

MEMÓRIA DESCRITIVA
DA
PATENTE DE INVENÇÃO

Nº 93.721

NOME: CHIRON CORPORATION

EPIGRAFE: "METODO DE PREPARAÇÃO DE UM POLINUCLEOTIDO QUE
CODIFICA UM ANALOGO IMUNOGENICO DA PROTEINA DO
CIRCUNSPOROZOITO (CS) DO PLASMODIUM"

INVENTORES: Philip J. Barr, Ian C. Bathurst e Helen L.Gibson

Reivindicação do direito de prioridade ao abrigo do artigo
4º da Convenção da União de Paris de 20 de Março de 1883.
11 de Abril de 1989, nos Estados Unidos da América, sob
o Nº.336.288



TITULAR: CHIRON CORPORATION

EPIGRAFE: "METODO DE PREPARAÇÃO DE UM POLINUCLEOTIDO QUE
CODIFICA UM ANALOGO IMUNOGENICO DA PROTEINA DO
CIRCUNSPOROZOITO (CS) DE PLASMODIUM"


M E M O R I A D E S C R I T I V A

Campo do invento

O presente invento situa-se no campo da biologia molecular e da bioquímica. Particularmente, diz respeito aos análogos imunogénicos de proteínas do circunsporozoito de Plasmodium (CS), aos quais faltam uma ou mais das sequências repetitivas das proteínas nativas.

Antecedentes do invento

A malária é uma doença debilitante causada através da infecção por um parasita do género Plasmodium. É transmitida através das picadas de mosquitos. A medida que o mosquito se alimenta, os esporozoitos, que são a forma diferenciada do parasita que se encontra na glândula salivar,, entram na circulação sanguínea do hospedeiro. Numa última fase, os esporozoitos entram nas primeiras células




alvo, i.e., as células do fígado. A maioria das espécies de Plasmodium que infectam os humanos são o P.falciparum e P.vivax. Outras espécies, tais como o P.ovale e o P.malariae, infectam humanos com muito menos frequência. Há ainda outras espécies do parasita que infectam mamíferos, tal como o P.berghei que infecta roedores.

Foi feito um primeiro e substancial esforço na tentativa de fazer vacinas de subunidade contra a malária. Uma parte significativa deste trabalho foi realizada por investigadores na Universidade de Nova Iorque (UNI) e centrou-se na proteína CS que cobre a superfície do esporozoíto. A UNI apresentou uma série de pedidos de patentes em tecnologia da proteína CS.

A sua primeira patente, dos Estados Unidos com o nº de série 4.446.917, descreve uma classe de proteínas do esporozoíto, designada P-44 (em que P significa Plasmodium, e 44 é o seu peso molecular aproximado, em kilodaltons), que foram purificadas a partir dos lisados de esporozoítos.

WO 84/02922 descreve a identificação de uma região imunodominante da proteína CS, consistindo em repetições tandem de um epítipo de sequência relativamente curta, e a síntese de péptidos compostos de repetições desses epítipos. WO 86/00911 descreve a identificação e caracterização da sequência repetitiva do epítipo da proteína CS do P.falciparum, e a síntese de péptidos compostos de sequências tandem. De forma similar, WO 87/00533 descreve a identificação e caracterização da sequência repetitiva do epítipo da proteína CS do P.vivax. Embora esse pedido de patente indique que



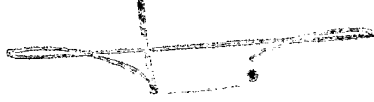
outros segmentos da proteína CS podem ser incluídos com as sequências tandem, o ensinamento claro da referência é que a imunogenicidade da proteína é atribuível à presença do epitopo repetido. Não é dado qualquer exemplo de péptido que inclua regiões delimitadoras.

WO 86/01721 ensina que certas regiões carregadas em cada lado da repetição tandem do epitopo da proteína CS, são imunogénicas. Os péptidos que correspondem em sequência a essas regiões foram sintetizados e mostraram ligar-se a anticorpos anti-CS. Conjugados desses péptidos sintéticos foram relatados como indutores de anticorpos anti-CS, em animais teste.

O pedido de patente norte-americano com o nº de série 032327, que é co-apresentada pela UNI e pelos assinantes do presente invento, descreve um análogo imunogénico da proteína CS de P.vivax, composto de toda a região repetitiva tandem da molécula nativa e de porções seleccionadas das sequências delimitadoras em cada lado da sequência repetitiva. Este polipéptido exibiu propriedades imunogénicas superiores aos péptidos que foram preparados pela requerente das patentes/pedidos norte-americanos acima descritos.

Eichinger, et al., Molecular and Cellular Biology (1968) 6:3965-3972, de la Cruz, et al., Molecular and Biochemical Parasitology (1988) 28:31-38, e Romero, et al., Vaccines 89 (1989) Cold Spring Harbor, pags 335-340, caracterizam a proteína CS do P.berghei e o gene que a codifica.

A parte de toda esta literatura pioneira sobre as



proteínas CS de Plasmodium , ainda não está bem definida qual a construção que fornecerá in vivo respostas em humanos que ofereçam protecção contra a malária. Neste aspecto, o papel da resposta imune a regiões excluindo repetições tandem, na neutralização do esporozoito, não foi extensivamente avaliada, dada a extrema imunodominância da região repetitiva.

Exposição do invento

O presente invento baseia-se no conceito de que uma redução da imunodominância da região repetitiva, relativa aos epitopos nas regiões que delimitam a região repetitiva, irá aumentar qualitativamente a actividade neutralizadora do esporozoito.

Por consequência, o invento diz respeito aos análogos imunogénicos da proteína CS de Plasmodium, onde faltam um ou mais dos epitopos repetitivos da proteína CS nativa, mas que têm pelo menos um epitopo delimitador não repetitivo de cada região delimitadora não repetitiva da proteína CS nativa, polinucleótidos recombinantes e hospedeiros transformados para fazer tais proteínas, e composições de vacinas que utilizam tais proteínas como imunogénicas.

Mais especificamente, um aspecto do invento é um análogo imunogénico da proteína CS de Plasmodium que compreenda:

(a) uma região ou domínio interno composto de pelo menos duas repetições de uma sequência repetitiva de uma proteína CS nativa de Plasmodium, mas menos repetições das



ditas sequências repetitivas do que as presentes na proteína CS nativa de Plasmodium;

(b) uma região amina terminal ou domínio delimitando a região interna, tendo uma sequência de aminoácidos que define, pelo menos, um epítipo da região amina delimitadora não repetitiva da proteína CS nativa de Plasmodium; e

(c) uma região carboxi-terminal, ou domínio, delimitando a região interna, tendo uma sequência de aminoácidos que define, pelo menos, um epítipo da região carboxilica delimitadora não repetitiva da proteína CS nativa de Plasmodium.


Outros aspectos do invento são polinucleótidos recombinantes que codificam os análogos da proteína CS de Plasmodium acima referidos, vectores de expressão para expressar tais polinucleótidos numa célula hospedeira compreendendo esse DNA ligado, de forma operativa, ao DNA que permita a expressão do DNA na célula hospedeira, e células hospedeiras que são transformadas com tais vectores e que são capazes de produzir tais análogos.

Ainda outros aspectos deste invento são as vacinas que contêm os análogos imunogénicos da proteína CS e métodos para protecção de indivíduos contra a malária, através da sua imunização com tais vacinas.

Breve Descrição das Figuras

Nas figuras:

A Figura 1 mostra a sequência de aminoácidos e a sequência de DNA que codifica um análogo da proteína CS de



P.vivax, o Vivax 3.

A Figura 2 mostra os oligómeros e o esquema de ligações que foram utilizados para construir a sequência de DNA que se mostra na Figura 1.

A Figura 3 é um diagrama esquemático do plasmideo pBS100.

A Figura 4 é um diagrama esquemático do plasmideo pBS24.1

A Figura 5 mostra a sequência de aminoácidos de um análogo da proteína CS de P.vivax, o Vivax 3.1, e uma sequência de DNA que codifica o mesmo análogo.

A Figura 6 mostra os oligómeros e o esquema de ligações que foram utilizados para construir a sequência de DNA que se mostra na Figura 5.


A Figura 7 é um diagrama esquemático do plasmideo pAB125.

A Figura 8 mostra a sequência de aminoácidos de um análogo da proteína CS de P.falciparum, o Falciparum 3, e uma sequência de DNA que codifica o mesmo análogo.

A Figura 9 mostra os oligómeros e o esquema de ligações que foram utilizados para construir a sequência de DNA que se mostra na Figura 8.

A Figura 10 mostra a sequência de aminoácidos de um análogo da proteína CS de P.falciparum, o Falciparum 4, e a sequência de DNA que codifica o mesmo análogo.

A Figura 11 mostra a sequência de aminoácidos de um análogo da proteína CS de P.berghei, o Berghei 3, e uma sequência de DNA que codifica o mesmo análogo.




A Figura 12 mostra os oligômeros e o esquema de ligações que foram utilizados para construir a sequência de DNA que se mostra na Figura 11.

Modos de realização do invento

Definições

O termo "sequência repetitiva de uma proteína CS nativa de Plasmodium", diz respeito à sequência repetitiva de aminoácidos que aparece nas regiões imunodominantes ou domínios centrais da proteína CS nativa de Plasmodium, e as modificações dessa sequência repetitiva que não destroem as suas propriedades imunogênicas. As sequências repetitivas nativas para P.vivax e P.falciparum estão descritas nos pedidos das patentes norte-americanas citados acima. As sequências repetitivas nativas para P.berghei estão descritas por Eichinger, et al., supra citado. Por exemplo, as sequências repetitivas do P.vivax são Asp-Arg-Ala-Asp/Ala-Gly-Glu-Pro-Ala-Gly, as do P.falciparum são Asn-Ala-Asn-Pro e Asn-Val-Asp-Pro, e as do P.berghei são Pro-Pro/Ala-Pro-Pro-Asn-Pro/Ala-Asn-Asp.

O termo "região amina delimitadora, não repetitiva, de uma proteína CS nativa de Plasmodium", diz respeito à porção dessa proteína que se estende do grupo amina terminal da proteína à região da sequência repetitiva da proteína. De forma semelhante, o termo "região carboxilica delimitadora, não repetitiva, de uma proteína CS nativa de Plasmodium", diz respeito à porção dessa proteína que se estende do grupo carboxi terminal da proteína à região de sequência repetitiva.



O termo "interno", como usado na descrição de uma sequência de aminoácidos de uma proteína, diz respeito a uma sequência que não inclui um aminoácido terminal da proteína.

"Células hospedeiras", "células", "linhas celulares", "culturas de células", e outros termos semelhantes que indiquem microorganismos procariotas ou linhas de células eucariotas cultivadas como entidades unicelulares, são usados alternadamente, e referem-se a células que foram ou podem ser usadas como recipientes para vectores recombinantes ou outro DNA transferido, e incluem a descendência da célula original que foi transfectada. Está entendido que a descendência de uma só célula parental pode não ser, necessariamente, completamente idêntica, em morfologia ou em genoma, ou em todo o complemento de DNA, ao progenitor original, devido a mutações acidentais ou deliberadas. A descendência da célula parental, que é suficientemente semelhante às células parentais, para ser classificada através de propriedades relevantes, como a presença de uma sequência de nucleótidos codificando um péptido desejado, está incluída no que se entende por descendência nesta definição, e está incluída no âmbito dos termos acima descritos.

Uma "réplica" é um elemento genético, e.g., um plasmídeo, um cromossoma, um vírus, que se comporta como uma unidade autónoma de replicação polinucleotídica na célula.

Um "vector" é uma réplica à qual está ligada um outro segmento polinucleotídico, para efectuar a replicação e/ou expressão do segmento ligado.

"Sequência controlo" refere-se às sequências




polinucleotídicas que são necessárias para efectivar a expressão das sequências codificantes às quais estão ligadas. A natureza de tais sequências controlo difere, dependendo do organismo hospedeiro; em procariotas, tais sequências controlo incluem, geralmente, um promotor, um local de ligação do ribossoma, e terminadores; em eucariotas, geralmente, tais sequências controlo incluem promotores, terminadores e, por vezes activadores. O termo "sequências controlo" é entendido como incluindo, no mínimo, todos os componentes cuja presença é necessária para a expressão, e pode também incluir componentes adicionais cuja presença é vantajosa, como por exemplo, as sequências " leader".

"Ligado de forma funcional" refere-se a uma justaposição onde os componentes descritos estão relacionados de forma a permitir que funcionem na forma pretendida. Uma sequência controlo "ligada de forma funcional" a uma sequência código, está ligada de tal modo que a expressão da sequência código é feita em condições compatíveis com as sequências controlo.

O termo "imunogénico", usado para descrever os análogos da invenção, diz respeito à capacidade de causar a neutralização do esporozoíto de Plasmodium pela indução de uma resposta humoral e /ou celular. O termo "neutralização" diz respeito a um estado incapaz de causar infecção e/ou sintomas de malária.

O termo "transformação", como é aqui usado, refere-se à inserção de um polinucleótido exógeno na célula hospedeira, à parte do método utilizado para a inserção, por exemplo,



inserção directa, transdução, ou mediado pelo factor f. O polinucleotido exógeno pode manter-se como um vector não integrado, por exemplo, um plasmideo, ou, alternativamente, pode estar integrado no genoma do hospedeiro.

O termo "tratamento", como é aqui usado, refere-se a profilaxia e/ou terapia.

O termo "indivíduo", como é aqui usado refere-se a um animal que é susceptível à infecção por uma ou mais espécies de Plasmodium, e inclui primatas e humanos, não estando no entanto limitado a estes.

Caracterização dos análogos da proteína CS

Como indicado acima, aquilo que distingue o análogo deste invento dos análogos da proteína CS precedentes, é que aos primeiros faltam uma ou mais sequências repetitivas das que se encontram na proteína nativa, mas retêm os epitopos sequencias de ambas as regiões delimitadoras da proteína nativa.


A proteína CS nativa de P.vivax consiste em 373 aminoácidos. Tem uma sequência "leader" N-terminal hidrofóbica, que se estende dos aminoácidos 1 a 22, (todos os aminoácidos referidos são inclusivé, a numeração dos aminoácidos na proteína CS nativa de P.vivax está de acordo com aquela que se mostra na sequência de Vivax 3.1, na figura 5). As repetições tandem, das quais há 19, começam no aminoácido 92 e estendem-se até ao aminoácido 263, e estão precedidas de uma região conservativa, chamada Região I, que se estende do aminoácido 77 até ao 91. A região C-terminal



acaba com uma sequência transmembranar, que se estende do aminoácido 356 até ao 373 e também contém uma região conservativa, chamada Região II, do aminoácido 306 ao 318. Nove das sequências repetitivas são Asp-Arg-Ala-Asp-Gly-Gln-Pro-Ala-Gly (designada sequência "A") e dez são Asp-Arg-Ala-Ala-Gly-Gln-Pro-Ala-Gly (designada sequência "B").

Correspondentemente, a proteína CS de P.falciparum consiste em 412 aminoácidos. A sua sequência "leader" N-terminal hidrofóbica, estende-se do aminoácido 1 até ao aminoácido 42 e a sua Região I está localizada nos aminoácidos 107-123, inclusivé. (A numeração dos aminoácidos na proteína CS de P.falciparum está de acordo com a que se mostra na sequência Falc3 mostrada na figura 8). Há 41 repetições na proteína, com quatro sendo Asn-Val-Asp-Pro (sequência "A") e as restantes sendo Asn-Ala-Asn-Pro (sequência "B"). A Região II da proteína CS de P.falciparum está localizada nos aminoácidos 344-356, e a sua sequência transmembranar está nos aminoácidos 391-412.


A proteína CS nativa de P.berghei, como descrita por Eichinger et al., supra, (como corrigida por de la Cruz et al., supra), consiste em 339 aminoácidos. As repetições tandem, das quais há 13, começam no aminoácido 93 e acabam no aminoácido 196. Cinco das sequências repetitivas são Pro-Pro-Pro-Pro-Asn-Pro-Asn-Asp (designada por sequência "A"), duas são Pro-Pro-Pro-Pro-Asn-Ala-Asn-Asp, e seis são Pro-Ala-Pro-Pro-Asn-Ala-Asn-Asp (designada por sequência "B"). A Região I estende-se, aproximadamente, do aminoácido 82 até ao aminoácido 92. A Região II estende-se do aminoácido 272 até ao



aminoácido 286. A sequência "leader" N-terminal hidrofóbica compreende os aminoácidos do 1 ao 23.

Os análogos deste invento contêm, pelo menos, duas sequências repetitivas da proteína CS nativa, tendo de preferência presentes ambos os tipos de sequências repetitivas. Na possibilidade mais extrema, as razões entre a sequência A e a sequência B são preservadas no análogo e a sua distribuição (i.e., o posicionamento respectivo dos dois tipos de repetições) é também preservada. Neste aspecto, as sequências A de P.vivax são predominantes na terminação N-terminal das séries de repetições tandem. O mesmo se aplica nos casos das proteínas de P.falciparum e de P.berghei. Conseqüentemente, nos análogos de P.vivax da invenção há 2-18 repetições, de preferência 2 a 10 repetições, com um número aproximadamente igual de repetições da sequência A e da sequência B, com a(s) sequência(s) A localizada(s) predominantemente na porção N-terminal das séries. Correspondentemente, os análogos de P.falciparum conterão 2 a 40 repetições, de preferência 2 a 20 repetições, com uma razão A/B de 1:10 a 3:10. Os análogos de P.berghei conterão 2 a 12 repetições.

A região delimitadora N-terminal compreenderá, tipicamente, a Região I e, opcionalmente, sequências delimitadoras adicionais para fora da sequência "leader" hidrofóbica. Deste modo, a região delimitadora N-terminal do análogo de P.vivax conterá tipicamente, pelo menos, os aminoácidos 77-91 da proteína CS nativa de P.vivax e, opcionalmente, todos ou uma porção dos aminoácidos 23-76. De



preferência, consiste nos aminoácidos 23-91. Correspondentemente, a região delimitadora N-terminal dos análogos de P.falciparum conterá, pelo menos, os aminoácidos 107-123 da proteína CS nativa de P.falciparum e, opcionalmente, todos ou uma porção dos aminoácidos 43-106. De preferência, consiste nos aminoácidos 43-123. No caso dos análogos de P.berghei, a região delimitadora N-terminal conterá tipicamente, pelo menos, os aminoácidos 82 a 92.

A região delimitadora C-terminal compreenderá, tipicamente, a sequência nativa que se estende do fim da região repetitiva até à Região II e, opcionalmente, toda ou uma parte da sequência que se estende do fim da Região II até à sequência transmembranar. Deste modo, a região delimitadora C-terminal do análogo de P.vivax conterá tipicamente, pelo menos, os aminoácidos 264-318 da proteína nativa e, opcionalmente, todos ou uma porção dos aminoácidos 319-355. De preferência consiste nos aminoácidos 264-335 da proteína nativa. A correspondente região C-terminal do análogo de P.falciparum conterá os aminoácidos 289-356 e, opcionalmente, todos ou uma porção dos aminoácidos 357-392.

As regiões N-terminal e C-terminal dos análogos podem variar na sequência de aminoácidos, em relação às correspondentes sequências CS nativas ou fragmentos destas, devido a uma ou mais substituições, adições, e/ou deleções ("deleção" refere-se à eliminação de resíduos internos e não das extremidades) prevenindo, no entanto, que cada uma continue a definir, pelo menos, um epitopo. Normalmente não haverá mais do que 10 variações por região, de preferência não



mais do que 5 variações por região.

Os análogos da proteína CS de P.vivax, P.falciparum e P.berghei deste invento podem ser representados pela seguinte fórmula:



onde (A_nB_m) representa uma região interna de repetições tandem, A representa uma das sequências repetitivas da proteína CS nativa, B representa outra das sequências repetitivas, n e m são números inteiros positivos, em que a soma de n e m é inferior ao número de sequências repetitivas na proteína CS nativa, X representa uma sequência delimitadora N-terminal e define, pelo menos, um epitopo sequencial da região N-delimitadora da proteína CS nativa, e Y representa uma sequência delimitadora C-terminal e define, pelo menos, um epitopo sequencial da região C-delimitadora da proteína CS nativa. No caso dos análogos de P.vivax, X representa pelo menos os aminoácidos 77-91 da proteína nativa e Y representa pelo menos os aminoácidos 264-318 da proteína nativa, enquanto que nos análogos de P.falciparum, X representa pelo menos os aminoácidos 107-123 da proteína nativa e Y representa pelo menos os aminoácidos 289-356 da proteína nativa.

Materiais recombinantes e métodos para produzir análogos da proteína CS

Os análogos da proteína CS deste invento são feitos através de técnicas de recombinação de DNA. Baseado no conhecimento das sequências de aminoácidos das sequências repetitivas nativas e das sequências delimitadoras, podem ser



preparados genes sintéticos codificando os análogos in vitro, através da síntese de fragmentos de oligonucleótidos apropriados do gene, do seu tratamento com cinases e ligando-os uns com os outros. Para a expressão num hospedeiro particular, pode ser desejável desenhar uma sequência de DNA sintético que utilize códons preferidos pelo hospedeiro particular, no qual o DNA vai ser expresso.

O DNA sintético que codifica os análogos pode ser clonado em qualquer réplica apropriada, de modo a criar um vector clone. São conhecidos numerosos vectores clone pelos especialistas nesta técnica, e a selecção de um vector clone apropriado é uma questão de escolha. Exemplos de vectores para clonar uma célula hospedeira que possam transformar, incluem o bacteriofago lambda (E.coli), pBR322 (E.coli), pACYC177 (E.coli), pKT 230 (bactéria gram-negativa), pME290 (bactéria gram-negativa que não a E.coli), pHV14 (E.coli e Bacillus subtilis), pBD9 (Bacillus), pIJ61 (Streptomyces), pUC6 (Streptomyces), actinofago phiC31 (Streptomyces), YIp5 (Saccharomyces), YCp19 (Saccharomyces), YEp24 e YEp13 (Saccharomyces), e o vírus bovino do papiloma (células de mamíferos).

As sequências polinucleotídicas que codificam os análogos das proteínas CS, são expressas pela inserção das sequências em réplicas apropriadas, criando, desse modo, vectores expressão, e introduzindo os vectores expressão resultantes em hospedeiros compatíveis.

Na criação de um vector expressão, a sequência que codifica o análogo da proteína CS está localizada no vector,



com as sequências controlo apropriadas. O posicionamento e a orientação da sequência código, em relação às sequências controlo, é tal que a sequência código é transcrita sob o controlo das sequências controlo: i.e., o promotor controla a transcrição do mRNA derivado da sequência código, e os ribossomas ligam-se ao local de ligação ribossómica para começar o processo de tradução, e o codão stop utilizado para terminar a tradução estará a montante do codão de terminação da transcrição. São preferíveis vectores de expressão de levedura. As sequências controlo para vectores de levedura incluem promotores para a síntese de enzimas glicolíticas (Hess et al., J Adv Enzyme Reg (1968) 7:149; Holland et al., Biochemistry (1978) 17:4900). Promotores adicionais conhecidos na técnica incluem o promotor do 3-fosfoglicerato cinase (Hitzeman et al., J Biol Chem (1980) 255:2073). Outros promotores, que têm a vantagem adicional de terem a transcrição controlada pelas condições de crescimento, e/ou património genético, são as regiões promotoras para o álcool desidrogenase 2 (ADH2), isocitocromo C, fosfatase ácida, enzimas degradativos associados com o metabolismo do nitrogénio, o sistema do factor alfa, e enzimas responsáveis pela utilização da maltose e da galactose. Sequências terminadoras da transcrição são desejáveis na terminação 3' das sequências código. Tais terminadores encontram-se na região não traduzida 3', a seguir à sequência código dos genes derivados da levedura.

Modificações da sequência que codifica o análogo, antes da sua inserção na réplica, podem ser desejáveis ou



necessárias, dependendo do sistema de expressão escolhido. Por exemplo, em alguns casos, pode ser necessário modificar a sequência, de forma a poder ligar-se às sequências controlo com a orientação apropriada, i.e., para manter a estrutura de leitura. Em alguns casos pode ser desejável adicionar sequências como a sequência "leader" do factor alfa da levedura, que causa a secreção do polipeptido a partir do organismo hospedeiro, com a subsequente clivagem do sinal secretor. As técnicas para modificar as sequências nucleotídicas, utilizando a clonagem, são bem conhecidas dos especialistas na técnica. Estas incluem, e.g., o uso de enzimas restritivos, enzimas como o Bal31, para remover os nucleótidos em excesso, oligonucleótidos quimicamente sintetizados para usar como adaptadores, para restituir os nucleótidos perdidos.

A sequência que codifica o análogo, modificada se necessário, pode ser ligada às sequências controlo antes da inserção num vector. Alternativamente, a sequência código pode ser clonada directamente num vector expressão que já contenha as sequências controlo e um local de restrição apropriado. Para expressão do análogo na levedura, as sequências controlo serão necessariamente heterólogas para a sequência código. Nos casos em que o gene análogo é para ser expresso numa linha celular derivada de vertebrados, as sequências controlo podem ser ou heterólogas ou homólogas, dependendo da linha celular específica.

Dependendo da célula hospedeira usada, a transformação é feita usando técnicas standard apropriadas



para tais células. Transformações na levedura podem ocorrer de acordo com o método de Beggs, J.D., Nature (1978) 275:104-109, ou de Hinnen, A., et al., Proc Natl Acad Sci (EUA) (1978) 75:1929.

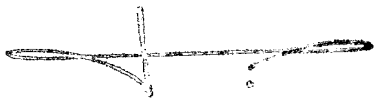
As células transformadas crescem, depois, em condições que permitem a expressão do gene do análogo da proteína CS e, se apropriado, conduzem à proteína madura. Porque o gene é expresso em organismos/células heterólogos, a proteína está livre das outras proteínas a que está associada nos mosquitos. O análogo recombinante da proteína CS assim sintetizado, é depois isolado das células hospedeiras e purificado. Se o sistema de expressão segrega o análogo para o meio de crescimento, o análogo é isolado directamente a partir do meio. Se o análogo recombinante não é segregado, ele é isolado a partir dos lisados das células. A selecção das condições de crescimento apropriadas e dos métodos de recolha encontra-se no âmbito dos especialistas na técnica. Em relação à purificação consulte, por exemplo, o pedido de patente da UNI descrito anteriormente.

Vacinas

As vacinas podem ser preparadas a partir dos análogos recombinantes da proteína CS de Plasmodium acima descritos que têm actividade neutralizadora do esporozoíto, ou melhorem ou previnam a malária por outros mecanismos. A formulação de vacinas que contêm polipeptido(s) imunogénicos como ingredientes activos é conhecida pelos especialistas na técnica. Tais vacinas são tipicamente preparadas como




injectáveis, ou em soluções líquidas ou em suspensões; formas sólidas apropriadas para solução ou suspensão num líquido antes da injeção, podem também ser preparadas. A preparação pode também ser emulsionada, ou a proteína encapsulada em liposomas. Os ingredientes imunogénicos activos são muitas vezes misturados com excipientes ou veículos que são farmacêuticamente aceitáveis e compatíveis com os ingredientes activos. Excipientes apropriados são, por exemplo, a água, soluções salinas ou de dextrose, glicerol, etanol, ou semelhantes e combinações destes. A concentração do análogo em formulações injectáveis será usualmente entre 0,2 e 5 mg/ml. Adicionalmente, se desejável, a vacina pode conter quantidades menores de substâncias auxiliares tais como agentes molhantes ou emulsionadores, agentes tampões de pH, e/ou adjuvantes que aumentem a eficácia da vacina. Exemplos de adjuvantes que podem ser efectivos incluem, não estando no entanto limitados a estes: hidróxido de alumínio, N-acetil-muramil-L-treonil-D-isoglutamina (thr-MDP), N-acetil-nor-muramil-L-alanil-D-isoglutamina (CGP 11637, referido como nor-MDP), N-acetilmuramil-L-alanil-D-isoglutamil-L-alanina-2-(1'-2'-dipalmitoil-sn-glicero-3-hidroxifosforiloxi)-etilamina (CGP 19835A, referido como MTP-PE), e RIBI, que contém três componentes extraídos de bactérias, monofosforil lípido A, dimicolato de trealose e o esqueleto da parede celular (MPL+TDM+CWS) numa emulsão a 2% de esqualeno/Tween 80. A eficiência do adjuvante pode ser determinada medindo directamente a quantidade de anticorpos contra os imunogénicos resultantes da administração do imunogénico em vacinas, que são



também constituídas por vários adjuvantes.

As vacinas são convencionalmente administradas parentericamente, por injeção, que pode ser, por exemplo, intravenosa, subcutânea ou intramuscular. Formulações adicionais, que são apropriadas para outros modos de administração incluem supositórios e, em alguns casos, formulações orais e liposomas. Para os supositórios, os veiculos e agentes de ligação tradicionais podem incluir, por exemplo, glicóis polialcalinos e triglicéridos; tais supositórios podem ser formados a partir de misturas contendo o análogo em cerca de 0,5 a 10 p%, de preferência 1-2p%. Formulações orais incluem os excipientes normalmente utilizados, tais como, por exemplo, classes farmacêuticas de manitol, lactose, amido, estearato de magnésio, sacarina de sódio, celulose, carbonato de magnésio, e semelhantes. Estas composições tomam a forma de soluções, suspensões, comprimidos, pilulas, cápsulas, formulações de libertação controlada, ou pós que contêm 10 a 95% do análogo, de preferência 25 a 70%.

Os análogos podem ser formulados nas vacinas em forma neutra ou em sal. Sais farmacêuticamente aceitáveis incluem os sais de adição ácidos (formados com grupos amina livres do péptido) e que são formados com ácidos inorgânicos tais como, por exemplo, os ácidos hidrolóricos e fosfóricos, ou com ácidos orgânicos tais como o acético, o oxálico, o tartárico, o maleico, e outros semelhantes. Sais formados com os três grupos carboxílicos livres podem também ser derivados de bases inorgânicas, tais como, por exemplo, os hidróxidos de



sódio, de potássio, de amónia, de cálcio ou de ferro, e bases orgânicas como a isopropilamina, trimetilamina, etanol de 2-etilamina, histidina, procaina, e semelhantes.

As vacinas são administradas de maneira compatível com a formulação da dose, a com uma quantidade tal que seja profilaticamente e/ou terapêuticamente eficiente. A quantidade a ser administrada, que é geralmente entre 5 microgramas e 250 microgramas do análogo por dose, depende do sujeito a ser tratado, capacidade do sistema imunológico do indivíduo para sintetizar anticorpos, e o grau de protecção desejado. As quantidades exactas de ingrediente activo requeridas para serem administradas podem depender da opinião do médico e podem ser peculiares a cada indivíduo.

A vacina pode ser aplicada segundo um calendário de uma só dose, ou, de preferência, com um calendário de múltiplas doses. Um calendário de múltiplas doses deverá ter uma primeira série de vacinações, que poderá ser de 1 a 10 doses separadas, seguida de outras doses dadas em intervalos de tempo subsequentes necessários para manter ou reforçar a resposta imunológica, por exemplo, entre 1-4 meses para uma segunda dose, e se necessário, dose(s) subsequente(s) passados alguns meses.

Exemplos


Os seguintes exemplos ilustram os análogos do invento e métodos para a sua preparação. Não pretendem limitar, de forma alguma, o invento.



Vector expressão do Vivax 3

O Vivax 3 é um análogo da proteína CS de P.vivax que consiste nos aminoácidos 23-91 da proteína CS nativa de P.vivax (com substituição Gly-->Arg na posição 91), seguida de sequência "A" simples (Asp-Arg-Ala-Asp-Gly-Gln-Pro-Ala-Gly), seguida de uma sequência "B" simples (Asp-Arg-Ala-Ala-Gly-Gln-Pro-Ala-Gly), seguida dos aminoácidos 264-335 (com uma Gly adicionada entre as posições 277-278, uma substituição Glu-->Ala na 290, uma substituição Val-->Leu na 300, uma substituição Arg-Glu na 301 e uma substituição Ala-->Thr na 302) da proteína CS nativa de P.vivax. A sequência de aminoácidos e a sequência de DNA codificante para o Vivax 3 estão ilustradas na figura 1.

Vinte e dois oligômeros sintéticos, que se mostram na figura 2, foram feitos num "Applied Biosystem 380A DNA synthetizer". Estes oligonucleótidos foram tratados com cinases, e hibridizados com fervura e arrefecimento lento até à temperatura ambiente e depois ligados para formar a sequência de DNA que codifica Vivax 3 (ver esquema da figura 2). O gene completo para Vivax 3 foi excisado de um gel de acrilamida a 7% e clonado no plasmideo pBS100, que foi digerido com NcoI e SalI. O pBS100 está esquematicamente representado na Figura 3 e contém a estrutura interna do promotor híbrido ADH2-GAPDH e do terminador GAPDH. Um fragmento de BamHI/SalI contendo o promotor, a sequência de DNA que codifica o Vivax 3, e o terminador foi isolado do vector de clonagem e inserido no plasmideo pBS24.1 para a



produção do vector de expressão pBS24vivax3 para a expressão intracelular em levedura. O pBS24.1 está esquematicamente representado na figura 4.

Vector de expressão do Vivax 3.1

O Vivax 3.1 é um análogo da proteína CS de P.vivax, semelhante ao Vivax 3 com a excepção de não ter as variações referidas nos aminoácidos das regiões delimitadoras da sequência nativa. A sequência de aminoácidos e a sequência de DNA codificadora do Vivax 3.1 estão ilustradas na figura 5.

O gene que codifica o Vivax 3.1 foi preparado como acima descrito, usando os oligómeros e o esquema de ligações que se mostra na figura 6. Este gene foi clonado no plasmídeo pAB125, que tinha sido digerido com XbaI e SalI. O pAB125 está representado esquematicamente na figura 7 e contém a estrutura interna do promotor híbrido ADH2-GAPDH, a sequência "leader" do factor alfa e o terminador do factor alfa de S.cerevisiae. Um fragmento BamHI/SalI contendo o promotor, a sequência "leader" do factor alfa, o gene de Vivax 3.1 e o terminador, foi isolado a partir do vector de clonagem e inserido no plasmídeo pBS24.1 para produção do vector de expressão pBS24vivax3.1 para expressão secretora em levedura.

Vector de expressão do Falciparum 3

Falciparum(Falc) 3 é um análogo da proteína CS de P.falciparum que consiste nos aminoácidos 43-123 da proteína CS nativa de P.falciparum, seguido por quatro sequências repetitivas (três "B", i.e., -Asn-Ala-Asn-Pro- e uma A,

i.e., Asn-Val-Asp-Pro, num padrão BABB), seguido pelos aminoácidos 289-374 da proteína nativa. A sequência de aminoácidos e a sequência de DNA codificadora para Falc 3 estão ilustradas na figura 8.


O gene que codifica o Falc 3 foi sintetizado a partir de oligômeros (ver figura 9), clonado no pAB125 e o resultante factor alfa-Falc 3-ADH2/GAPDH construído e clonado no pBS24.1 para a produção do vector de expressão pBS24falc3 para expressão secretora do Falc 3 em levedura.

Vector de expressão do Falciparum 4

O Falc 4 é um análogo da proteína CS tendo uma sequência similar ao Falc 3, com a excepção de ter uma região amina delimitadora mais pequena e uma região carboxilica delimitadora mais longa do que Falc 3. A sequência de aminoácidos e a sequência de DNA que codifica o Falc 4 estão ilustradas na figura 10. Como indicado, a sequência inclui os aminoácidos 68-123 da sequência da proteína CS nativa de P.falciparum, quatro repetições e os aminoácidos 289-392 da sequência nativa. Os aminoácidos adicionais no N-terminal da sequência são parte da sequência "leader" do factor alfa.

O DNA que codifica a sequência do Falc 4, ilustrado na figura 10, foi sintetizado e clonado no pAB125 e depois no pBS24.1, como descrito acima, para providenciar o vector de expressão pBS24falc4, para expressão secretora em levedura.

Vector de expressão do Berghei 3

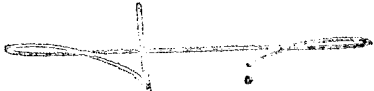


O Berghei 3 é um gene sintético que contém regiões delimitadoras amina e carboxi terminal e apenas 2 das 13 repetições tandem de oito aminoácidos, que se encontram no gene CS de P.berghei. As 16 repetições QP que seguem as repetições de aminoácidos são também eliminadas. A sequência para o gene e os aminoácidos nela codificados estão ilustrados na figura 11.

Então, a construção começa na Gly 63 e continua através da primeira repetição PPPNPND terminando no aminoácido 100. De notar que os aminoácidos 69-74, publicados originalmente por Eichinger et al. como PMLRRK, foram mudados para a sequência "correcta" de de la Cruz et al., ADAPEG. A segunda repetição, PAPPNAND, é derivada da primeira que se encontra mais à frente na sequência. A terminação carboxilica continua então na Gly 239 até à Ser 318. Como se mostra na figura 12, o gene foi feito com 20 48-mers e 2 28-mers para facilitar longas sobreposições na hibridização do gene. Os codões preferidos da levedura foram usados onde possível, e os locais de NcoI e SallI incorporados no codão de iniciação 5' e no codão stop 3', respectivamente.

O gene foi montado através do tratamento por cinases de cada oligómero, mergulhando-os num tampão com ligases e depois dissociando-os a 100 °C e deixando-os arrefecer à temperatura ambiente. A ligase e o ATP foram adicionados e o gene ligado foi depois cortado com enzimas restritivas NcoI e SallI, e purificado por electroforese de gel de agarose .


O gene está previsto para ser expresso no promotor



regulável ADH2-GAPDH da levedura para expressão interna (ver acima). O Berghei 3 pode também ser expresso através de secreção por clonagem no pAB125, um vector contendo a sequência promotora/"leader" do factor alfa-ADH2 (ver acima).

Transformação da levedura com os vectores de expressão do análogo da proteína CS

A estirpe AB110 de Saccharomyces cerevisiae (Mat, leu2-04, ou leu2-3 e leu2-112 (ambos), pep3-4, his4-580, cir^o) foi transformada com os vectores de expressão (à excepção da construção de Falc 4) de acordo com Hinnen et al., Proc Natl Acad Sci (USA) 75:1929-1933 (1978). Colónias de transformantes simples contendo vectores reguladores ADH2/GAP cresceram em 2 ml de meio selectivo para a leu- (deficiente em leucina) até à fase log tardia ou fase estacionária. Apenas as leveduras que contêm tais vectores, crescem neste meio. As culturas foram subsequentemente diluídas de 1:20 (v/v) em YEP (1% p/v de extrato de levedura, 2% p/v de peptona) com 1% de glucose, e crescem até à saturação (cerca de 36 h) neste meio. As células lisaram na presença de dodecil sulfato de sódio (SDS) e ditioneitol (DTT) e os lisados foram limpos por centrifugação. Os lisados limpos foram sujeitos a electroforese de gel de poliacrilamida (Laemmli, U. K., Nature (1970) 277:680). Após a coloração com Coomassie blue, foram observadas bandas de tamanho apropriado em estratos de transformantes contendo o plasmideo de P.vivax e o plasmideo de P.falciparum. Estas bandas foram detectadas nas células transformadas com os respectivos vectores de expressão,




enquanto se encontram ausentes nos estratos de células contendo plasmídeos de controlo.

Para confirmar a identidade das bandas, as proteínas foram submetidas a análises Western. Os lisados de levedura limpos e preparados como descrito acima foram sujeitos a electroforeses em geles de poliacrilamida (Laemmli, supra) e as proteínas foram subsequentemente electrotransferidas em filtros de nitrocelulose (Towbin et al., Proc Natl Acad. Sci (USA) (1979) 76:3450). O filtro foi preincubado durante 1 hora com 1% (BSA) em PBS e subsequentemente tratado com anticorpos para as proteínas CS nativas durante 12 horas a 4 °C . Os filtros foram lavados com 1% BSA/PBS e foi adicionado um segundo anticorpo de cabra anti-rato conjugado com peroxidase de rábano (não puro) (Bio-Rad Laboratories, Richmond, California). Finalmente, os filtros foram incubados com um reagente de desenvolvimento de coloração de peroxidase de rábano (não puro) (Bio-Rad, Richmond, California) e lavados. As análises da Western mostraram que as proteínas reagiram com os anticorpos.

Produção em leveduras de análogo da proteína CS
(intracelular) e purificação

Os transformantes cresceram em meio selectivo (tanto para Leu- como para Ura-) para aumentar o número de cópias de plasmídeo e cresceram, por outro lado, em altas concentrações de glucose (8%) para manter reprimida a expressão.

A levedura passou então para um meio de YEP com 2% de glucose (um meio rico para a expressão intracelular), onde



ficou durante 48 horas (a expressão vê-se a partir das 24 horas).


A levedura foi retirada do meio por filtração ou centrifugação e a levedura foi quebrada num tampão apropriado. A purificação das proteínas foi iniciada através de precipitação selectiva ao calor, após diluição com igual volume de água e o ajuste do pH para 8.0 (a 100 °C durante 20 minutos; isto não serve apenas de etapa de purificação mas também de etapa de inibição de proteases). O precipitado foi removido por filtração ou centrifugação.

O análogo da proteína CS no sobrenadante foi de seguida purificado por cromatografia de troca iónica num permutador de catiões, através do abaixamento do pH para 4.0, de forma a permitir a ligação. Após a lavagem da coluna a pH 5.0, é usado um gradiente de sal para remover selectivamente as proteínas, incluindo o análogo da proteína CS. O análogo é recolhido, concentrado e sujeito a cromatografia de filtração molecular.

A proteína CS purificada foi concentrada e diafiltrada. Foi armazenada, ou por congelamento a menos 80⁰, ou por secagem fria, e guardada sem humidade num frasco selado.

Produção em levedura do análogo da proteína CS
(secreção) e purificação

Os transformantes cresceram em meio selectivo (tanto para Leu- como para Ura-) para aumentar o número de cópias de plasmídeo e cresceram, por outro lado, em altas concentrações



de glucose (8%) para manter reprimida a expressão.

A levedura passou então para um meio de secreção de levedura com 2% de glucose (baixo nível de proteínas) e cresceram durante 90 horas (a expressão vê-se a partir das 36 horas, mas as 90 horas permitem a acumulação do análogo da proteína CS no meio).


A levedura é removida do meio por filtração ou centrifugação, o sobrenadante é retido e o pH é ajustado para 8.0 (NaOH). A purificação das proteínas é iniciada por precipitação selectiva ao calor (a 100 °C durante 20 minutos, o que não serve apenas como etapa de purificação mas também como etapa de inibição das proteases). Este precipitado é recolhido por filtração ou centrifugação.

A purificação do análogo a partir do sobrenadante foi realizada como acima descrito.

Novas modificações de formas de realização do invento acima descrito que são evidentes para os especializados nos campos da biologia molecular, bioquímica, imunologia e campos relacionados, incluem-se no espírito do invento.

R E I V I N D I C A Ç O E S

1- Um método de preparação de um polinucleótido que codifica uma proteína do circunsporozoito de Plasmodium (CS), caracterizado pelo facto de compreender:



- a síntese de uma sequência de polinucleótido que codifica, pelo menos, duas repetições da sequência repetitiva da proteína CS nativa, mas que codifica menos repetições da sequência repetitiva do que as que estão presentes na proteína CS nativa, e pelo menos um epítipo da região carboxi-terminal, não repetitiva, que flanqueia a proteína CS nativa.


2- Um método, conforme reivindicado na reivindicação 1, caracterizado pelo facto de compreender ainda as fases de:

(a) síntese ou obtenção de uma sequência nucleotídica amina terminal que codifica, pelo menos, um epítipo da região amina terminal, não repetitiva, que flanqueia a proteína CS nativa;

(b) ligação à extremidade 3' da sequência nucleotídica amina terminal, de uma sequência nucleotídica interna que codifica pelo menos, duas repetições de uma sequência repetitiva da proteína CS nativa, mas que codifica menos repetições da referida sequência repetitiva, do que as que estão presentes na proteína CS nativa; e

(c) ligação à extremidade 3' da sequência nucleotídica interna de uma sequência nucleotídica carboxi-terminal, que codifica, pelo menos, um epítipo da região não repetitiva carboxi-terminal, que flanqueia a proteína CS nativa.

3- Um método, conforme reivindicado na reivindicação 2, caracterizado ainda pelo facto de na fase (a) a sequência nucleotídica amina terminal codificar, pelo menos, os aminoácidos 23 a 76 da proteína CS nativa do *P. vivax* e, na fase (b) a sequência nucleotídica interna codifica ambos os tipos de sequências repetitivas da proteína CS nativa do *P. vivax*, e existindo 2 a 18 repetições de sequências repetitivas, e na fase (c) a sequência nucleotídica carboxi terminal codifica,




pelo menos, os aminoácidos 264 a 318 da proteína CS nativa do P. vivax.

4- Um método, conforme reivindicado na reivindicação 2, caracterizado pelo facto de, na fase (a), a sequência nucleotídica amina terminal codificar, pelo menos, os aminoácidos 23 a 91 da proteína CS nativa do P. vivax e, em que, na fase (b) a sequência nucleotídica interna codifica ambos os tipos de sequências repetitivas da proteína CS nativa do P. vivax, e existindo 2 a 10 repetições de sequências repetitivas, e, na fase (c) a sequência nucleotídica carboxi-terminal codifica, pelo menos, os aminoácidos 264 a 335 da proteína CS nativa do P. vivax.

5- Um método, conforme reivindicado na reivindicação 2, caracterizado pelo facto de, na fase (a), a sequência nucleotídica amina terminal codificar, pelo menos, os aminoácidos 107 a 123 da proteína CS nativa do P. falciparum, e, na fase (b) a sequência nucleotídica interna codificar ambos os tipos de sequências repetitivas da proteína CS nativa do P. falciparum, e existindo 2 a 40 repetições de sequências repetitivas, e, na fase (c) a sequência nucleotídica carboxi-terminal codifica, pelo menos, os aminoácidos 289 a 374 da proteína CS nativa do P. falciparum.

6- Um método, conforme reivindicado na reivindicação 2, caracterizado também pelo facto de, na fase (a), a sequência nucleotídica amina terminal codificar, pelo menos, os aminoácidos 68 a 123 da proteína CS nativa do P. falciparum, e, na fase (b) a sequência nucleotídica interna codificar ambos os tipos de sequências repetitivas da proteína CS nativa do P. falciparum, e existindo 2 a 20 repetições da sequência repetitiva, e, na fase (c), a sequência nucleotídica carboxi-terminal codificar, pelo



menos, os aminoácidos 289 a 374 da proteína CS nativa do *P. falciparum*.

7- Um método, conforme reivindicado na reivindicação 2, caracterizado pelo facto da sequência polinucleotídica codificar a sequência de aminoácidos representada nas Figuras 1, 5, 8, 10 ou 11.

8- Um método para a preparação de um vector de expressão para a expressão de um análogo imunogénico da proteína CS do *Plasmodium*, numa célula, caracterizado pelo facto de compreender:

(a) a síntese de uma sequência polinucleotídica, conforme reivindicado nas reivindicações de 1 a 7; e

(b) a ligação de sequências de controlo, unidas, de modo operável, à sequência polinucleotídica, e eficazes na condução da expressão da sequência nucleotídica do análogo da proteína, na célula.

9- Um método para a preparação de uma célula recombinante capaz de produzir um análogo da proteína CS do *Plasmodium*, caracterizado pelo facto de compreender a inserção do vector de expressão da reivindicação 8 numa célula hospedeira compatível.

10- Um método para a preparação de um análogo imunogénico da proteína CS do *Plasmodium*, caracterizado pelo facto de compreender o crescimento da célula, conforme reivindicado na reivindicação 8, sob condições que permitam a produção da proteína codificada pelo polinucleótido.

11- Um método para a preparação de uma vacina contra a malária, caracterizado pelo facto de compreender:

(a) a obtenção de uma quantidade de um análogo imunogénico da proteína CS do *plasmodium*, preparado pelo método da reivindicação

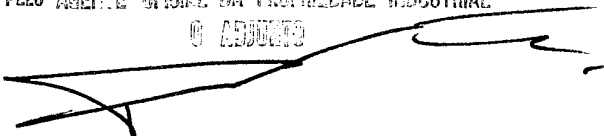
10; e

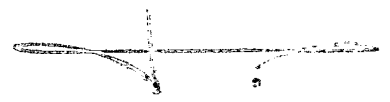
(b) misturar o análogo da proteína CS do plasmodium com um veículo farmacêuticamente aceitável.

Lisboa, 10 de Abril de 1990

PELO AGENTE OFICIAL DA PROPRIEDADE INDUSTRIAL

O ADJUNTO





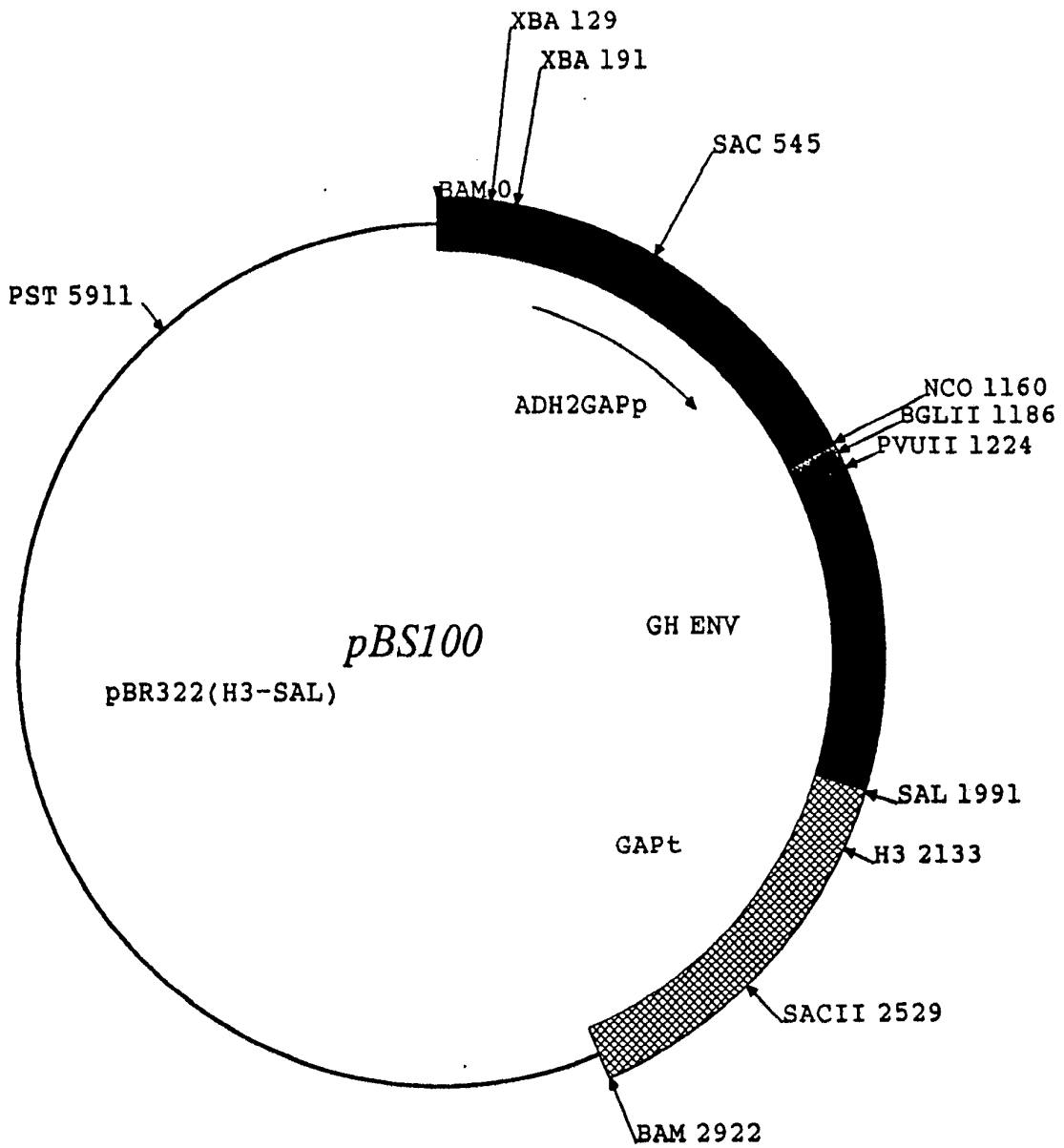
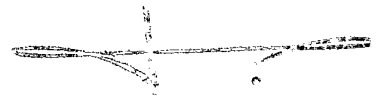
Oligonucleótidos Sintéticos para o Vivax3

	5'		3'
1-27	CATGGGTCACAACGTAGATTTGTGTAA		
2-46	GGCTATCAACTTGAACGAAGTTGACGCTAGCTCTTTGGGTGCTGCT		
3-46	CACGTTGGTCAATCTGCTTCTAGAGGTAGAGGTTTGGGTGAAAACC		
4-46	CAGACGACGAAGAAGGTGACGCTAAGAAGAAGAAGGACGGTAAGAA		
5-46	GGCTGAACCAAAGAACCCTCGAGAAAACAAGTTGAAGCAACCAAGA		
6-46	GACAGAGCTGACGGTCAACCAGCTGGTGATAGGGCCGCTGGCCAGC		
7-46	CTGCCGGCAACGGTGCTGGTGGTCAAGCCGCGGGTGGTAACGCTGG		
8-46	TGGTGGTCAAGGTCAAAAACAACGAAGGTGCTAACGCTCCAACGCT		
9-46	AAGTCTGTTAAGGAATACTGGACAAGTTGGAAACCACCGTTGGTA		
10-46	CCGAATGGACCCCATGTTCTGTTACCTGTGGTGTGGTGTAGAGT		
11-46	TAGATCCAGAGTTAACGCTGCTAACAGAAGCCAGAAGATCTATAG		
12*46	CAACTTCGTTCAAGTTGATAGCCTTACACAAATCTACGTTGTGACC		
13*46	CTAGAAGCAGATTGACCAACGTGAGCAGCACCCAAAGAGCTAGCGT		
14*46	AGCGTCACCTTCTTCGTCGCTCTGGGTTTTCACCCAAACCTCTACCT		
15*46	CTCGAGGGTTCCTTGGTTTCAGCCTTCTTACCGTCCTTCTTCTTCTT		
16*46	GCTGGTTGACCGTCAGCTCTGTCTCTTGGTTGCTTCAACTTGTGTTT		
17*46	ACCACCAGCACCGTTGCCGGCAGGCTGGCCAGCGGCCCTATACCA		
18*46	CGTTGTTTTGACCTTGACCACCACCAGCGTTACCACCCGCGGCTTG		
19*46	TCCAAGTATTCCTTAACAGACTTAGCGTTTGGAGCGTTAGCACCTT		
20*46	AACAGAACATGGGGTCCATTCCGGTACCAACGGTGGTTTCCAACCTG		
21*46	TAGCAGCGTTAACTCTGGATCTAACTCTAACACCAACACCACAGGT		
22*27	TCGACTATAGATCTTCTGGCTTCTTGT		

Vivax 3

NcoI	1-27	2-46	3-46	4-46	5-46	6-46	7-46	8-46	9-46	10-46	11-46	Sall
	12*46	13*46	14*46	15*46	16*46	17*46	18*46	19*46	20*46	21*46	22*27	

FIG. 2



BAM:	0
XBA:	129
XBA:	191
SAC:	545
NCO:	1160
BGLII:	1186
PVUII:	1224
SAL:	1991
H3:	2133
SACII:	2529
BAM:	2922
PST:	5911





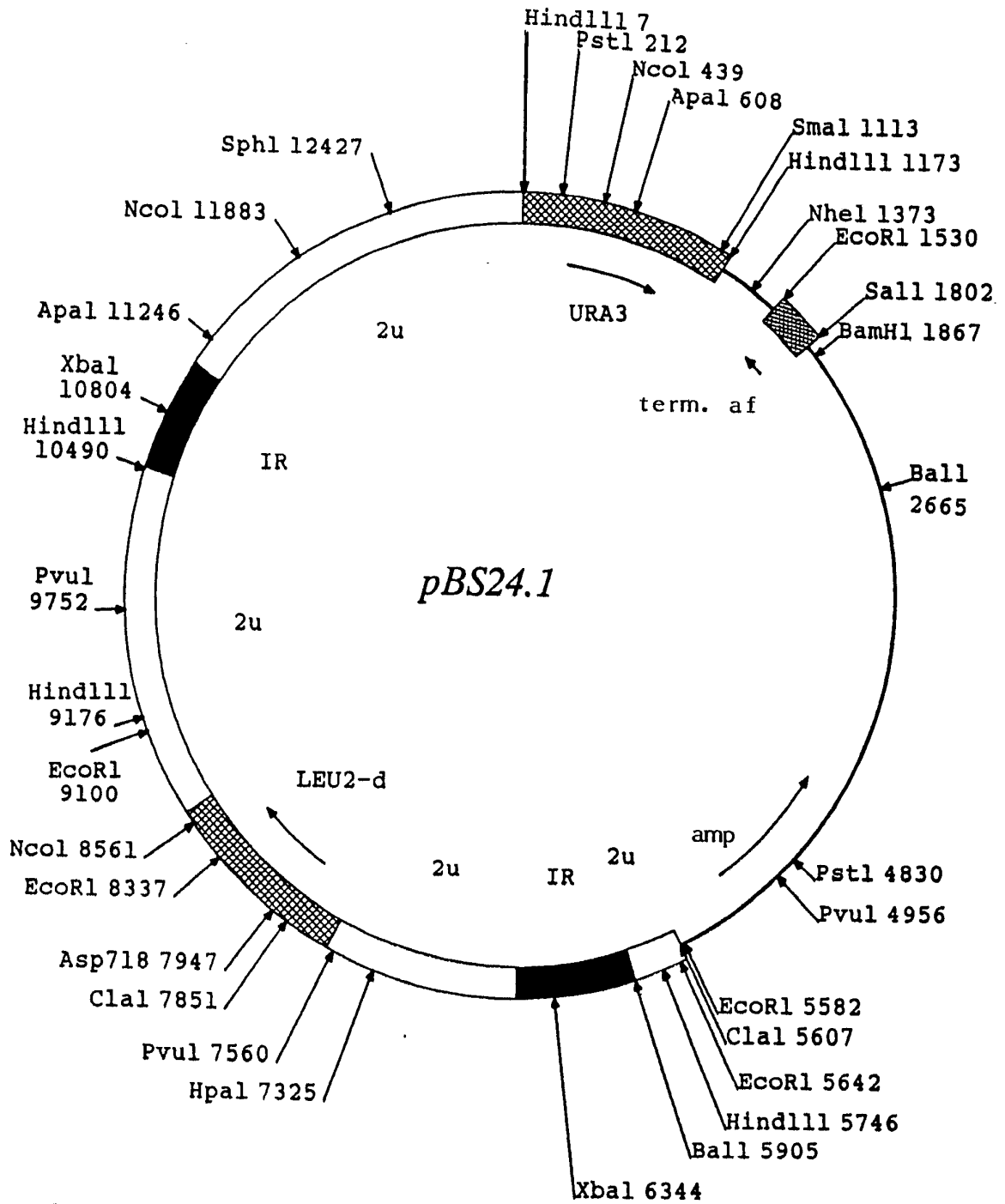
	ADH2GAPp:	0	1169
	GH ENV:	1169	1987
	GAPt:	1987	2922
	pBR322 (H3-SAL) :	2922	6663

FIG. 3














Clal:	1	Clal:	5607		URA3:	1	1173
HindIII:	7	EcoRI:	5642		term. Af:	1530	1802
PstI:	212	HindIII:	5746		amp:	4000	5607
NcoI:	439	Ball:	5905		2u:	5607	5900
Apal:	608	XbaI:	6344		IR:	5900	6550
SmaI:	1113	HpaI:	7325		2u:	6550	7600
HindIII:	1173	PvuI:	7560		LEU2-d:	7600	8630
NheI:	1373	ClaI:	7851		2u:	8630	10500
EcoRI:	1530	Asp718:	7947		IR:	10500	11100
Sall:	1802	EcoRI:	8337		2u:	11100	13143
BamHI:	1867	NcoI:	8561				
Ball:	2665	EcoRI:	9100				
PstI:	4830	HindIII:	9176				
PvuI:	4956	PvuI:	9752				
EcoRI:	5582	HindIII:	10490				

FIG. 4



FIG. 5

23 41
 3 MetGlyHisAsnValAspLeuCysLysAlaIleAsnLeuAsnGluValAspAlaSerSer
 ATGGGTCAACAACGTAGATTTGTGTAAGGCTATCAACTTGAACGAAGTTGACGCTAGCTCT
 TACCCAGTGTTCATCTAAACACATTCGGATAGTTGAACTTGCTTCAACTGCGATCGAGA
 1 NCOI, 54 NHEI,
 61
 63 LeuGlyAlaAlaHisValGlyGlnSerAlaSerArgGlyArgGlyLeuGlyGluAsnPro
 TTGGGTGCTGCTCACGTTGGTCAATCTGCTTCTAGAGGTAGAGGTTTGGGTGAAAACCCA
 AACCCACGACGAGTGCAACCAGTTAGACGAAGATCTCCATCTCCAAACCCACTTTTGGGT
 93 XBAI,
 81
 123 AspAspGluGluGlyAspAlaLysLysLysLysAspGlyLysLysAlaGluProLysAsn
 GACGACGAAGAAGGTGACGCTAAGAAGAAGAAGGACGGTAAGAAGGCTGAACCAAAGAAC
 CTGCTGCTTCTTCCACTGCGATTCTTCTTCTTCTGCCATTCTTCCGACTTGGTTCTTG
 183 ProArgGluAsnLysLeuLysGlnProGlyAspArgAlaAspGlyGlnProAlaGlyAsp
 CCTCGAGAAAACAAGTTGAAGCAACCAGGAGACAGAGCTGACGGTCAACCAGCTGGTGAT
 GGAGCTCTTTTGTTCACCTTCGTTGGTCTCTGCTCGACTGCCAGTTGGTTCGACCACTA
 184 XHOI, 232 PVU2,
 264 275
 243 ArgAlaAlaGlyGlnProAlaGlyAsnGlyAlaGlyGlyGlnAlaAlaGlyGlyAsnAla
 AGGGCCGCTGGCCAGCCTGCCGGCAACGGTGCTGGTGGTCAAGCCGCGGGTGGTAACGCT
 TCCCGCGACCGGTCGGACCGCCGTTGCCACGACCACCAGTTCGGCGCCACCATTGCGA
 251 BALI, 261 NAEI, 286 SAC2,
 295
 303 GlyGlyGlnGlyGlnAsnAsnGluGlyAlaAsnAlaProAsnGluLysSerValLysGlu
 GGTGGTCAAGGTCAAACAACGAAGGTGCTAACGCTCAAACGAGAAGTCTGTAAAGGAA
 CCACCAGTTCAGTTTTGTTGCTTCCACGATTGCGAGGTTTGTCTTCAGACAATTCCTT
 315
 363 TyrLeuAspLysValArgAlaThrValGlyThrGluTrpThrProCysSerValThrCys
 TACTTGGACAAGGTTAGAGCTACCGTTGGTACCGAATGGACCCCATGTTCTGTTACCTGT
 ATGAACCTGTTCCAATCTCGATGGCAACCATGGCTTACCTGGGGTACAAGACAATGGACA
 390 KPNI,
 335
 423 GlyValGlyValArgValArgSerArgValAsnAlaAlaAsnLysLysProGluAspLeu
 GGTGTTGGTGTAGAGTTAGATCCAGAGTTAACGCTGCTAACAAGAAGCCAGAAGATCTA
 CCACAACCACAATCTCAATCTAGGTCTCAATTGCGACGATTGTTCTTCGGTCTTCTAGAT
 450 HPAI, 476 BGL2,
 AM Ser
 483 TAGTCGAC
 ATCAGCTG
 485 SALI,

Oligonucleótidos Sintéticos para o Vivax3.1

	5'		3'
5-28	TCGAGAAAACAAGTTGAAGCAACCAGGA		
8-43	TGGTCAAGGTCAAAACAACGAAGGTGCTAACGCTCCAAACGAG		
9-43	AAGTCTGTTAAGGAATACTTGGACAAGGTTAGAGCTACCGTTG		
16*47	GCTGGTTGACCGTCAGCTCTGTCTCCTGGTTGCTTCAACTTGTTTTT		
18*43	CGTTGTTTTGACCTTGACCACCAGCGTTACCACCCGCGGCTTG		
19*46	TCCAAGTATTCCTTAACAGACTTCTCGTTTGGAGCGTTAGCACCTT		
20*24	GTACCAACGGTAGCTCTAACCTTG		

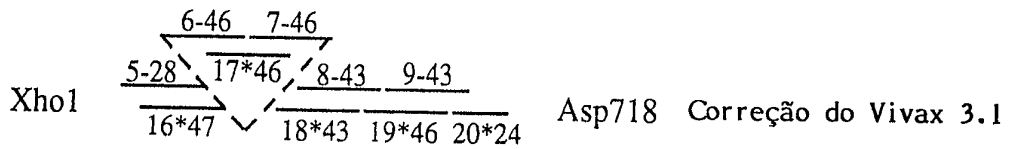


FIG. 6

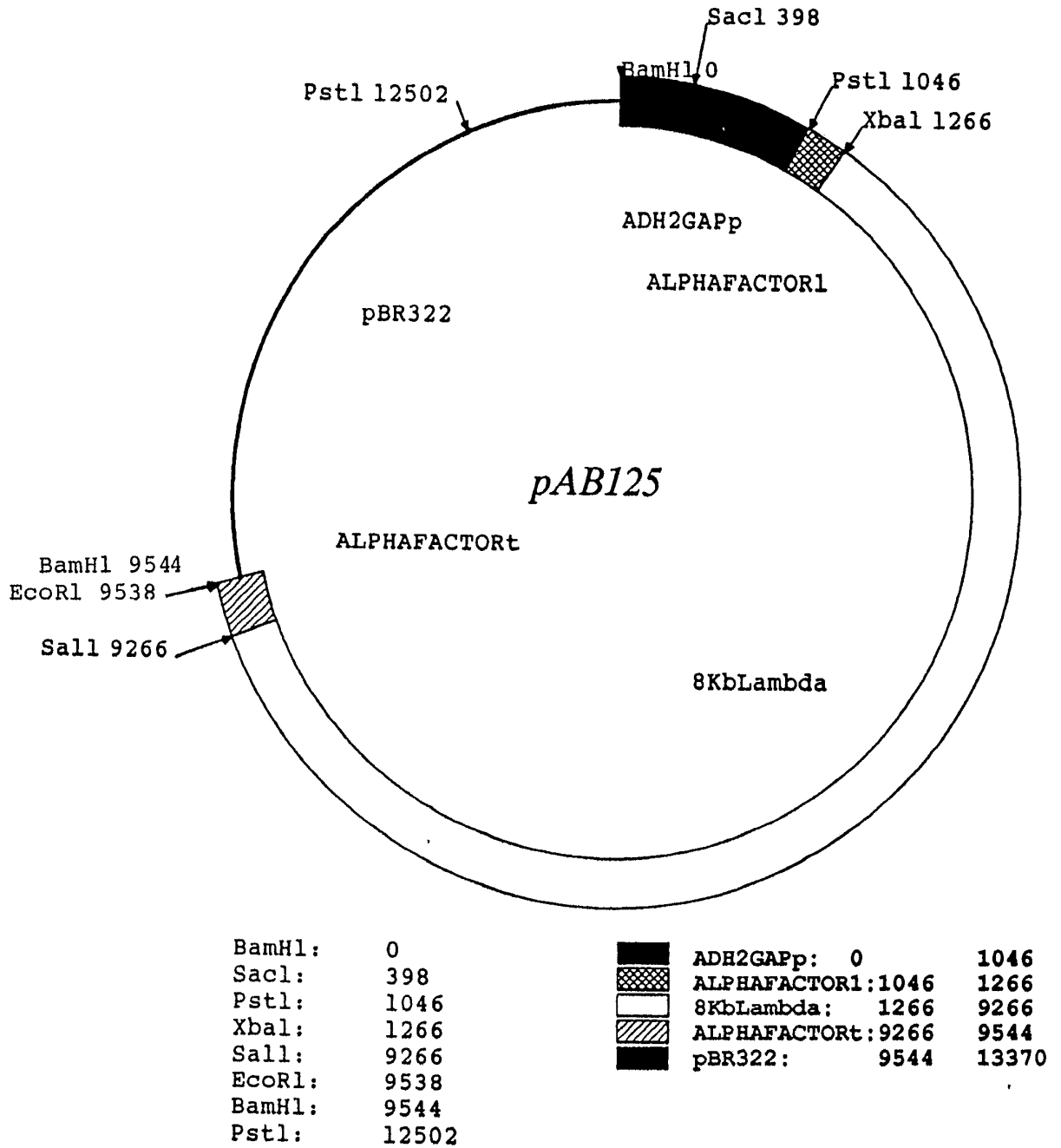
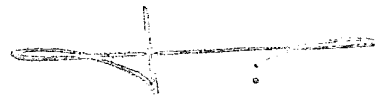


FIG. 7



FIG. 8

43
 MetAlaGlyThrAsnLeuTyrAsnGluLeuGluMetAsnTyrTyrGlyLysGlnGluAsn
 3 ATGGCTGGTACCAACTTGTACAACGAATTGGAAATGAACTACTACGGTAAGCAAGAAAAC
 TACCGACCATGGTTGAACATGTTGCTTAACCTTTACTTGGATGATGCCATTTCGTTCTTTTG
 1 NCOI, 9 KPNI,

61
 TrpTyrSerLeuLysLysAsnSerArgSerLeuGlyGluAsnAspAspAlaAsnAsnAsn
 63 TGGTACTCTTTGAAGAAGAAGCTTAGGTCTTTGGGTGAAAACGATGACGCTAACAACAAC
 ACCATGAGAAACTTCTTCTTGAGATCCAGAAACCCACTTTTGCTACTGCGATTGTTGTTG
 81

101
 AsnGlyAspAsnGlyArgGluGlyLysAspGluAspLysArgAspGlyAsnAsnGluAsp
 123 AACGGTGACAACGGTAGAGAAGGTAAGGACGAAGACAAGAGACGGTAACAACGAAGAC
 TTGCCACTGTTGCCATCTCTTCCATTCTGCTTCTGTTCTCTGCCATTGTTGCTTCTG
 121

183 AsnGluLysLeuArgLysProLysHisLysLysLeuLysGlnProGlyAspGlyAsnPro
 AACGAAAAGTTGAGAAAGCCAAAGCACAAGAAGCTTAAGCAACCAGGTGACGGTAACCCA
 TTGCTTTTCAACTCTTTCGGTTTCGTGTTCTTCGAATTGTTGTTGCCACTGCCATTGGGT
 213 HIND3, 216 AFL2, 234 BSTE2,

289290
 AspProAsnAlaAsnProAsnValAspProAsnAlaAsnProAsnAlaAsnProAsnLys
 243 GACCCAAACGCTAACCCAAACGTGGATCCTAACGCCAACCCCTAATGCTAATCCAAACAAG
 CTGGGTTTTCGATTGGGTTTGACACTAGGATTGCGGTTGGGATTACGATTAGGTTTGTTC
 266 BAMHI,

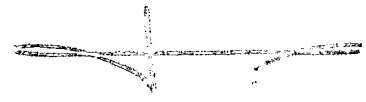
310
 AsnAsnGlnGlyAsnGlyGlnGlyHisAsnMetProAsnAspProAsnArgAsnValAsp
 303 AACCAACCAAGGTAACGGTCAAGGTCACAACATGCCAAACGACCCAAACAGAAACGTTGAC
 TTGTTGGTTCCATTGCCAGTTCAGTGTGACGGTTTGCTGGGTTTGTCTTTGCAACTG
 330

363 GluAsnAlaAsnAlaAsnAsnAlaValLysAsnAsnAsnAsnGluGluProSerAspLys
 GAGAACGCTAACGCTAACAACGCTGTTAAGAACAACAACAACGAAGAACCATCTGACAAG
 CTCTTGCGATTGCGATTGTTGCGACAATTCTTGTTGTTGTTGCTTCTTGGTAGACTGTTC
 350

423 HisIleGluGlnTyrLeuLysLysIleLysAsnSerIleSerThrGluTrpSerProCys
 CACATCGAACAATACTTGAAGAAGATCAAGAAGCTATCTCTACCGAATGGTCTCCATGT
 GTGTAGCTTGTATGAAGTCTTCTAGTTCCTTGAGATAGAGATGGCTTACCAGAGGTACA
 370

483 SerValThrCysGlyAsnGlyIleGlnValArgIleLysProGlySerAlaAsnLysPro
 TCTGTTACCTGTGGTAACGGTATCCAAGTTAGAATCAAGCCAGGTTCTGCTAACAAGCCT
 AGACAATGGACACCATTGCCATAGGTTCAATCTTAGTTCCGGTCCAAGACGATTGTTCCGA
 523 ALWN1, 540 MST2,

374
 LysAspGluLeuAM Ser
 543 AAGGACGAATTGTAGTCGAC
 TTCTGCTTAACATCAGCTG
 557 SALI,



Oligonucleótidos Sintéticos para Falc3

	5'		3'
1-28	CATGGCTGGTACCAACTTGTACAACGAA		
2-48	TTGGAAATGAACTACTACGGTAAGCAAGAAAACCTGGTACTCTTTGAAG		
3-48	AAGAACTCTAGGTCTTTGGGTGAAAACGATGACGCTAACCAACAACAAC		
4-48	GGTGACAACGGTAGAGAAGGTAAGGACGAAGACAAGAGAGACGGTAAC		
5-48	AACGAAGACAACGAAAAGTTGAGAAAGCCAAAGCACAAGAAGCTTAAG		
6-48	CAACCAGGTGACGGTAACCCAGACCCAAACGCTAACCCAAACGTGGAT		
7-48	CCTAACGCCAACCCTAATGCTAATCCAAACAAGAACAACCAAGGTAAC		
8-48	GGTCAAGGTCACAACATGCCAAACGACCCAAACAGAAACGTTGACGAG		
9-48	AACGCTAACGCTAACACGCTGTTAAGAACAACAACAACGAAGAACCA		
10-48	TCTGACAAGCACATCGAACAATACTTGAAGAAGATCAAGAACTCTATC		
11-48	TCTACCGAATGGTCTCCATGTTCTGTTACCTGTGGTAACGGTATCCAA		
12-48	GTTAGAATCAAGCCAGGTTCTGCTAACAAGCCTAAGGACGAATTGTAG		
13*48	CTTACCGTAGTAGTTCATTTCCAATTCGTTGTACAAGTTGGTACCAGC		
14*48	TTCACCCAAAGACCTAGAGTTCTTCTTCAAAGAGTACCAGTTTTCTTG		
15*48	CTTACCTTCTCTACCGTTGTCACCGTTGTTGTTGTTAGCGTCATCGTT		
16*48	TCTCAACTTTTCGTTGCTTTCGTTGTTACCGTCTCTCTTGTCTTCGTC		
17*48	GTCTGGGTTACCGTCACCTGGTTGCTTAAGCTTCTTGTGCTTTGGCTT		
18*48	ATTAGCATTAGGGTTGGCGTTAGGATCCACGTTGGGTTAGCGTTTGG		
19*48	GTTTGGCATGTTGTGACCTTGACCGTTACCTTGGTTGTTCTTGTGTTGG		
20*48	AACAGCGTTGTTAGCGTTAGCGTTCTCGTCAACGTTTCTGTTTGGGTC		
21*48	GTATTGTTGATGTGCTTGTGATGTTCTTTCGTTGTTGTTGTTCTT		
22*48	AGAACATGGAGACCATTCCGGTAGAGATAGAGTTCTTGATCTTCTTCAA		
23*48	AGCAGAACCTGGCTTGATTCTAACTTGGATAACCGTTACCACAGGTAAC		
24*28	TCGACTACAATTCGTCCTTAGGCTTGTT		

Falc 3

Nco1	1-28	2-48	3-48	4-48	5-48	6-48	7-48	8-48	9-48	10-48	11-48	12-48	Sal1
	13*48	14*48	15*48	16*48	17*48	18*48	19*48	20*48	21*48	22*48	23*48	24*28	

FIG. 9



FIG. 10

↓68

83

2 LeuAspLysArgAsnSerArgSerLeuGlyGluAsnAspAspAlaAsnAsnAsnGly
 CTAGATAAAAGAACTCTAGGTCTTTGGGTGAAAACGATGACGCTAACAACAACAACGGT
 GATCTATTTCTTTGAGATCCAGAAACCCACTTTTGCTACTGCGATTGTTGTTGTTGCCA

1 XBAI,

103

62 AspAsnGlyArgGluGlyLysAspGluAspLysArgAspGlyAsnAsnGluAspAsnGlu
 GACAACGGTAGAGAAGGTAAGGACGAAGACAAGAGAGACGGTAACAACGAAGACAACGAA
 CTGTTGCCATCTCTTCCATTCCCTGCTTCTGTTCTCTCTGCCATTGTTGCTTCTGTTGCTT

123

122 LysLeuArgLysProLysHisLysLysLeuLysGlnProGlyAspGlyAsnProAspPro
 AAGTTGAGAAAGCCAAAGCACAAGAAGCTTAAGCAACCAGGTGACGGTAACCCAGACCCA
 TTCAACTCTTTCGGTTTCGTGTTCTTCGAATTCGTTGGTCCACTGCCATTGGGTCTGGGT

146 HIND3, 149 AFL2, 167 BSTE2,

289 292

182 AsnAlaAsnProAsnValAspProAsnAlaAsnProAsnAlaAsnProAsnLysAsnAsn
 AACGCTAACCCAAACGTGGATCCTAACGCCAACCTAATGCTAATCCAAACAAGAACAAC
 TTGCGATTGGGTTTGCACCTAGGATTGCGGTTGGGATTACGATTAGGTTTGTCTTGTGTTG

199 BAMHI,

312

242 GlnGlyAsnGlyGlnGlyHisAsnMetProAsnAspProAsnArgAsnValAspGluAsn
 CAAGGTAACGGTCAAGGTCACAACATGCCAAACGACCCAAACAGAAACGTTGACGAGAAC
 GTTCCATTGCCAGTTCAGTGTGTACGGTTTGTCTGGGTTTGTCTTTGCAACTGCTCTTG

332

302 AlaAsnAlaAsnAsnAlaValLysAsnAsnAsnAsnGluGluProSerAspLysHisIle
 GCTAACGCTAACAACGCTGTTAAGAACAACAACAACGAAGAACCATCTGACAAGCACATC
 CGATTGCGATTGTTGCGACAATTCTTGTGTTGTTGCTTCTTGGTAGACTGTTGCGTGTAG

352

362 GluGlnTyrLeuLysLysIleLysAsnSerIleSerThrGluTrpSerProCysSerVal
 GAACAATACTTGAAGAAGATCAAGAACTCTATCTCTACCGAATGGTCTCCATGTTCTGTT
 CTTGTTATGAACCTTCTTAGTTCTTGAGATAGAGATGGCTTACCAGAGGTACAAGACAA

372

422 ThrCysGlyAsnGlyIleGlnValArgIleLysProGlySerAlaAsnLysProLysAsp
 ACCTGTGGTAACGGTATCCAAGTTAGAATCAAGCCAGGTTCTGCTAACAAGCCTAAGGAC
 TGGACACCATGCCATAGGTTCAATCTTAGTTCGGTCCAAGACGATTGTTGCGATTCTG

456 ALWN1, 473 MST2,

392

482 GluLeuAspTyrGluAsnAspAsnGluLysLysIleCysLysMetGluLysCysSerSer
 GAATTGGACTACGAAAACGACAACGAAAAGAAGATTTGTAAGATGGAAAAGTGTCTTCC
 CTTAACCTGATGCTTTTGTGCTTTTCTTCTAAACATTCTACTTTTTACAAGAAGG

542 AM Ser
 TAGTCGAC
 ATCAGCTG

544 SALI,

602

FIG. 11

3 MetGlyGlnAsnLysIleIleGlnAlaGlnArgAsnLeuAsnGluLeuCysTyrAsnGlu
 ATGGGTCAAACAAGATTATCCAAGCCCAAAGAACTTGAACGAATTGTGTTACAACGAA
 TACCCAGTTTTGTTCTAATAGGTTTCGGGTTTCTTTGAACTTGCTTAACACAATGTTGCTT
 ~
 1 NCOI,

63 GlyAsnAspAsnLysLeuTyrHisValLeuAsnSerLysAsnGlyLysIleTyrAsnArg
 GGTAACGACAACAAGTTGTACCACGTCTTGAACCTAAGAACGGTAAGATTTACAACAGA
 CCATTGCTGTTGTTCAACATGGTGCAGAACTTGAGATTCTTGCCATTCTAAATGTTGTCT

123 AsnThrValAsnArgLeuLeuAlaAspAlaProGluGlyLysLysAsnGluLysLysAsn
 AACACTGTCAACAGATTGTTGGCCGACGCTCCAGAAGGTAAGAAGAACGAAAAGAAGAAC
 TTGTGACAGTTGTCTAACAACCGGCTGCGAGGTCTTCCATTCTTCTTGCTTTTCTTCTTG

183 GluLysIleGluArgAsnAsnLysLeuLysGlnProProProProProAsnProAsnAsp
 GAAAAGATTGAAAGAAACAACAAGTTGAAGCAACCACCACCACCCCAACCCAAACGAC
 CTTTTCTAACTTTCTTTGTTGTTCAACTTCGTTGGTGGTGGTGGGGGGTTGGGTTTGCTG

243 ProAlaProProAsnAlaAsnAspGlyGlyAsnAsnAsnAsnLysAsnAsnAsnAsp
 CCAGCTCCACCAACGCTAACGACGGTGGTAACAACAATAATAAGAACAACAATAATGAC
 GGTGAGGTGGTTTTGCGATTGCTGCCACCATTGTTGTTATTATTCTTGTGTTATTACTG

303 AspSerTyrIleProSerAlaGluLysIleLeuGluPheValLysGlnIleArgAspSer
 GACTCTTACATCCCATCTGCTGAAAAGATCTTGGAAATTCGTTAAGCAAATCAGAGACTCT
 CTGAGAATGTAGGGTAGACGACTTTTCTAGAACCCTAAGCAATTCGTTTAGTCTCTGAGA
 ~

328 BGL2, 336 ECORI,

363 IleThrGluGluTrpSerGlnCysAsnValThrCysGlySerGlyIleArgValArgLys
 ATTACCGAAGAATGGTCCCAATGTAACGTCACCTGTGGTTCTGGTATCAGAGTTAGAAAG
 TAATGGCTTCTTACCAGGGTTACATTGCAGTGGACACCAAGACCATAGTCTCAATCTTTC

423 ArgLysGlySerAsnLysLysAlaGluAspLeuThrLeuGluAspIleAspThrGluIle
 AGAAAGGGTTCTACAAGAAGGCCGAAGACTTGACTTTGGAAGACATCGACACTGAAATC
 TCTTTCCCAAGATTGTTCTTCCGGCTTCTGAACTGAAACCTTCTGTAGCTGTGACTTTAG
 ~

424 XMNI,

483 CysLysMetAspLysCysSerSerAM
 TGTAAGATGGACAAGTGTCTTCTTAGTCGAC
 ACATTCTACCTGTTCAAGAAGAATCAGCTG
 ~

509 SALI,

7432 M 0199 1990/04/10 16:42:01 PAT. 93721 U

(a/

"ANALOGOS DA PROTEINA DO CIRCUNSPOROZOITO DE PLASMODIUM
DESPROVIDO DE ALGUMAS SEQUENCIAS REPETITIVAS"

Esta invenção refere-se a análogos imunogénicos recombinantes de proteínas do circunsporozoito de Plasmodium (CS), úteis como vacinas de subunidades contra a malária, aos quais faltam um ou mais dos epitopos repetitivos da proteína CS nativa, mas que possuem, pelo menos, um epitopo flanqueador não repetitivo de cada uma das regiões flanqueadoras amino-terminal e carboxi-terminal, da proteína CS nativa.