagggtct	tatgtatttt	aatgtcaatc	aaaccataca	agaagttgaa	960
aatgtagatc	ttagcgaatt	tatgcgaaga	attagtggta	gtgaaattgt	ttttttgttt	1020
tctttgtttg	gttttgtag	gcttttgaga	aaacataaaa	gtatgattat	ggctttacct	1080
atattgggtg	ttgggttttt	agccttaaaa	ggggggctta	gattttaccat	ttattctgta	1140
cctgtaatgg	ccttaggatt	tggtttttta	ttgagcgagt	ttaaggctat	aatgggtaaa	1200
aaatatagcc	aattaacttc	aaatgtttgt	attgtttttg	caactatttt	gacttttagct	1260
ccagtattta	tccatattta	caactataaa	gcgccaacag	ttttttctca	aaatgaagca	1320
tcattattaa	atcaattaaa	aaatatagcc	aatagagaag	attatgtggg	aacttggggc	1380

gcttatgggtt atcctgtgcg ttattatagc gatgtgaaaa ctttagtaga tgggtggaaag	1440
catttaggta aggataatth tttcccttct tttgctttta gcaaagatga acaagctgca	1500
gctaataatgg caagacttag tgtagaatat acagaaaaaa gcttttatgc tccgcaaaat	1560
gatattttta aaacagacat tttgcaagcc atgatgaaag attataatca aagcaatgtg	1620
gatttggtttc tagottcatt atcaaaacct gatttttaaaa tcgatacgcc aaaaactcgt	1680
gatattttatc tttatatgcc cgctagaatg tctttgattt tttctacggt ggctagtttt	1740
tcttttatta atttagatac aggagttttg gataaacctt ttacctttag cacagcttat	1800
ccacttgatg ttaaaaatgg agaaatttat cttagcaacg gagtgggttt aagcgatgat	1860
tttagaagtt ttaaaatagg tgataatgtg gtttctgtaa atagtatcgt agagattaat	1920
tctattaaac aagggtgaata caaaatcact ccaattgatg ataaggctca gttttatatt	1980
ttttatttaa aggatagtgc tattccttac gcacaattta ttttaatgga taaaaccatg	2040
tttaaatagt cttatgtgca aatgtttttt ttaggaaatt atgataagaa tttatttgac	2100
ttggtgatta attctagaga tgctaagggt tttaaaacta aaatttacct atacgatgtt	2160

Lisboa, 30 de janeiro de 2018

REIVINDICAÇÕES

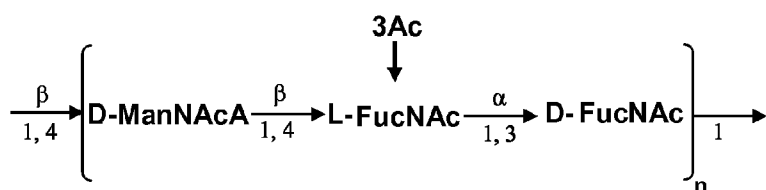
1. Organismo procariótico hospedeiro Gram-negativo, em que o referido organismo procariótico hospedeiro é *E. coli*, compreendendo:
 - (i) uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-positiva, em que a referida bactéria Gram-positiva é *Staphylococcus aureus*;
 - (ii) uma sequência nucleotídica que codifica, pelo menos, uma glicosiltransferase de uma bactéria Gram-negativa, em que a referida bactéria Gram-negativa é *Pseudomonas aeruginosa*;
 - (iii) uma sequência nucleotídica que codifica uma proteína transportadora, em que a referida proteína transportadora compreende a sequência consenso de aminoácidos D/E-X-N-Z-S/T, em que X e Z podem ser qualquer aminoácido natural, exceto prolina; e
 - (iv) uma sequência nucleotídica que codifica uma oligossacariltransferase.
2. Organismo procariótico hospedeiro da reivindicação 1, em que o referido *Staphylococcus aureus* é uma estirpe de polissacárido capsular 5 ou uma estirpe de polissacárido capsular 8.
3. Organismo procariótico hospedeiro da reivindicação 1 ou 2, em que o referido *Staphylococcus aureus* é uma estirpe resistente à meticilina.

4. Organismo procariótico hospedeiro de qualquer uma das reivindicações 1-3, compreendendo, pelo menos, duas glicosiltransferases de diferentes estirpes de bactérias Gram-positivas.
5. Organismo procariótico hospedeiro de qualquer uma das reivindicações 1-4, em que a referida proteína transportadora é a Exotoxina de *Pseudomonas aeruginosa*, a hemolisina alfa de *Staphylococcus aureus* ou o fator de aglutinação A de *Staphylococcus aureus*.
6. Organismo procariótico hospedeiro de qualquer uma das reivindicações 1-5, em que a referida oligossacariltransferase é de *Campylobacter jejuni*.
7. Método de produção de uma proteína N-glicosilada recombinante utilizando o organismo procariótico hospedeiro de qualquer uma das reivindicações 1-6.
8. Vacina de *Staphylococcus aureus* compreendendo:

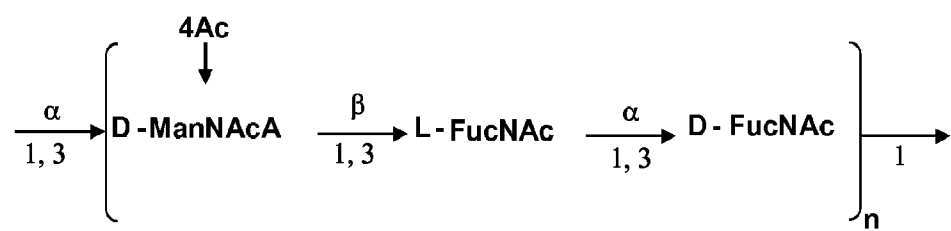
uma proteína compreendendo uma sequência consenso D/E-X-N-Z-S/T inserida, em que X e Z podem ser qualquer aminoácido natural, exceto prolina;
pelo menos, um polissacárido de *Staphylococcus aureus* ligado à referida sequência consenso por ligação N-glicosídica; e,
opcionalmente, um veículo ou adjuvante farmacêuticamente aceitável.
9. Vacina de *Staphylococcus aureus* da reivindicação 8, compreendendo um adjuvante.

10. Vacina de *Staphylococcus aureus* da reivindicação 9, em que o adjuvante contém sais de alumínio, emulsões, tais como óleo em água (MF59, AS03), lípido e combinações de sais (AS04), ISCOMS, lipossomas/virossomas, nano e micropartículas, saponinas (QS21), MPL A, citocinas, derivados de ADN ou toxinas bacterianas.
11. Vacina de *Staphylococcus aureus* de qualquer uma das reivindicações 8-10, compreendendo duas ou mais das referidas sequências consenso inseridas e dois ou mais dos referidos polissacáridos de *Staphylococcus aureus*.
12. Vacina de *Staphylococcus aureus* de qualquer uma das reivindicações 8-11, em que o referido, pelo menos, um polissacárido de *Staphylococcus aureus* compreende o polissacárido 5 capsular ou o polissacárido 8 capsular.
13. Vacina de *Staphylococcus aureus* da reivindicação 12, em que a referida proteína é a Exotoxina de *Pseudomonas aeruginosa*, ou a hemolisina alfa de *Staphylococcus aureus* ou o fator de aglutinação A de *Staphylococcus aureus*.
14. Vacina de *Staphylococcus aureus* da reivindicação 13, em que o referido, pelo menos, um polissacárido de *Staphylococcus aureus* compreende a estrutura:

(A)



ou (B)



Lisboa, 30 de janeiro de 2018

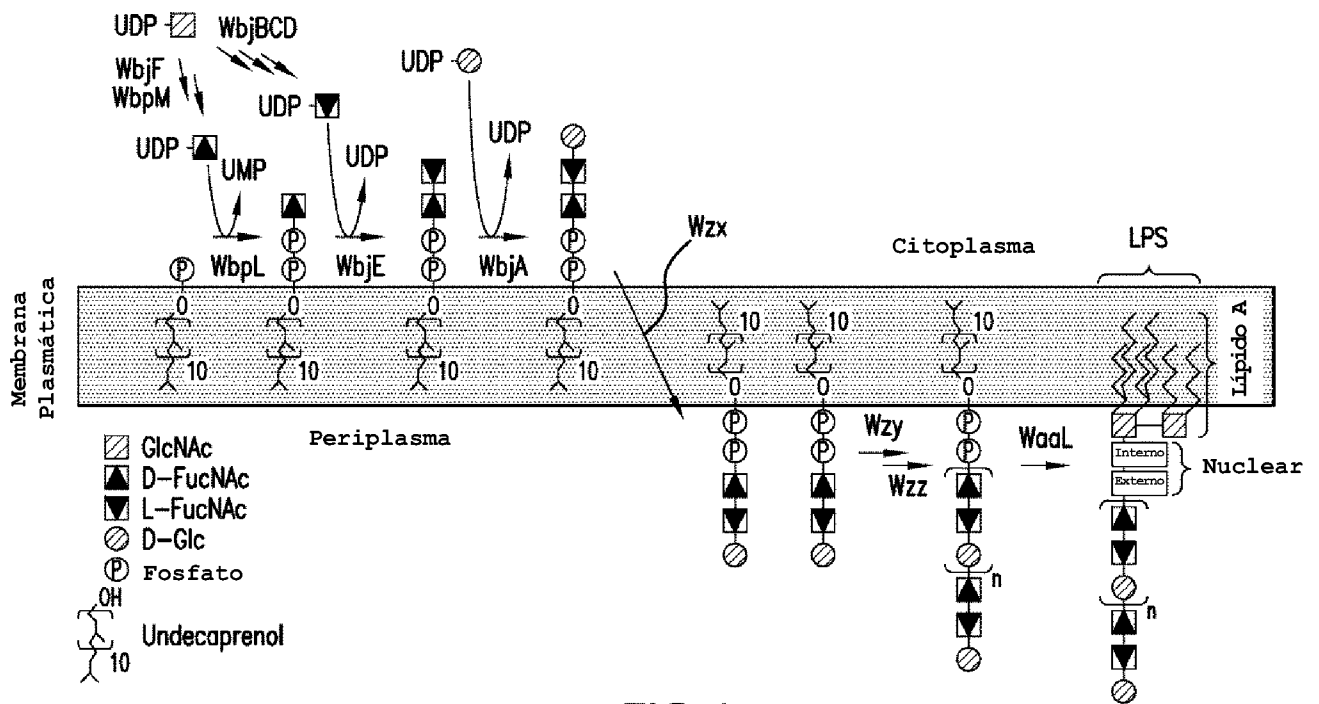


FIG.1

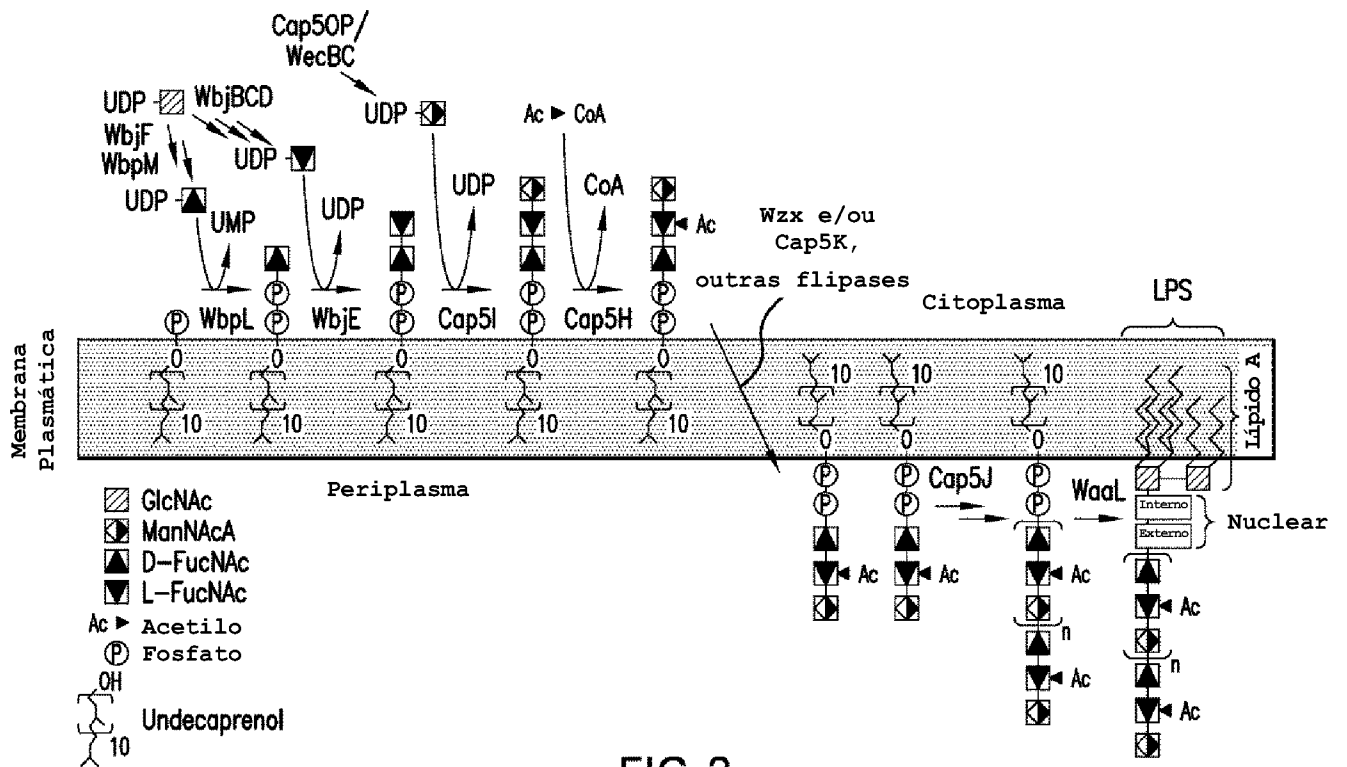


FIG.2

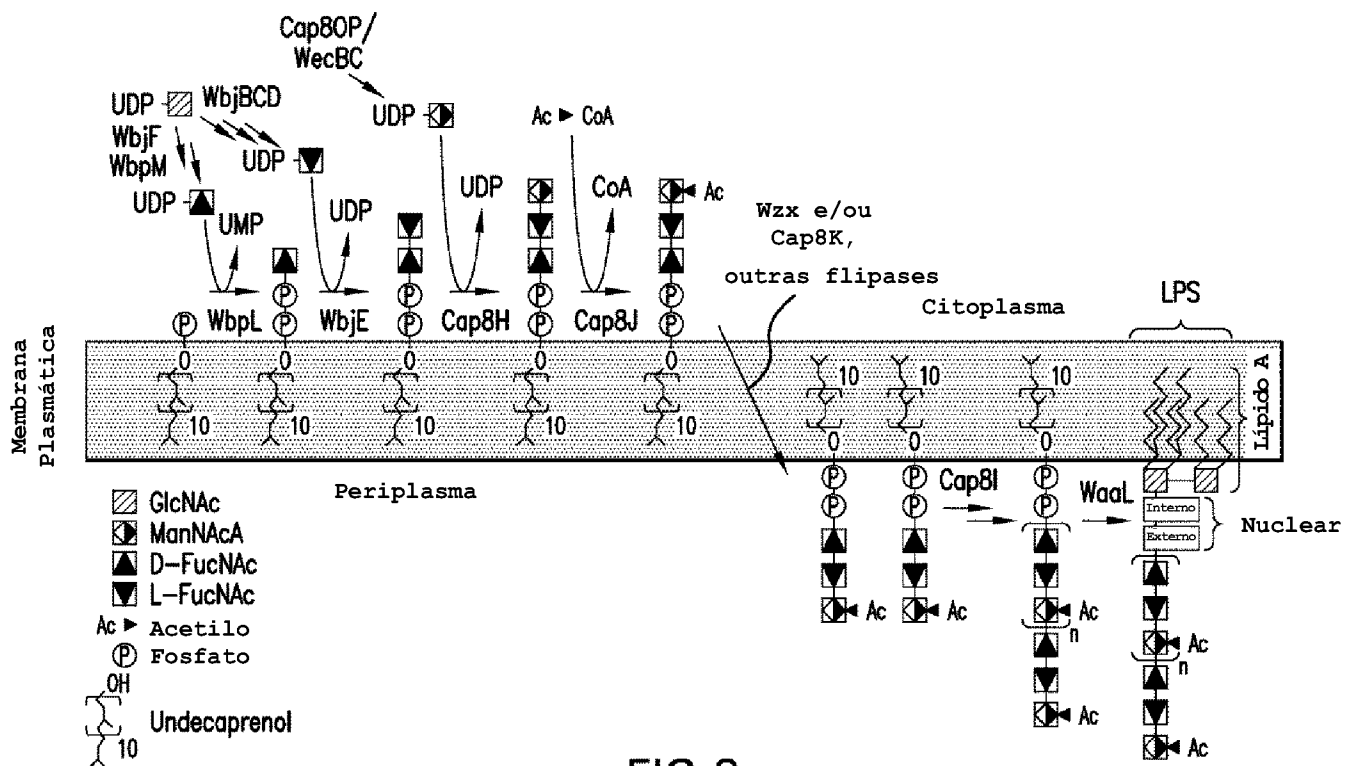


FIG.3

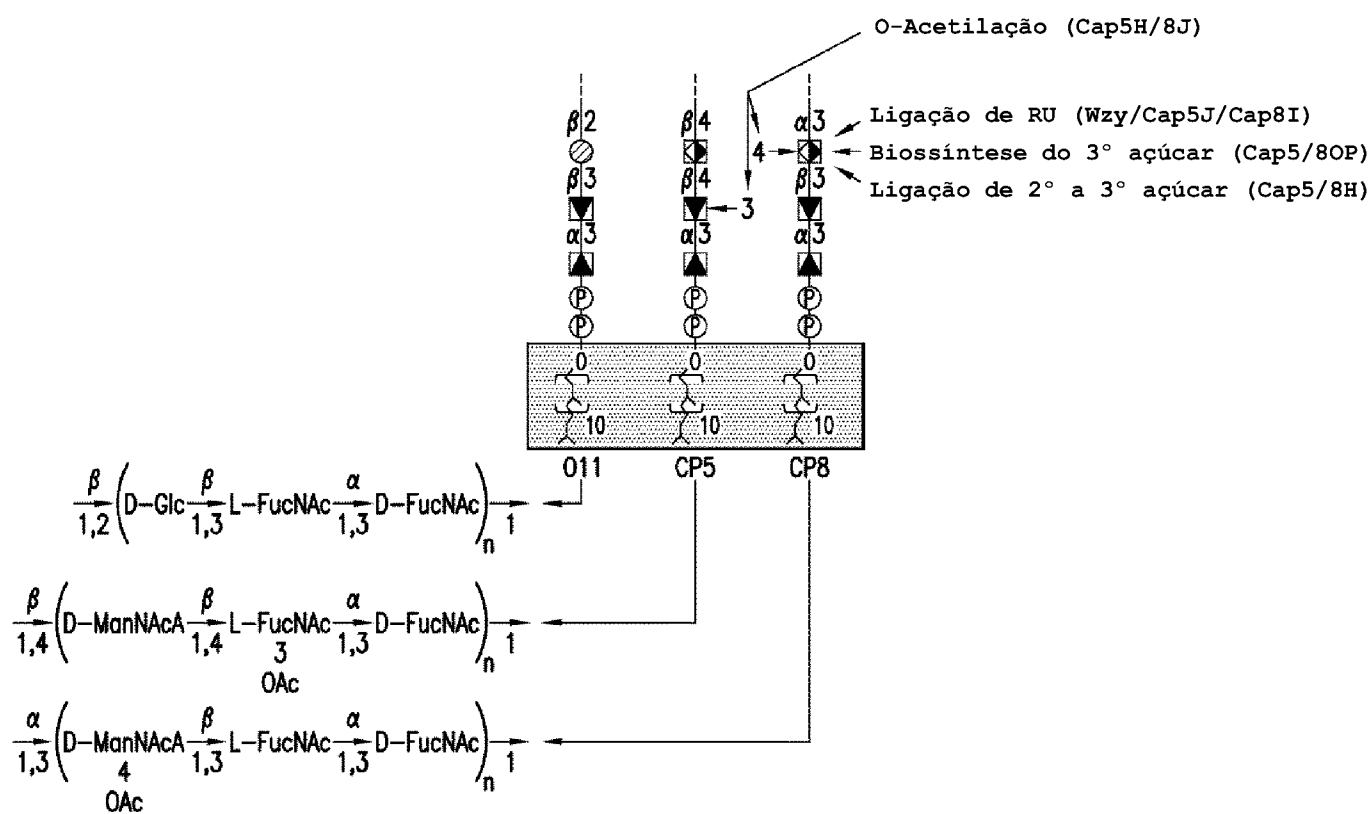


FIG.4

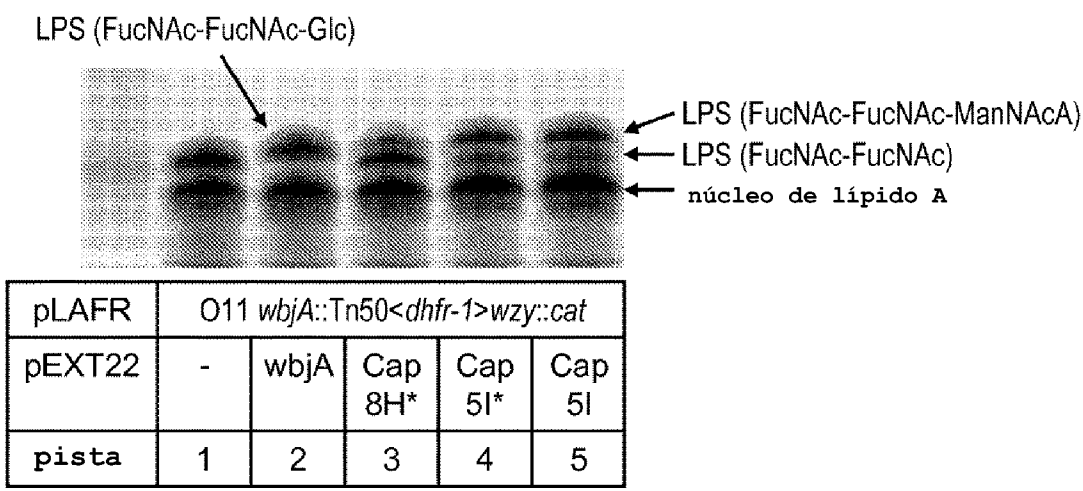


FIG.5A

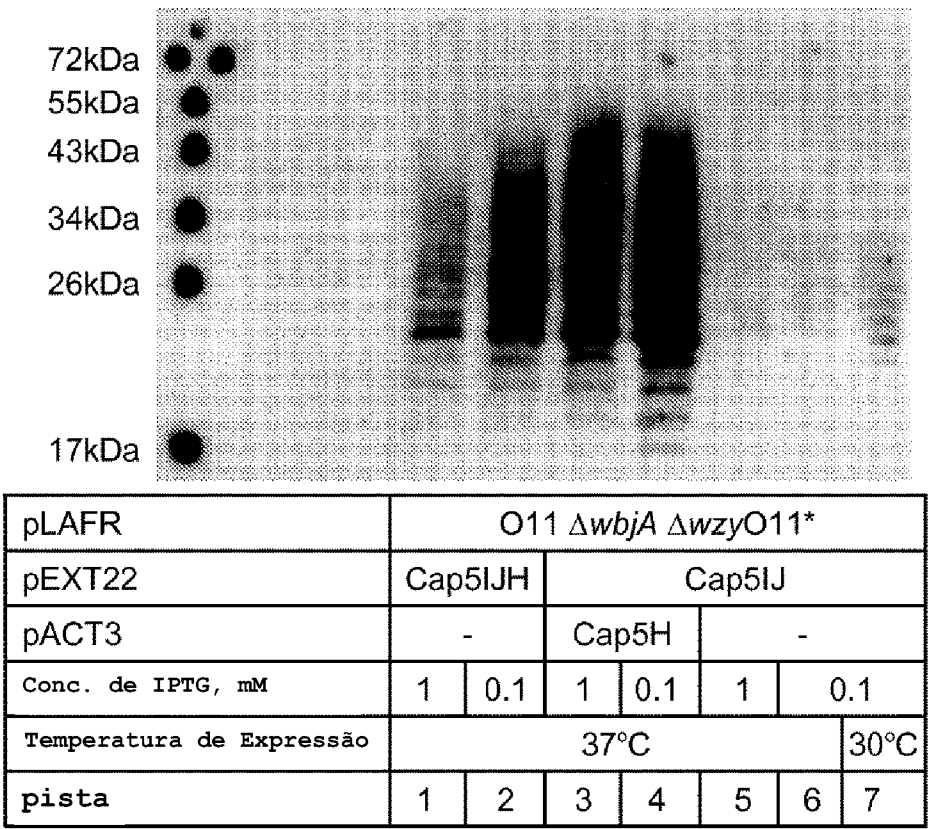


FIG.5B

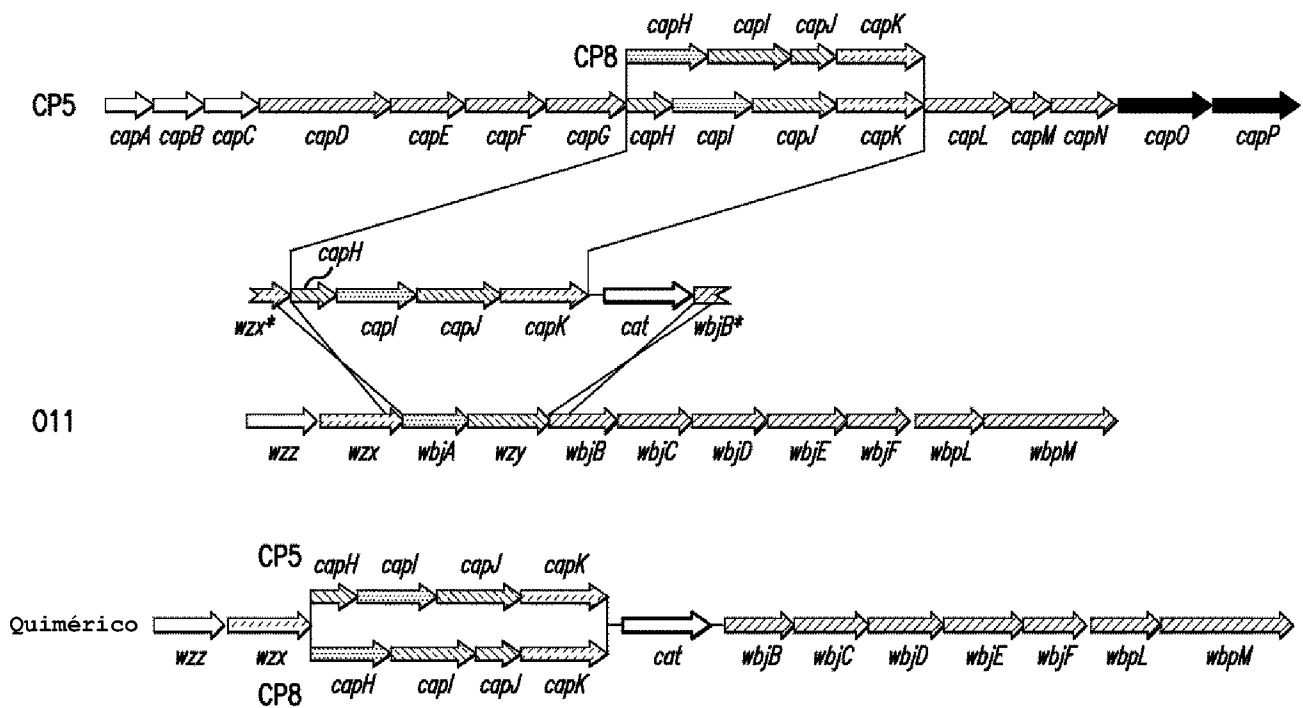
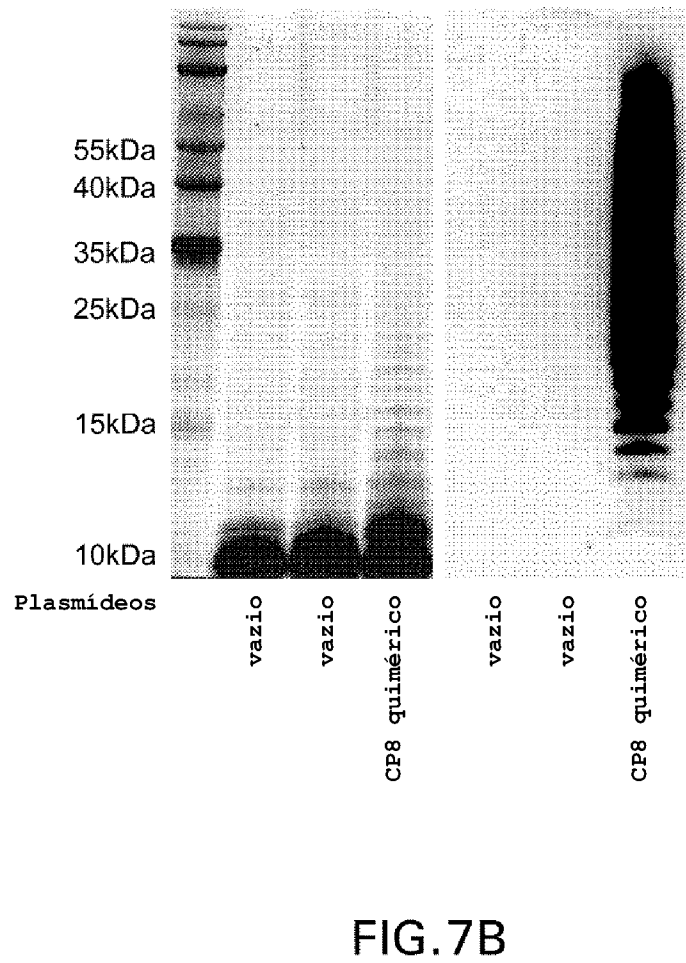
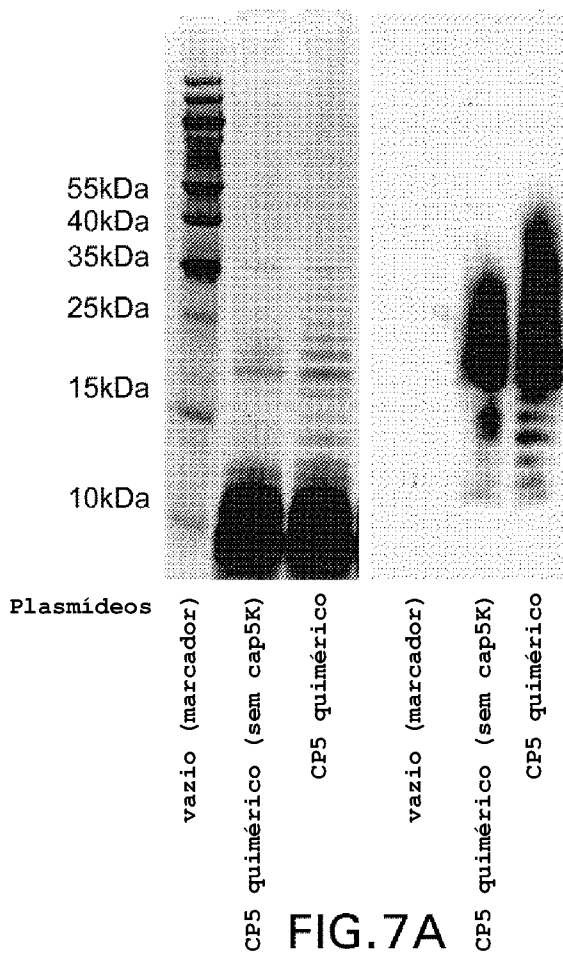


FIG. 6



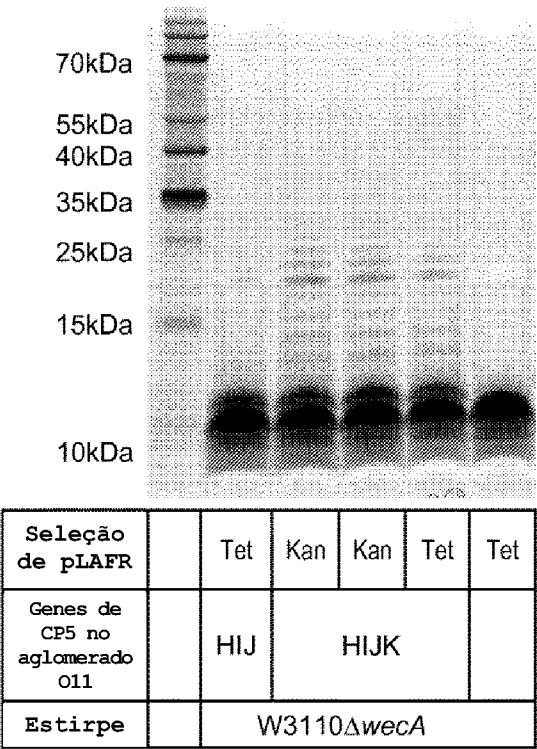


FIG.8A

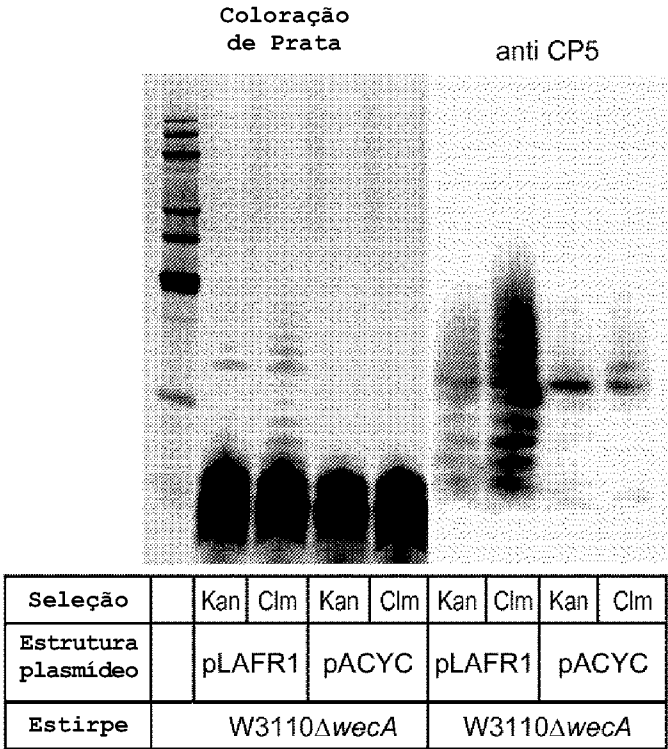


FIG.8B

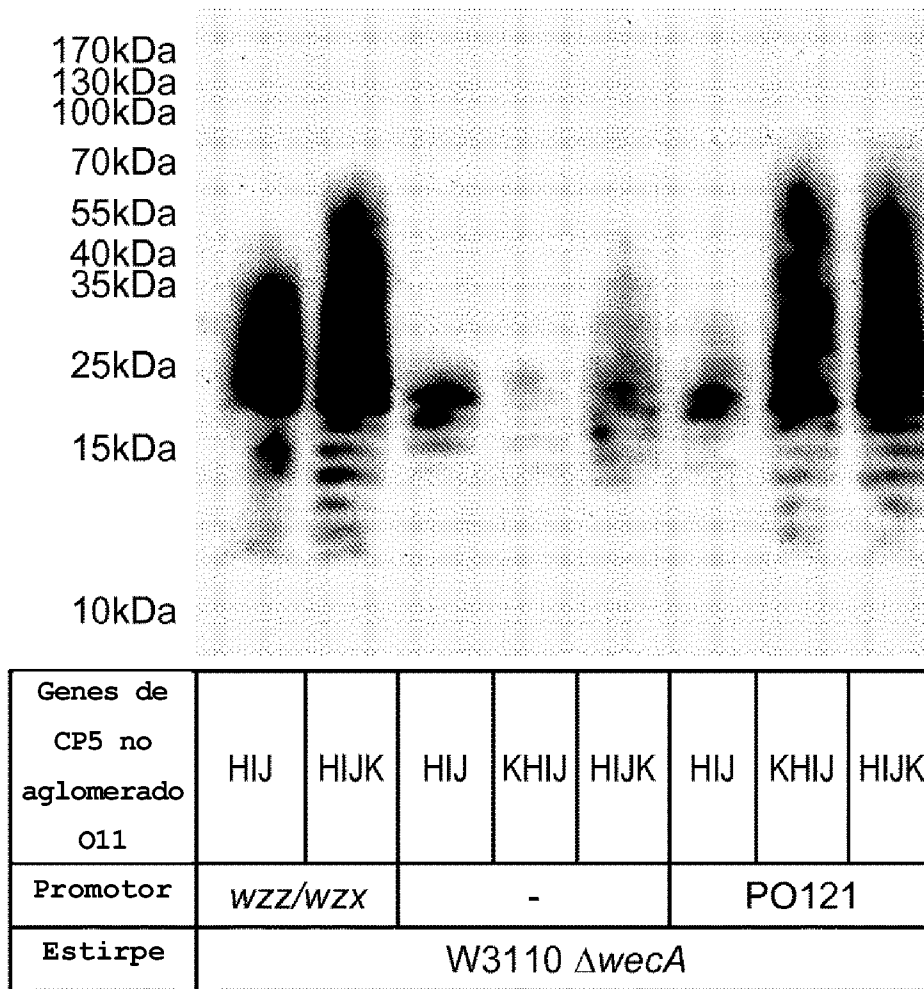


FIG.9

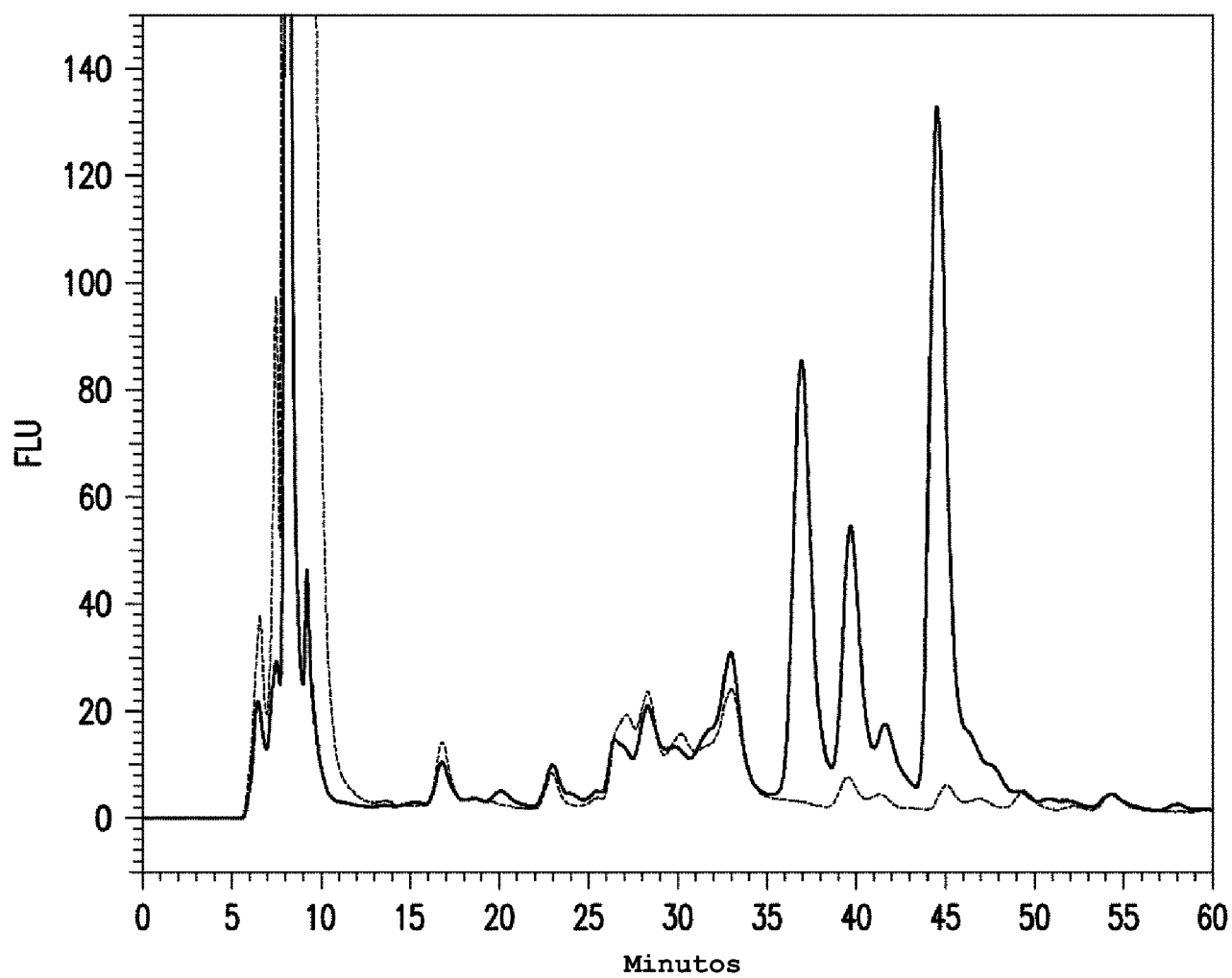


FIG.10A

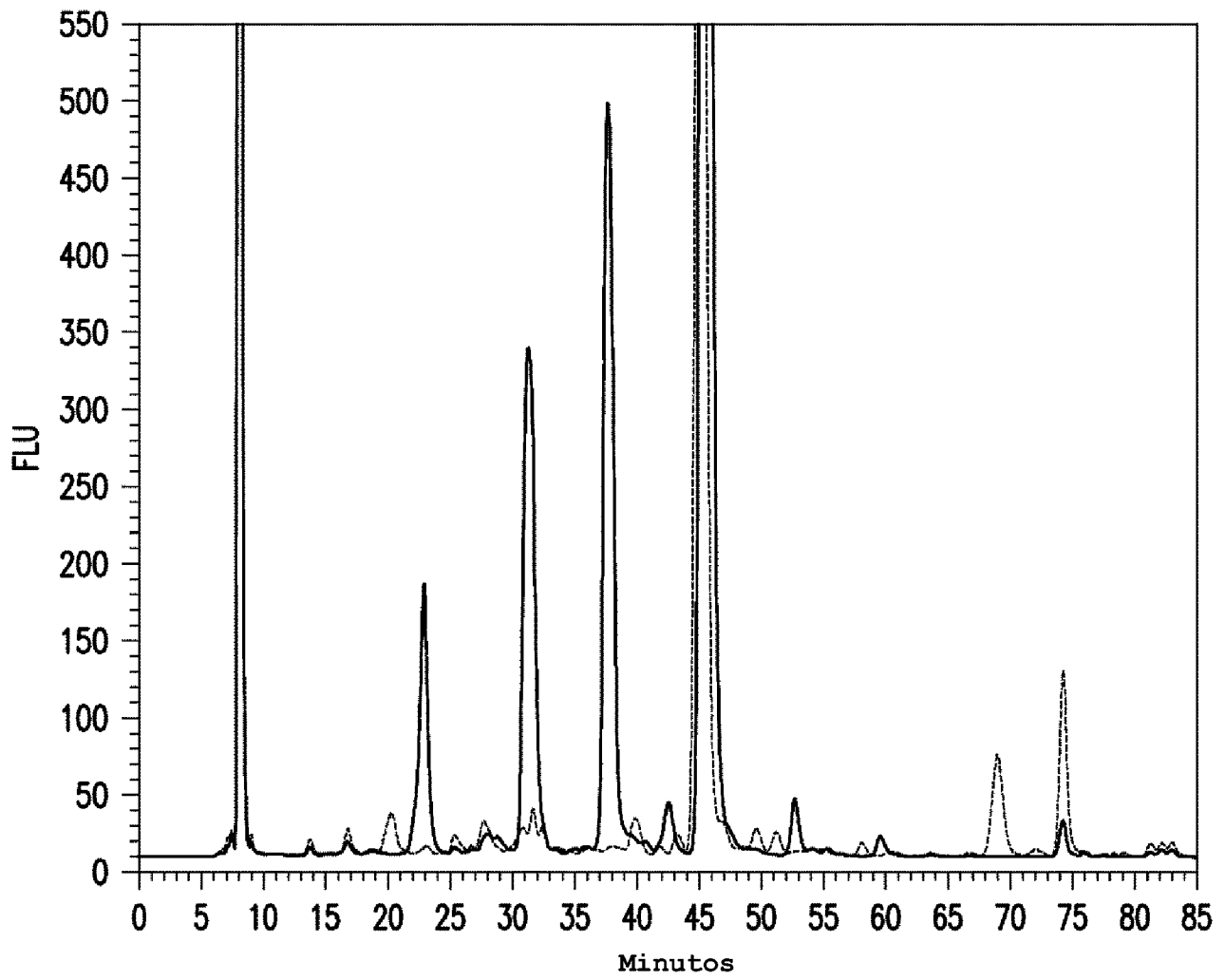


FIG.10B

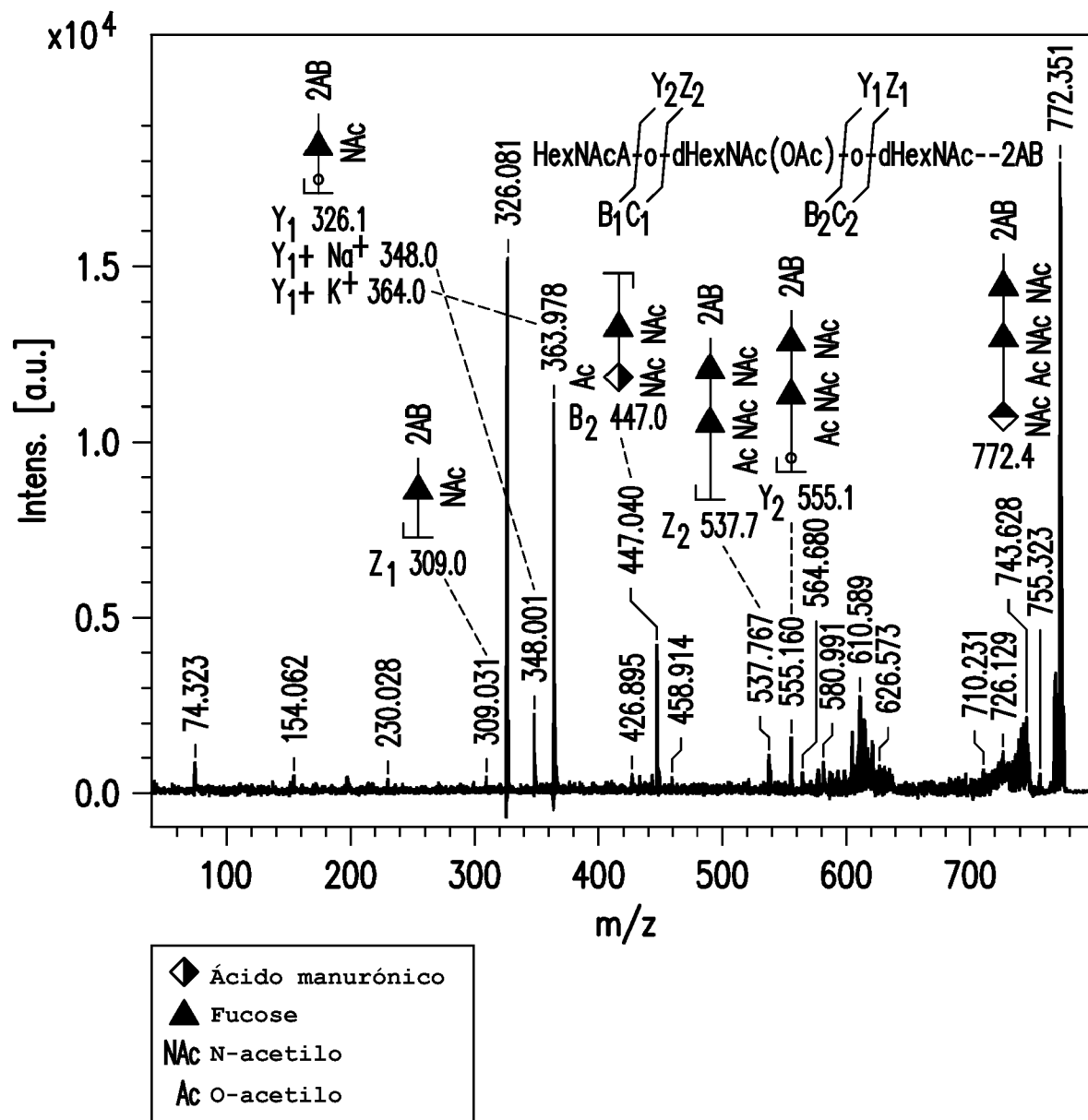


FIG.11A

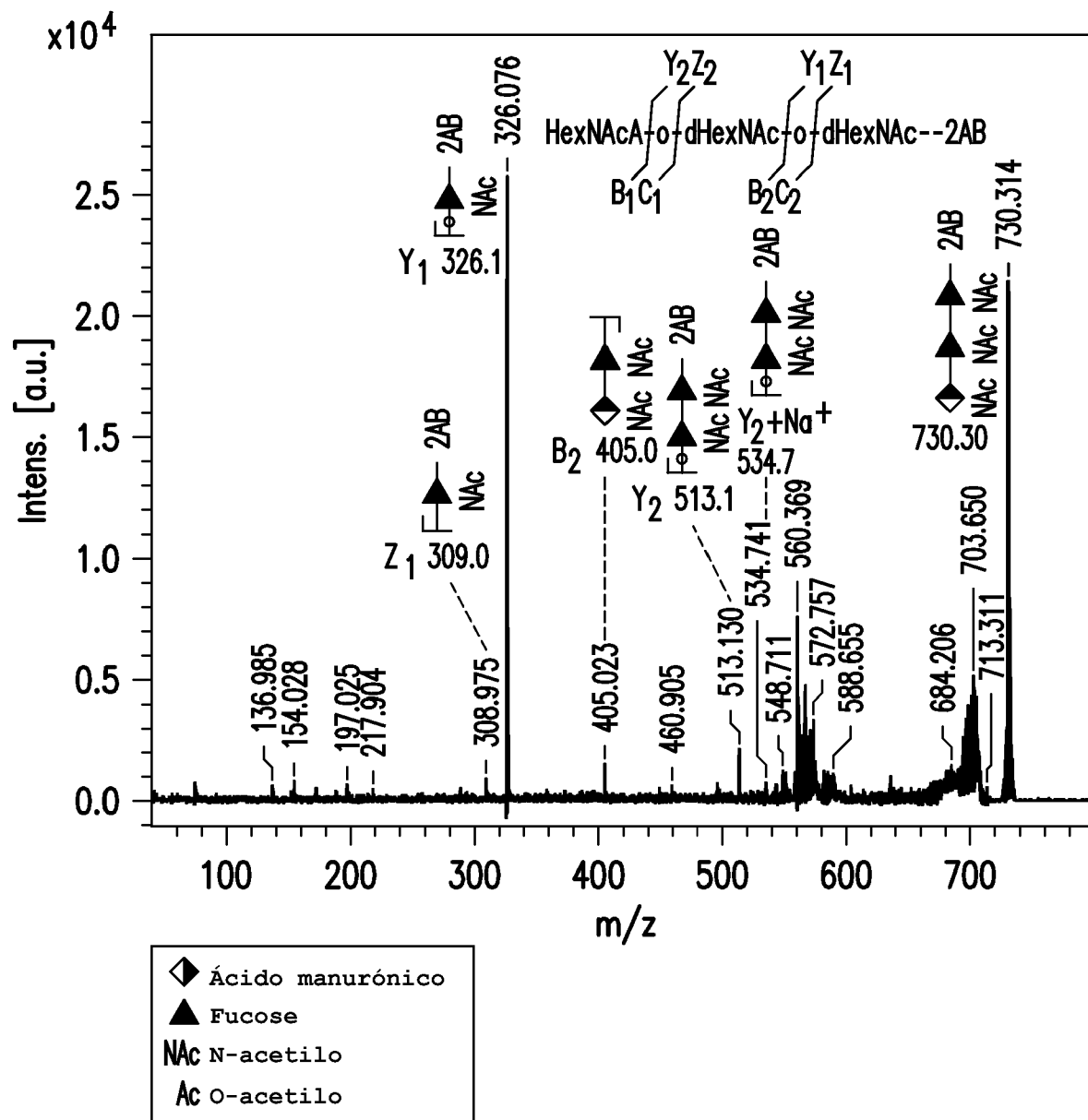


FIG.11B

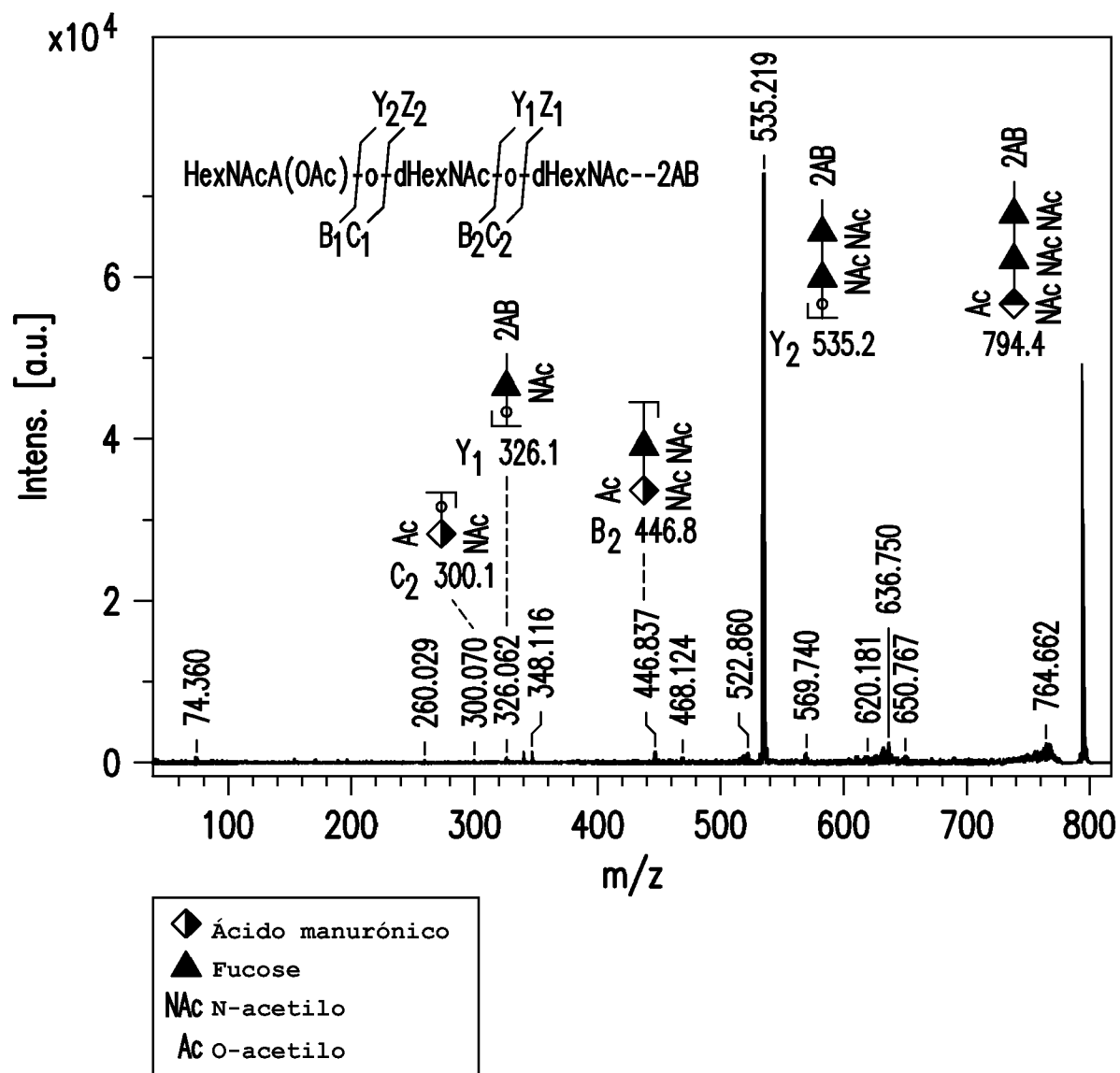


FIG.11C

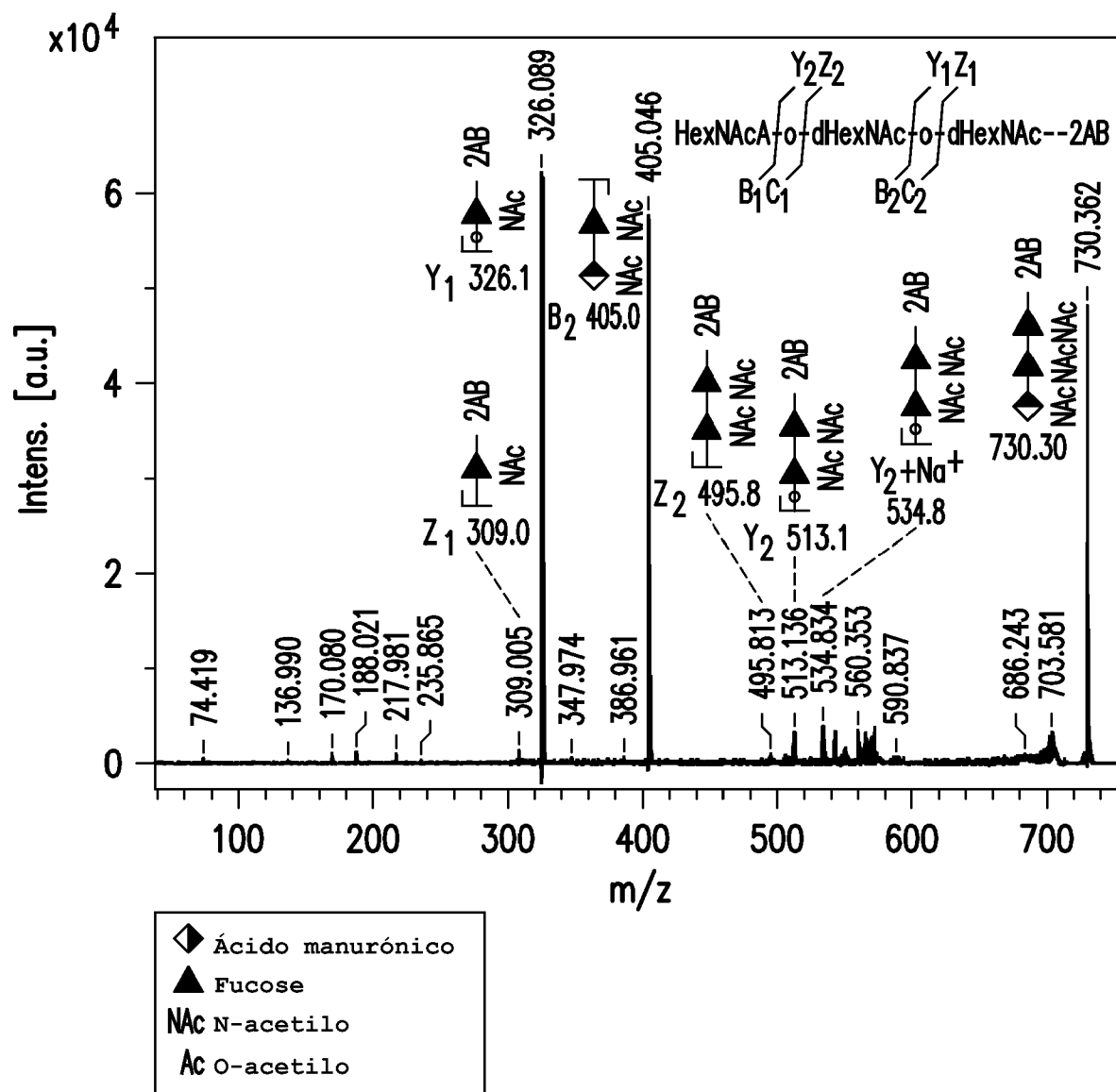


FIG.11D

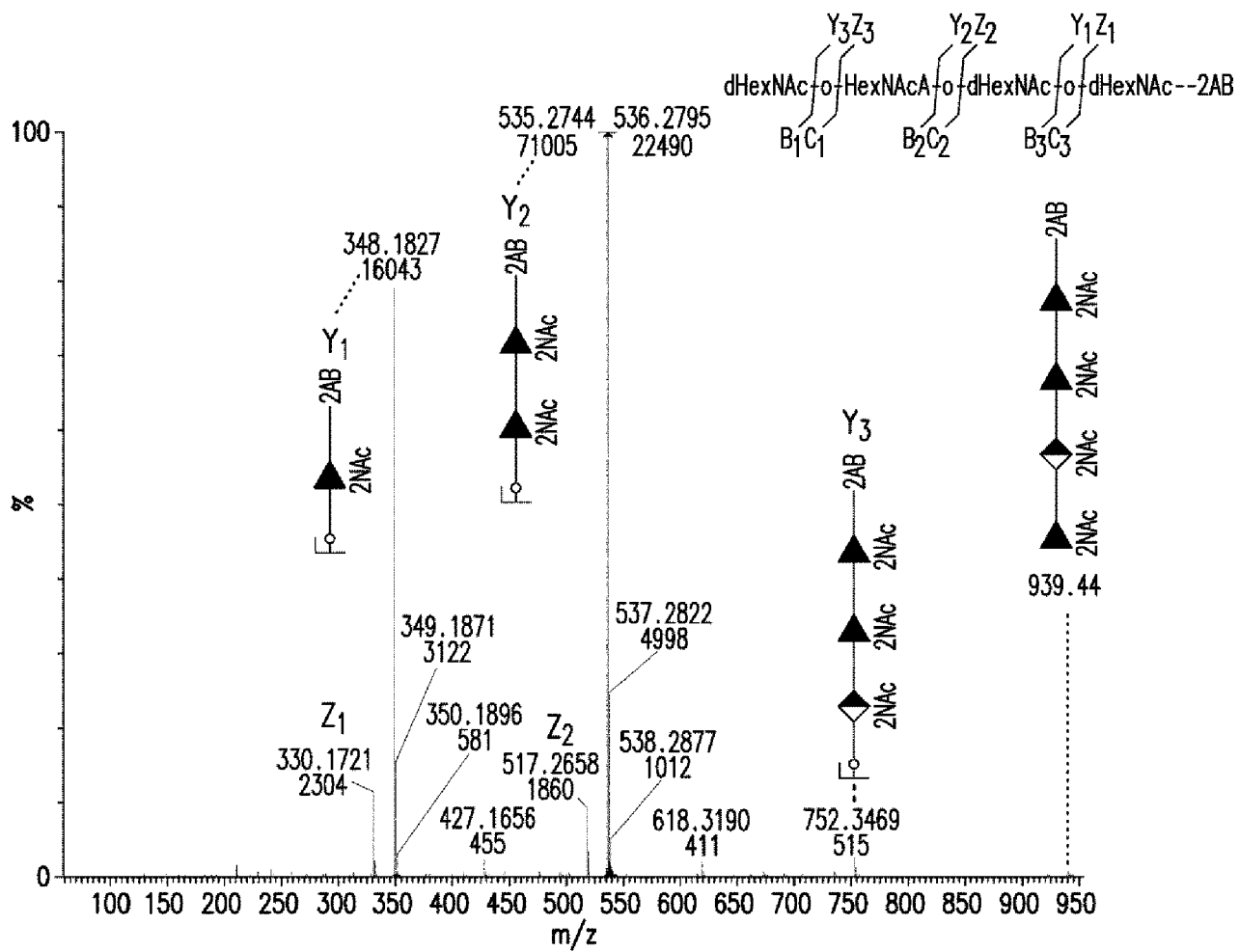


FIG. 11E

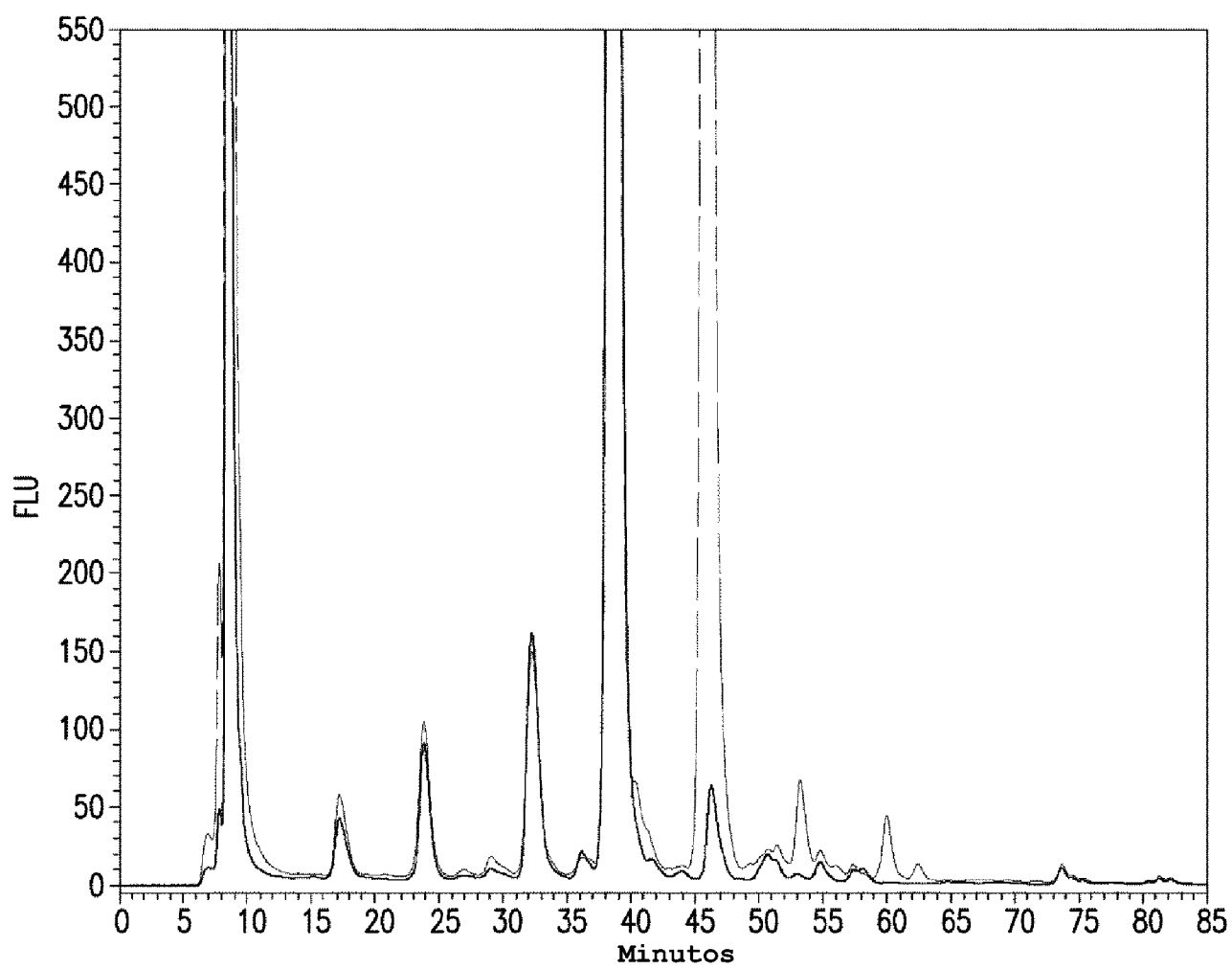


FIG.11F

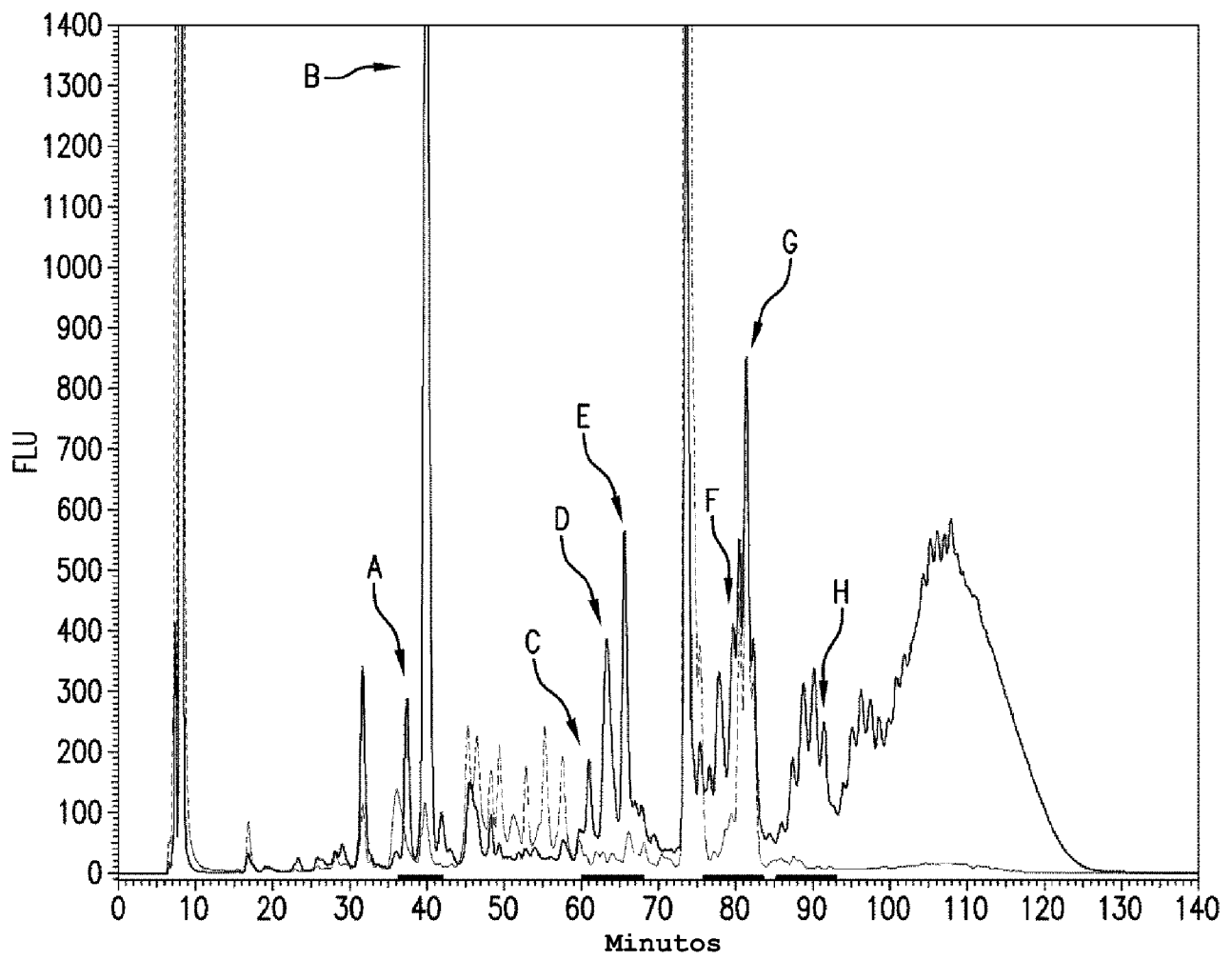


FIG.11G

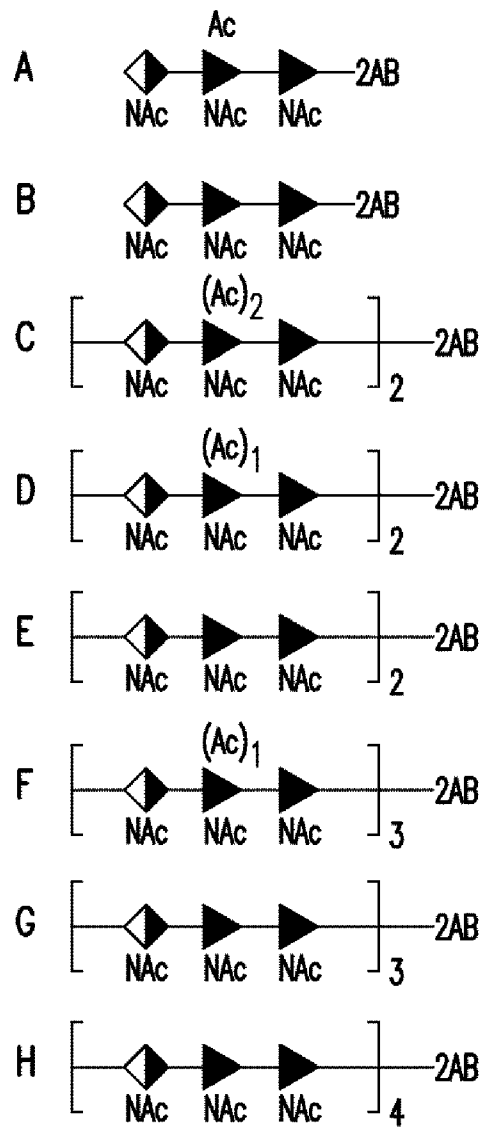


FIG. 11G-1

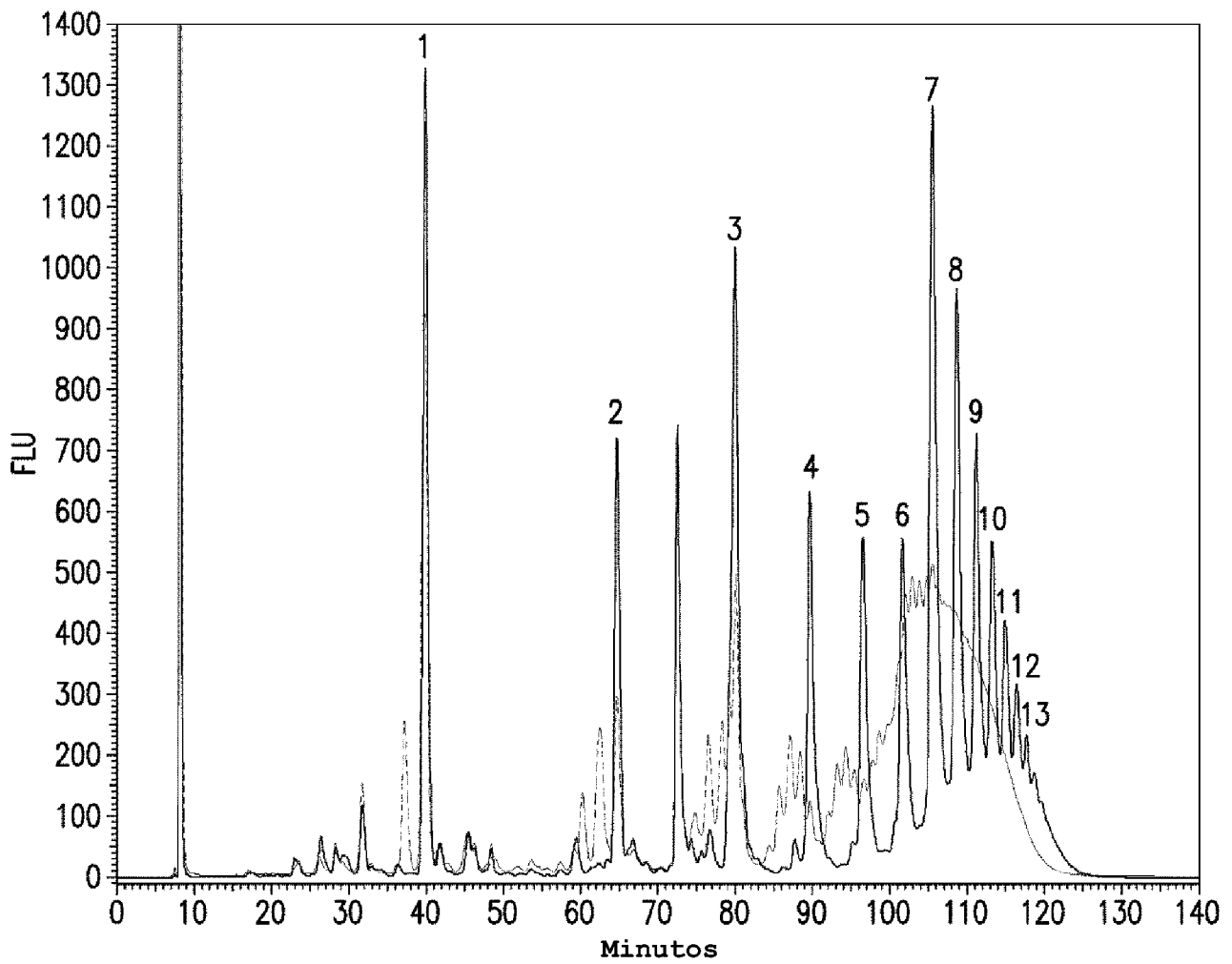


FIG.11H

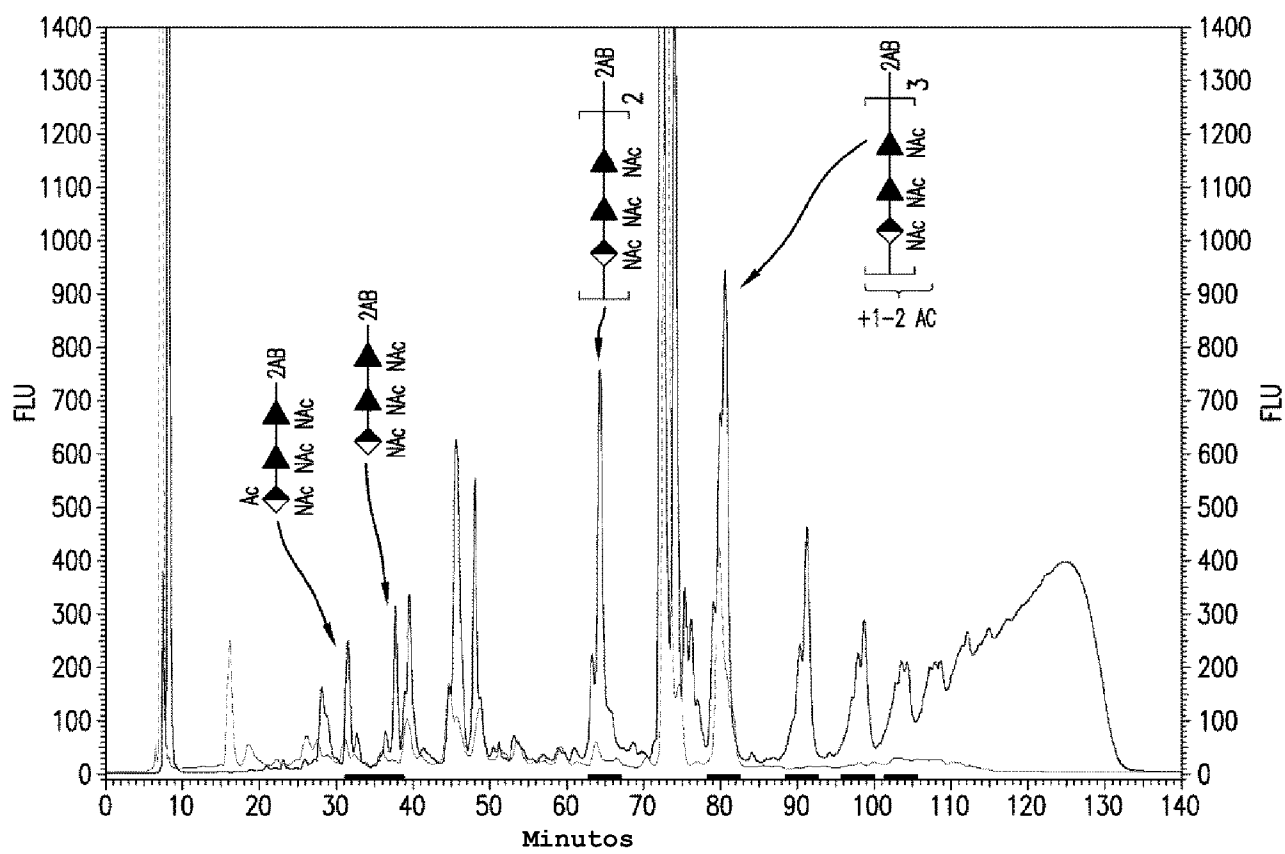


FIG. 11I

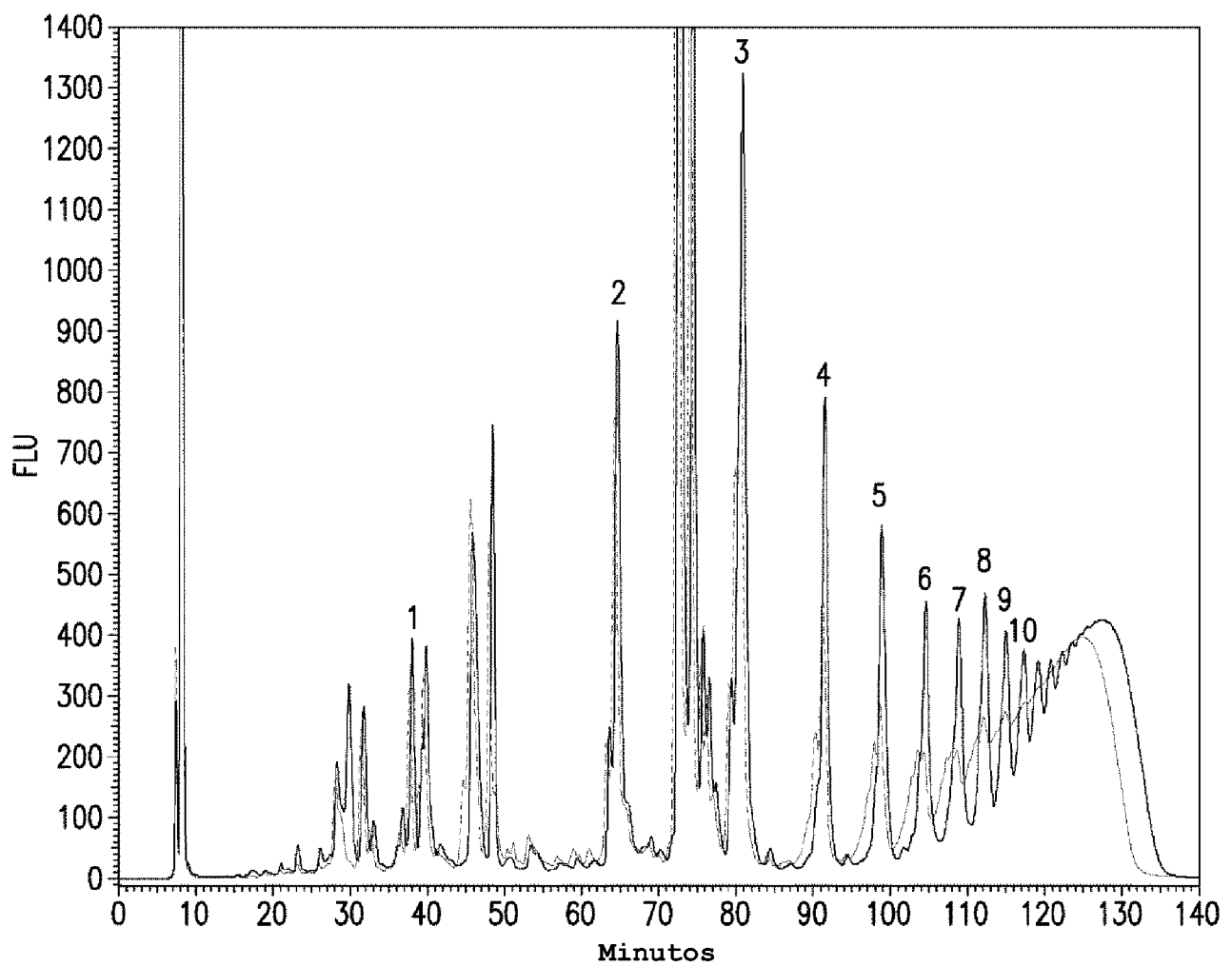


FIG.11J

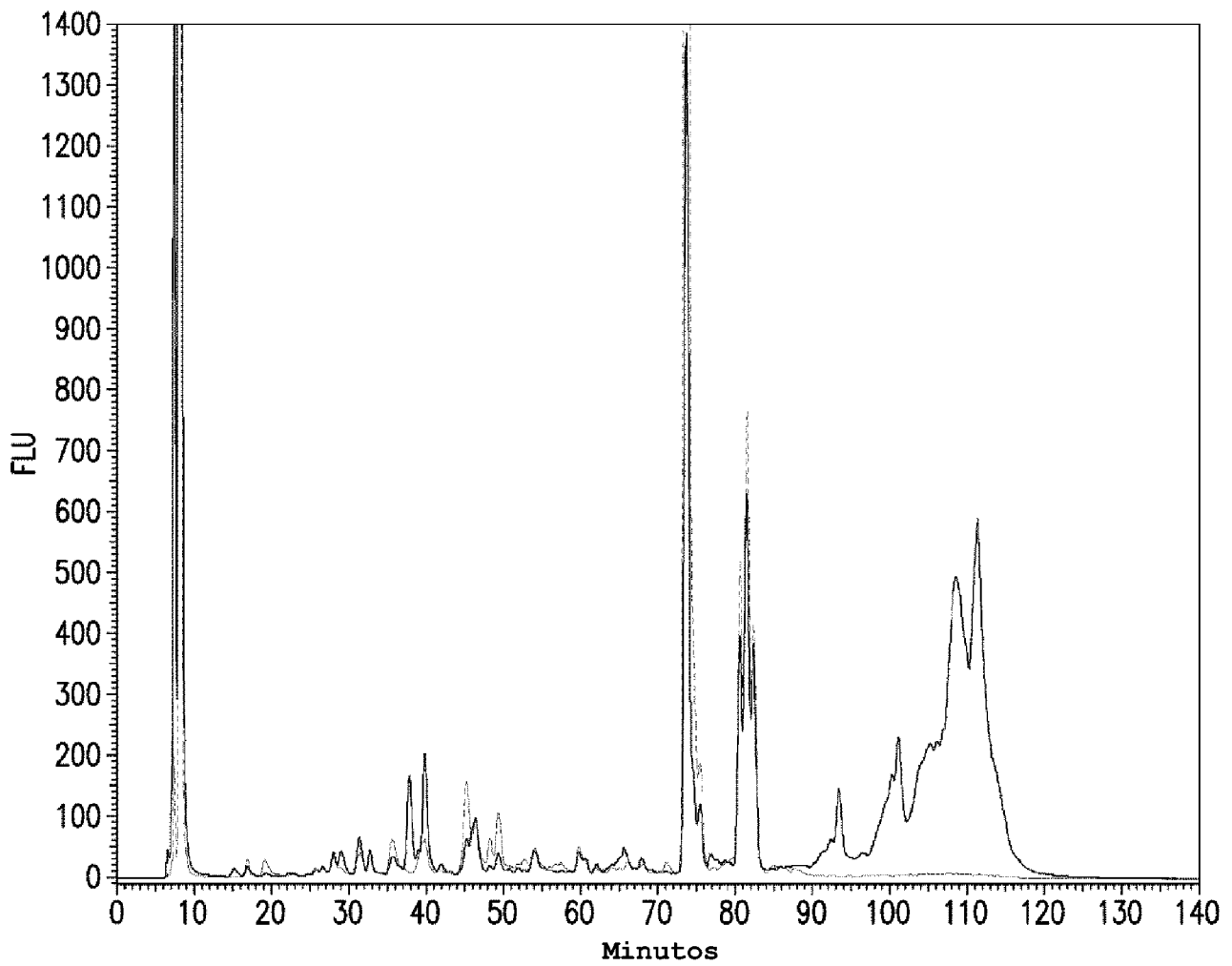


FIG.11K

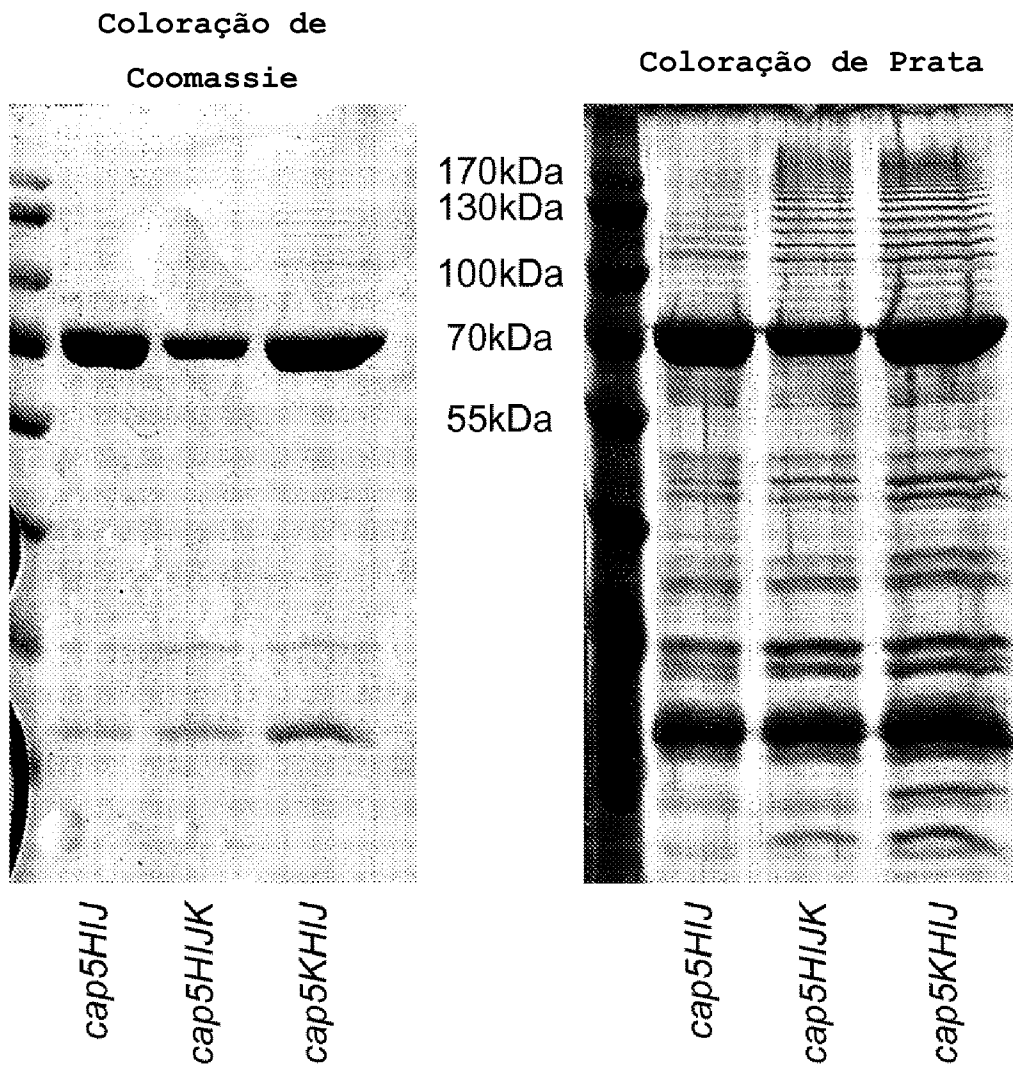


FIG.12

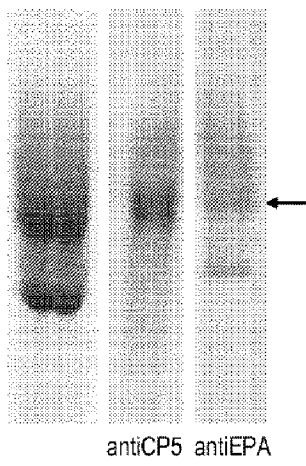


FIG.13A

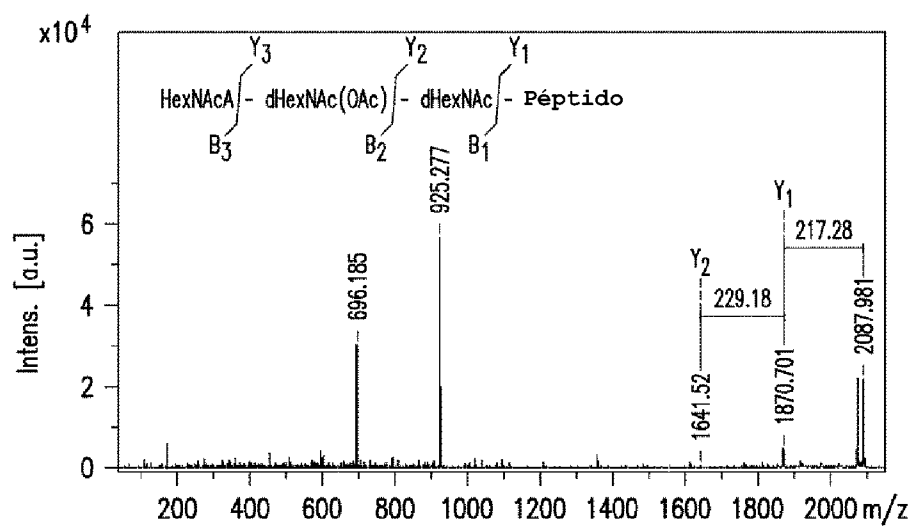


FIG.13B

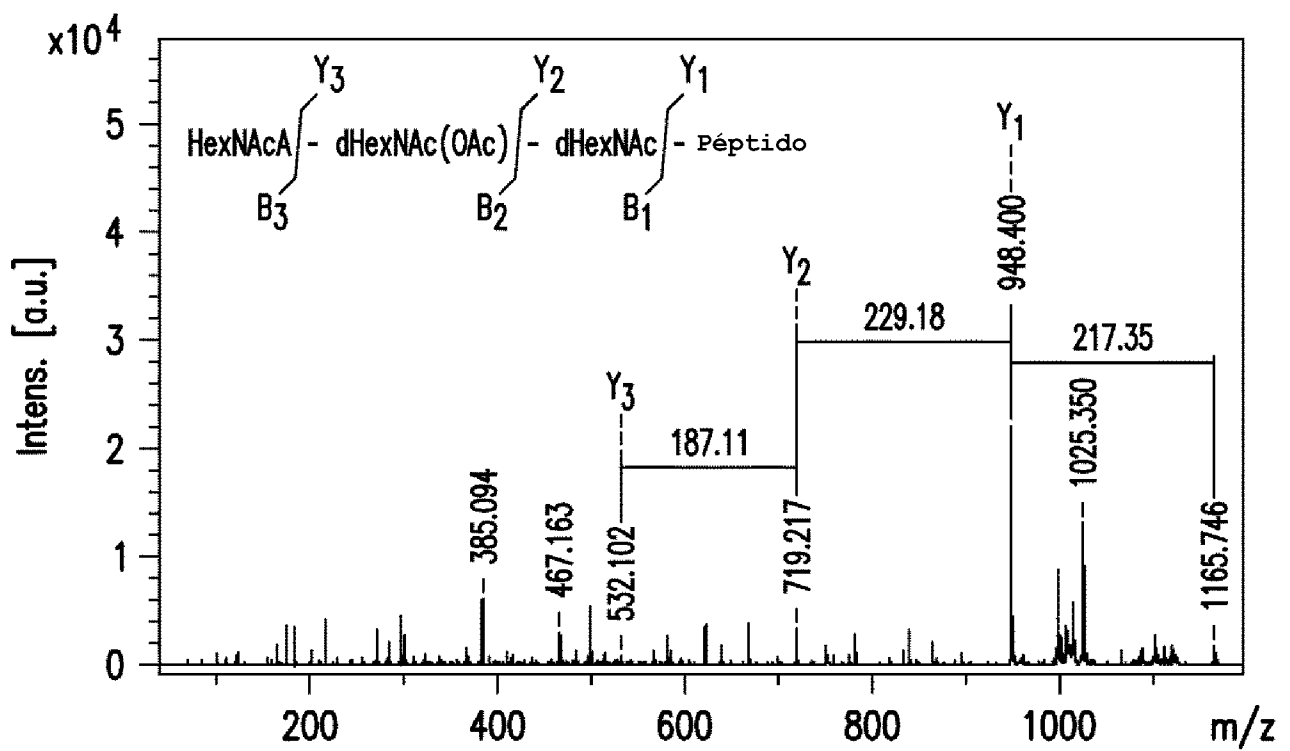


FIG.13C

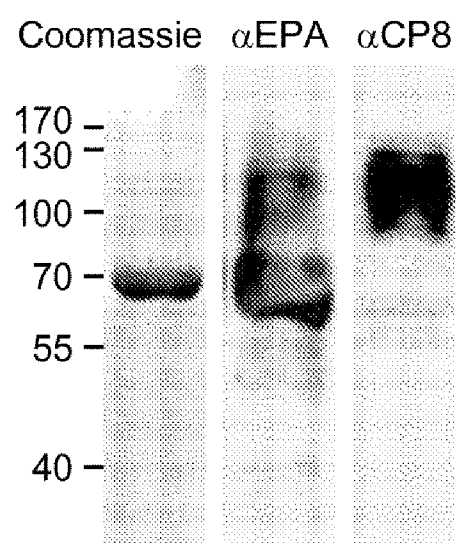


FIG.13D

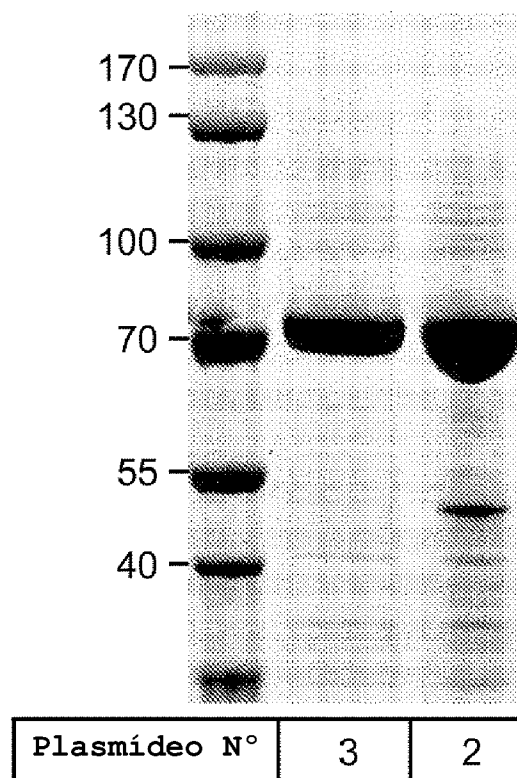


FIG.13E

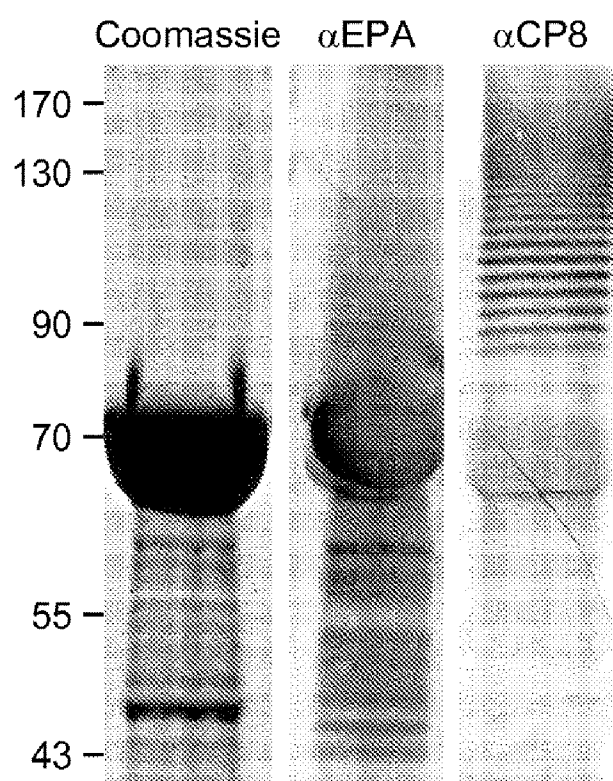


FIG.13F

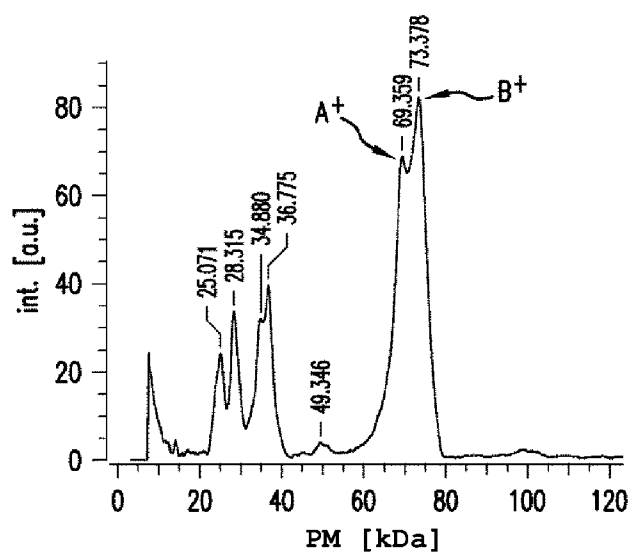


FIG.14A

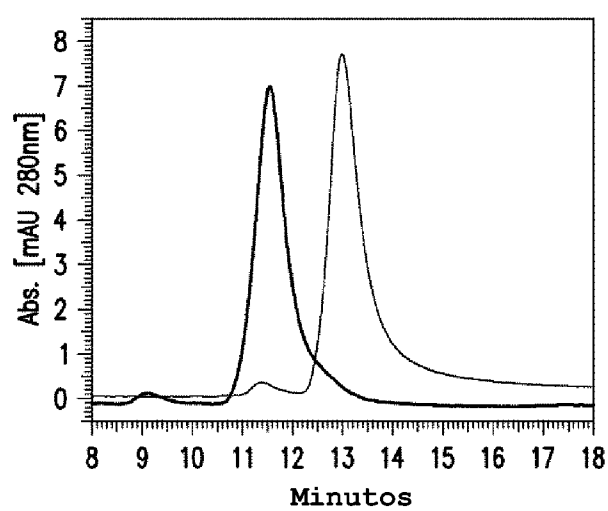


FIG.14B

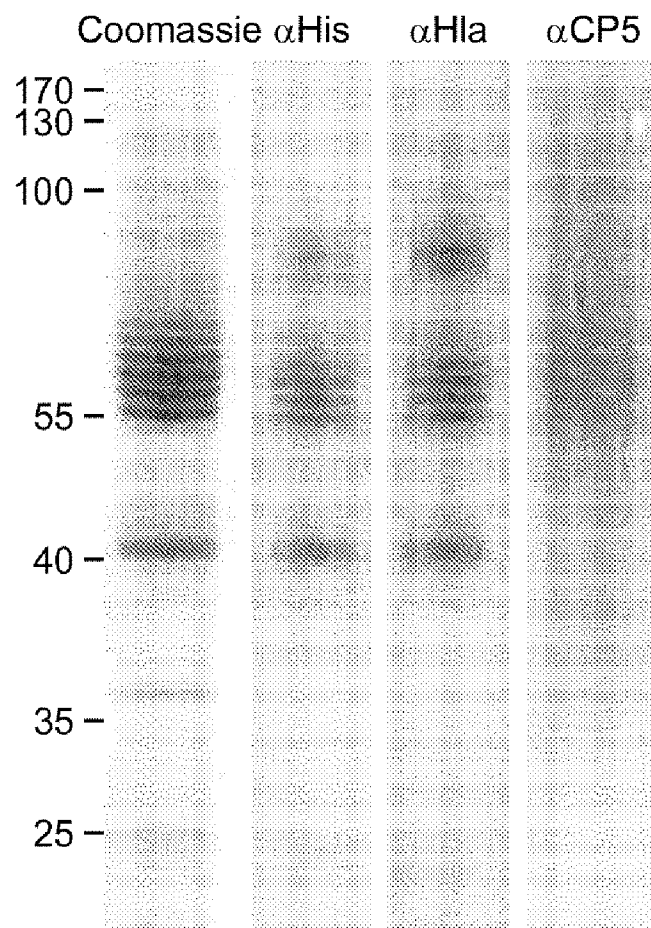


FIG.14C

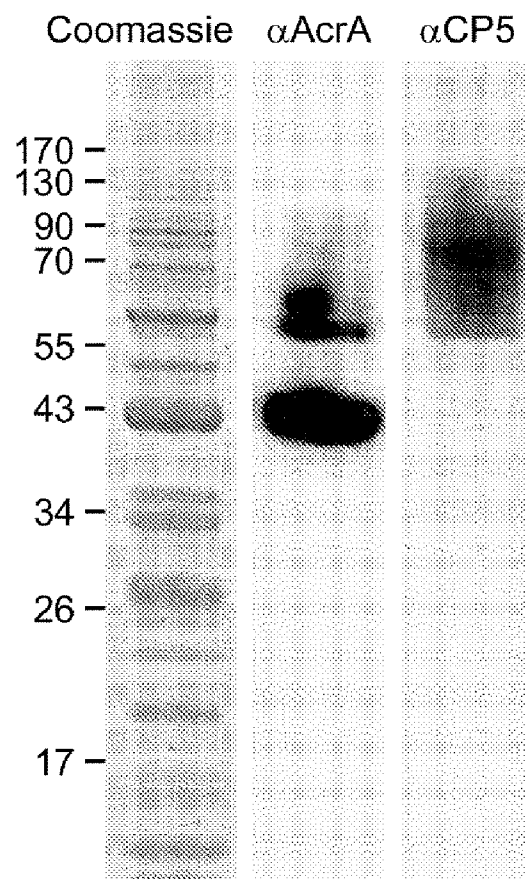


FIG. 14D

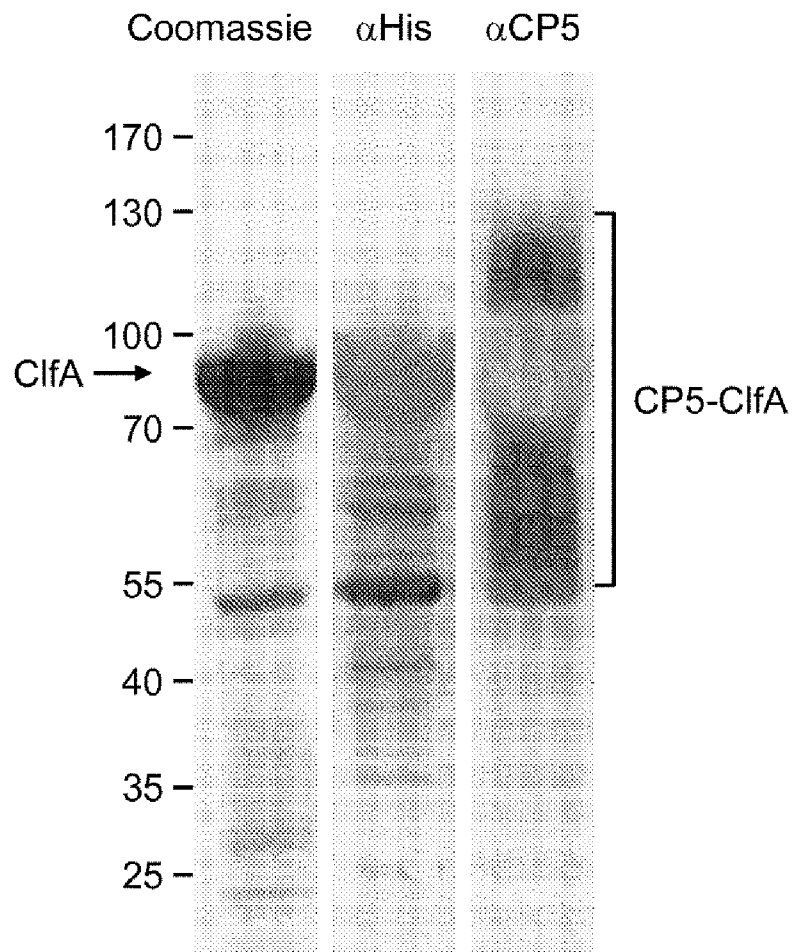


FIG.14E

ELISA em soro de murganhos imunizados
com CP5-EPA

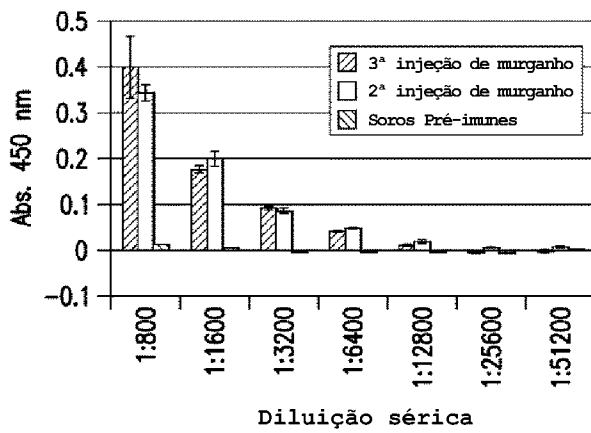


FIG.15A

ELISA em soro de coelhos imunizados

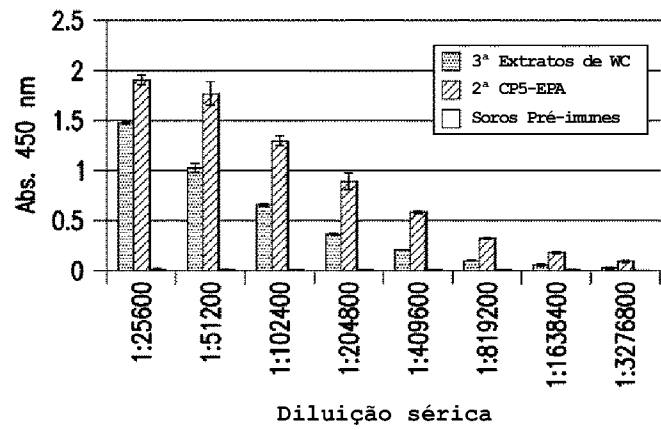


FIG.15B

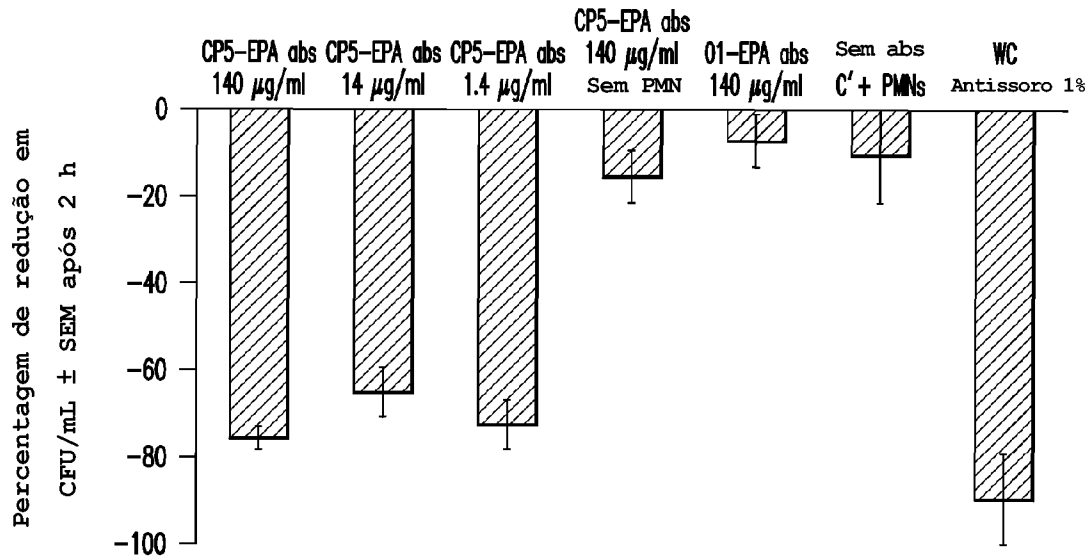
Neutralização Opsonofagocítica de *S. aureus* Reynolds

FIG.16A

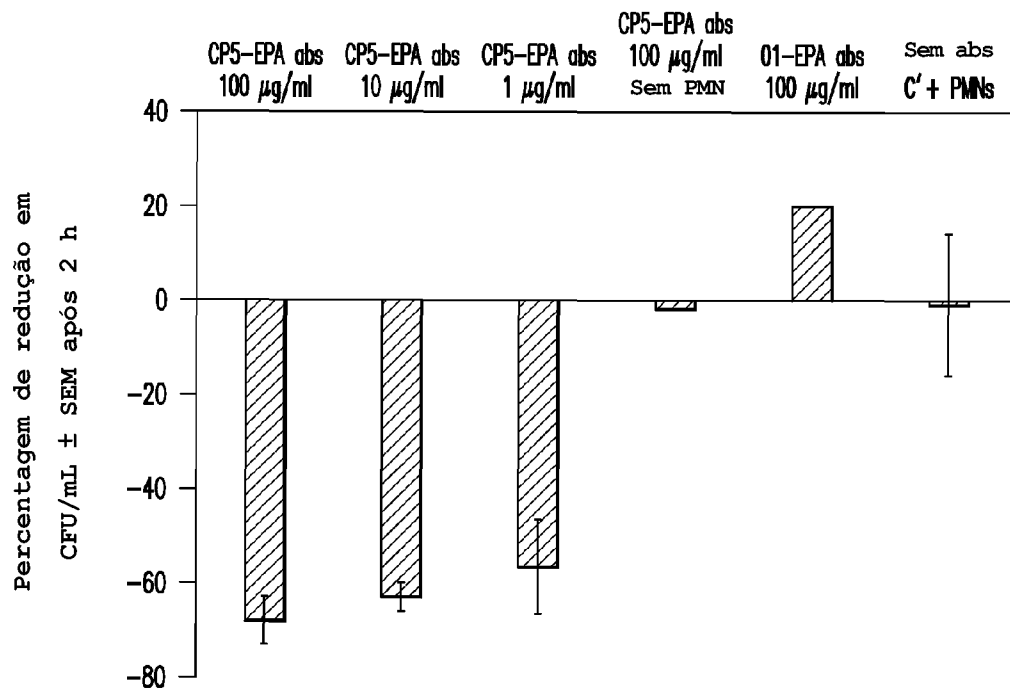
Neutralização Opsonofagocítica de *S. aureus* USA 100

FIG.16B

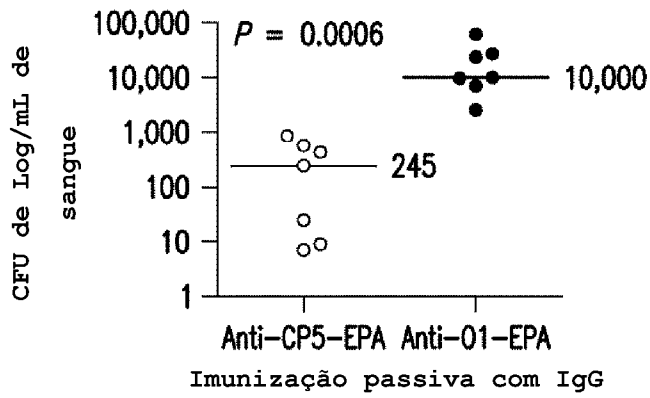


FIG.17A

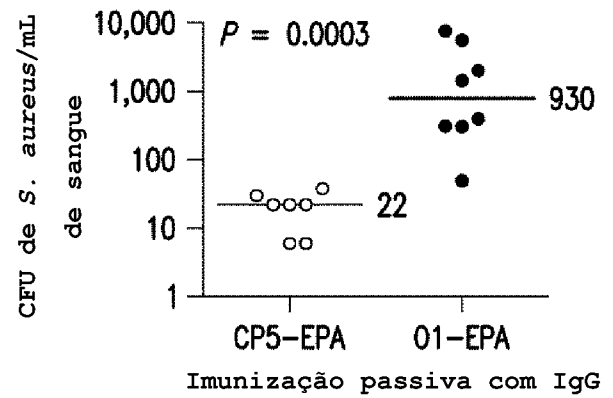


FIG.17B

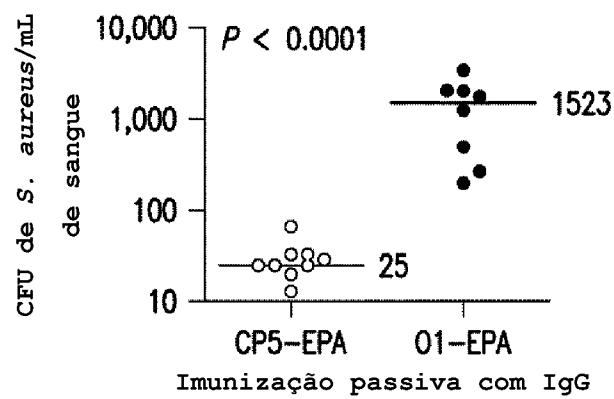


FIG.17C

Bacteremia de 2 h - imunização ativa contra a estirpe Reynolds (CP5)

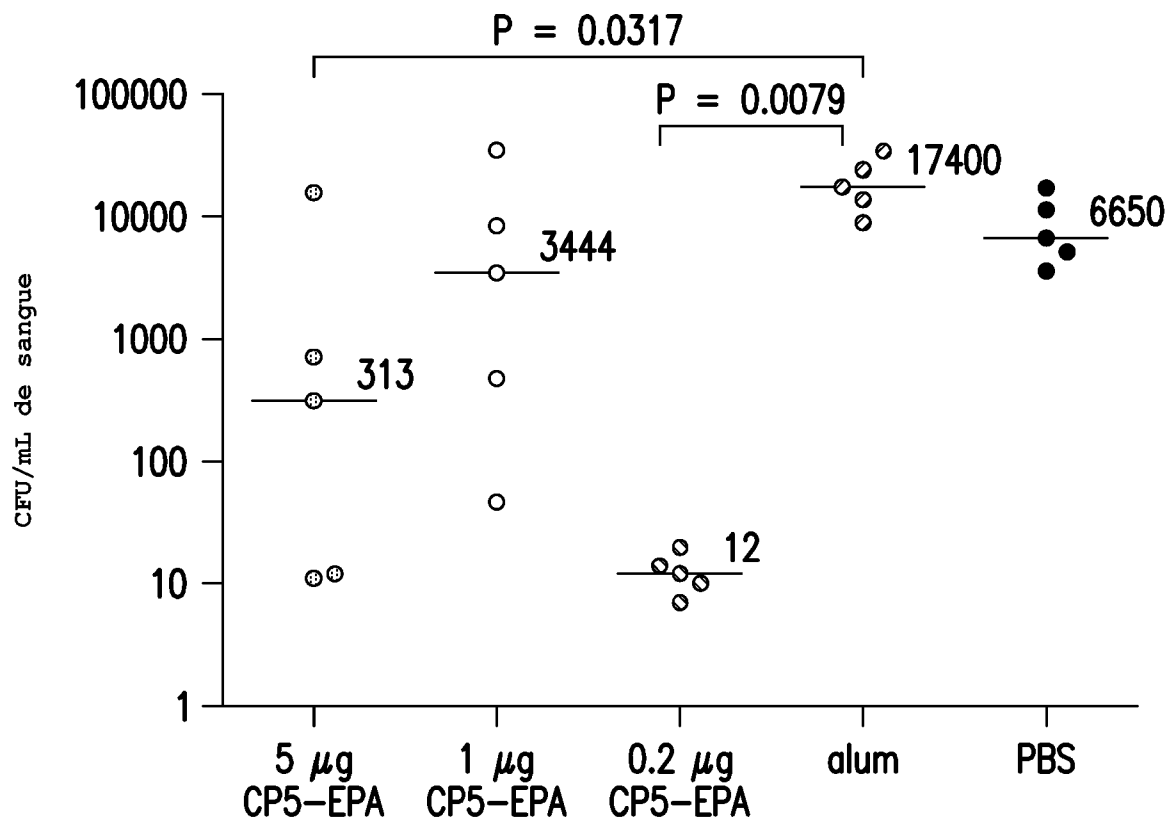


FIG.18