

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

(22) Data de pedido: 2003.10.30	(73) Titular(es): IMMUNOCORE LTD.	
(30) Prioridade(s): 2002.11.09 GB 0226227	57C MILTON PARK ASBINGDON OXFORD OX14	
2003.01.25 GB 0301814	4RX	GB
2003.02.22 GB 0304067	(72) Inventor(es):	
2003.04.16 US 463046 P	BENT KARSTEN JAKOBSEN	GB
2003.05.16 GB 0311397	TORBEN BENT ANDERSEN	GB
(43) Data de publicação do pedido: 2005.08.03	PETER EAMON MOLLOY	GB
(45) Data e BPI da concessão: 2009.05.27	YI LI	GB
163/2009	JONATHAN MICHAEL BOULTER	GB
	(74) Mandatário:	
	ALBERTO HERMÍNIO MANIQUE CANELAS	
	RUA VÍCTOR CORDON, 14 1249-103 LISBOA	PT

(54) Epígrafe: **APRESENTAÇÃO DE UM RECEPTOR DAS CÉLULAS T**

(57) Resumo:

RESUMO**"APRESENTAÇÃO DE UM RECEPTOR DAS CÉLULAS T"**

Uma partícula proteica, por exemplo um bacteriófago, um ribossoma ou uma célula, que apresenta na sua superfície um receptor de células T (TCR). O TCR apresentado é, de preferência, um heterodímero tendo uma ligação dissulfureto não nativa entre resíduos do domínio constante. Tais partículas apresentadoras podem ser usadas para a criação de bibliotecas diversificadas de TCR para a identificação de TCRs de elevada afinidade. São descritas várias afinidades elevadas.

DESCRIÇÃO

"APRESENTAÇÃO DE UM RECEPTOR DAS CÉLULAS T"

O presente invento está relacionado com partículas proteicas, por exemplo partículas fágicas ou ribossomas, que apresentam receptores de células T (TCRs) e bibliotecas diversificadas dos mesmos.

Fundamento do invento

TCRs nativos

Como descrito, por exemplo, em WO99/60120 os TCRs medeiam o reconhecimento de complexos major de histocompatibilidade (MHC)-péptido pelas células T e, como tal, são essenciais ao funcionamento da vertente celular do sistema imune.

Os anticorpos e TCRs são os únicos dois tipos de moléculas que reconhecem antígenios de forma específica e, assim, o TCR é o único receptor de antígenios peptídicos particulares apresentados por MHC, o péptido estranho sendo por vezes o único sinal de uma anormalidade dentro de uma célula. O reconhecimento pelas células T ocorre quando uma célula T e uma célula apresentadora de antígeno (APC) estão em contacto físico directo e inicia-se pela ligação de TCRs específicos de antígeno a complexos pMHC.

O TCR nativo é uma proteína heterodimérica da superfície celular, da superfamília das imunoglobulinas, que está associada a proteínas invariantes do complexo CD3 envolvidas na mediação do sinal de transdução. Os TCRs existem em duas formas, $\alpha\beta$ e $\alpha\delta$, as quais são estruturalmente semelhantes mas possuem localizações anatómicas e, provavelmente, também funções muito distintas. Os ligandos de MHC classe I e de classe II são igualmente proteínas da superfamília das imunoglobulinas mas são especializados na apresentação de antígeno, com um local altamente polimórfico de ligação ao péptido, o que permite apresentar um arranjo diverso de pequenos fragmentos peptídicos na superfície da célula APC.

Duas outras classes de proteínas são conhecidas como sendo capazes de funcionar como ligandos de TCR. (1) Os antígenos CD1 são moléculas relacionadas com MHC classe I, cujos genes estão localizados num cromossoma diferente dos antígenos MHC classe I e classe II clássicos. As moléculas CD1 são capazes de apresentar grupos peptídicos e não peptídicos (e.g., lípidos, glicolípidos) a células T, numa forma análoga aos complexos convencionais de MHC classe I e classe II-pep. Ver por exemplo (Barclay et al., (1997) "The Leucocyte Antigen Factsbook" 2nd Edition, Academic Press) e (Bauer (1997) Eur J Immunol 27 (6) 1366-1373)); (2) Os superantígenos bacterianos são toxinas solúveis que são capazes de se ligar a moléculas MHC classe II e a uma sub-série de TCRs. (Fraser (1989) Nature 339:

221-223). Muitos superantigénios apresentam especificidade para um ou dois segmentos Vbeta, enquanto que outros apresentam ligação mais promíscua. Em qualquer dos casos, os superantigénios são capazes de induzir um aumento da resposta imune devido à sua capacidade para estimular sub-séries de células T de forma policlonal.

A porção extracelular dos TCRs heterodiméricos nativos $\alpha\beta$ e $\alpha\delta$ consiste em dois polipéptidos, cada um dos quais possui um domínio constante próximo da membrana e um domínio variável afastado da membrana. Cada um dos domínios constante e variável inclui uma ligação dissulfureto intracadeia. Os domínios variáveis possuem ansas altamente polimórficas análogas a regiões determinantes de complementaridade (CDRs) dos anticorpos. CDR3 dos TCRs $\alpha\beta$ interagem com o péptido apresentado por MHC e as CDRs 1 e 2 dos TCRs $\alpha\beta$ interagem com o péptido e com o MHC. A diversidade das sequências de TCR é gerada através do rearranjo somático dos genes ligados das regiões variável (V), de diversidade (D), de ligação (J) e constante.

Os polipéptidos das cadeias α e γ funcionais são formados por rearranjo das regiões V-J-C, enquanto que as cadeias β e δ consistem em regiões V-D-J-C. O domínio constante extracelular possui uma região próxima da membrana e uma região de imunoglobulina. Existem domínios constantes simples das cadeias α e δ , conhecidos como TRAC e TRDC respectivamente. O domínio constante da cadeia β é composto por um de dois domínios constantes β diferentes,

conhecidos como TRBC1 e TRBC2 (nomenclatura IMGT). Existem quatro diferenças entre os aminoácidos destes domínios constantes β , três das quais estão dentro dos domínios usados para produzir os TCRs de cadeia simples apresentados nas partículas fágicas do presente invento. Estas diferenças estão todas dentro do exão 1 de TRBC1 e TRBC2: $N_4K_5 \rightarrow K_4N_5$ e $F_{37} \rightarrow Y$ (numeração IMTG, diferenças TRBC1 $\rightarrow$ TRBC2), a última alteração de aminoácidos entre as duas regiões constantes da cadeia β de TCR sendo no exão 3 de TRBC1 e TRBC2: $V_1 \rightarrow E$. O domínio constante γ é composto por um de entre TRGC1, TRGC2(2x) ou TRGC2(3x). Os dois domínios constantes TRGC2 diferem apenas no número de cópias dos aminoácidos codificados pelo exão 2 deste gene que estão presentes.

A extensão de cada um dos domínios extracelulares de TCR é algo variável. No entanto, os familiarizados com a área podem facilmente determinar a posição das fronteiras do domínio usando uma referência como seja "The T Cell Receptor Facts Book", Lefranc, Publ. Academic Press 2001.

TCRs recombinantes

A produção de TCRs recombinantes é benéfica, uma vez que proporciona análogos solúveis de TCR adequados para os seguintes fins:

- estudo das interações TCR/ligando (e.g., pMHC para TCRs $\alpha\beta$)

- rastreio de inibidores das interacções associadas a TCR
- proporcionar a base para potenciais terapias

Até agora, foi projectada uma série de construções para a produção de TCRs recombinantes. Estas construções classificam-se em duas classes, TCRs de cadeia simples e TCRs diméricos, a literatura relevante destas construções está resumida abaixo. As TCRs de cadeia simples (scTCR) são construções artificiais consistindo numa única cadeia de aminoácidos, a qual à semelhança de TCRs heterodiméricos nativos se liga a complexos de MHC-péptido. Infelizmente, foram infrutíferas as tentativas para produzir análogos scTCRs alfa/beta de cadeia simples ligando simplesmente as cadeias alfa e beta, de forma que ambas fossem expressas numa única grelha de leitura aberta, presumivelmente devido à instabilidade natural do emparelhamento dos domínios solúveis alfa-beta.

Assim, técnicas especiais usando várias trunca-gens de uma ou de ambas as cadeias alfa e beta têm sido necessárias para a produção de scTCRs. Estes formatos parecem ser aplicáveis apenas a uma gama muito limitada de sequências de scTCR. Soo Hoo *et al.* (1992) PNAS, 89 (10):4759-63 descrevem a expressão de um TCR de murganho, no formato de cadeia simples, a partir do clone 2C de células T usando uma cadeia beta e uma cadeia alfa truncada ligadas através de um elemento de ligação de 25 aminoácidos e expressão no periplasma bacteriano (ver também Schodin *et al.* (1996) Mol. Immunol. 33 (9): 819-29). Esta estratégia

também constitui a base do TCR m6 de cadeia simples descrito por Holler *et al.* (2000) PNAS 97 (10):5387-92 que deriva do scTCR 2C e se liga ao mesmo aloepitopo restringido por H2-Ld. Shusta *et al.* (2000) Nature Biotechnology 18:754-759 e US 6423538 descrevem a utilização de uma construção de TCR de cadeia simples 2C murina em experiências de apresentação em levedura, as quais produziram TCRs mutados com maior estabilidade térmica e solubilidade. Esta descrição também demonstrou a capacidade destes TCRs 2C para selectivamente se ligarem a células que expressam o seu pMHC cognato. Khandekar *et al* (1997) J. Biol. Chem. 272 (51):32190-7 descrevem uma estratégia semelhante para o TCR D10 murino, ainda que este scTCR seja fundido com MBP e expresso no citoplasma bacteriano (ver também Hare *et al.* (1999) Nat. Struct. Biol. 6(6):574-81). Hilyard *et al.* (1994) PNAS 91(19):9057-61 descrevem um scTCR humano específico para a proteína da matriz da gripe-HLA-A2, usando uma construção V α -elemento de ligação-V β e expressão no periplasma bacteriano.

Chung *et al.* (1994) PNAS 91(26) 12654-8 descrevem a produção de scTCR humano usando uma construção V α -elemento de ligação-V β -C β e expressão na superfície de uma linha celular de mamífero. Esta descrição não inclui qualquer referência à ligação específica péptido-HLA da scTCR. Plaksin *et al* (1997) J. Immunol. 158 (5):2218-27 descrevem uma construção V α -elemento de ligação-V β -C β semelhante para a produção de um scTCR murino específico de um epitopo gp120 de HIV-H-2D^d. Este scTCR foi expresso como corpos de inclusão bacterianos e renaturado *in vitro*.

Uma série de publicações descrevem a produção de heterodímeros de TCR que incluem a ligação dissulfureto nativa que liga as respectivas subunidades (Garboczi, *et al.*, (1996), *Nature* 384(6605): 134-41; Garboczi, *et al.*, (1996), *J Immunol* 157(12):5403-10; Chang *et al.*, (1994), *PNAS USA* 91: 11408-11412; Davodeau *et al.*, (1993), *J. Biol. Chem.* 268(21):15455-15460; Golden *et al.*, (1997), *J. Imm. Meth.* 206:163-169; Patente uS N° 6080840). No entanto, ainda que tais TCRs possam ser reconhecidos por anticorpos específicos de TCR, nenhum reconheceu o seu ligando nativo sem ser em concentrações relativamente elevadas e/ou não eram estáveis.

Em WO 99/60120, foi descrito um TCR solúvel que atinge a sua conformação correcta de forma a ser capaz de reconhecer o seu ligando nativo, é estável durante um longo período de tempo e pode ser produzido em quantidades razoáveis. Este TCR compreende um domínio extracelular da cadeia α ou δ de TCR dimerizado com um domínio extracelular de uma cadeia β ou δ respectivamente, por meio de um par de péptidos de dimerização C-terminais, tais como fechos de leucina. Esta estratégia de produção de TCRs é, de um modo geral, aplicável a todos os TCRs.

Reiter *et al.*, *Immunity*, 1995, 2:281-287, descreve detalhadamente a construção de uma molécula solúvel compreendendo domínios variáveis de TCR α e β estabilizados com ligações dissulfureto, um dos quais está

ligado a uma forma truncada da exotoxina de *Pseudomonas* (PE38). Uma das razões estabelecidas para a produção desta molécula foi ultrapassar a instabilidade inerente dos TCRs de cadeia simples. A posição da nova ligação dissulfureto nos domínios variáveis de TCR foi identificada via homologia com os domínios variáveis de anticorpos, nos quais esta foi previamente introduzida (por exemplo ver Brinkmann, *et al.* (1993), *Proc. Natl. Acad. Sci. USA* 90:7538-7542 e Reiter, *et al.* (1994) *Biochemistry* 33:5451-5459). No entanto, uma vez que não existe tal homologia entre os domínios constantes de anticorpos e de TCR, tal técnica poderá não ser empregue para identificar os locais adequados para novas ligações inter-cadeias entre os domínios constantes de TCR.

Conforme referido atrás Shusta *et al.* (2000) *Nature Biotechnology* 18:754-759 descrevem a utilização de construções do TCR 2C de cadeia simples em experiências de apresentação em levedura. O princípio de apresentação de scTCRs em partículas fágicas foi anteriormente discutido. Por exemplo, WO/19129 descreve detalhadamente a produção de scTCRs e sumariza um potencial método para a produção de partículas fágicas que apresentam scTCRs do formato V α -elemento de ligação-V β C β . No entanto, este pedido não possui exemplos que demonstrem a produção das referidas partículas fágicas apresentadoras de TCR. O pedido, no entanto, refere-se a um pedido de patente copendente: "A construção de vectores de DNA incluindo um segmento de DNA codificador de uma molécula de sc-TCR fundida com uma

proteína externa do bacteriófago (gene II ou gene VIII) foi descrita no referido pedido de patente U.S. copendente N° 08/813781."

Ainda, este pedido de patente baseia-se na capacidade dos anticorpos anti-TCR ou complexos de superantígeno e MHC para reconhecerem scTCRs solúveis, não apresentados por fagos, produzidos para verificar a sua conformação correcta. Assim, a verdadeira especificidade da ligação a péptido-MHC dos scTCRs, em qualquer formato, não foi conclusivamente demonstrada. Finalmente, um outro estudo (Onda et al., (1995) Molecular Immunology 32 (17-18) 1387-1397) descreve a apresentação por fagos de duas cadeias α de TCR murino na ausência das suas respectivas cadeias β . Este estudo demonstrou que as partículas fágicas apresentadoras de uma das cadeias α de TCR (derivada do hibridoma murino A1.1) ligaram-se preferencialmente aos mesmos péptidos imobilizados em alvéolos de microtitulação a que o TCR completo normalmente responderia quando apresentados por MHC I classe I-A^d murino.

Utilização do rastreio

Várias interacções celulares e respostas celulares importantes, incluindo a sinapse imune mediada por TCR, são controladas por contactos feitos entre os receptores da superfície celular e ligandos apresentados nas superfícies de outras células. Estes tipos de contactos moleculares específicos são de importância crucial para a

regulação bioquímica correcta do corpo humano e estão, consequentemente, a ser intensamente estudados. Em muitos casos, o objectivo de tais estudos é procurar um meio de modular as respostas celulares de forma a prevenir ou combater doenças.

Assim, métodos com os quais identificar compostos que se liguem, com algum grau de especificidade, a moléculas de receptores humanos ou a moléculas de ligandos são importantes como novas moléculas promissoras para a descoberta e desenvolvimento de novos fármacos. Em particular, os compostos que interferem com determinadas interacções de receptor-ligando possuem potencial imediato como fármacos ou veículos.

Os avanços na química combinatória, permitindo a produção, relativamente fácil e eficiente em termos de custos, de bibliotecas muito grandes de compostos aumentaram enormemente o âmbito da testagem de compostos. Agora, as limitações dos programas de rastreio geralmente residem na natureza dos ensaios que podem ser empregues, na produção de moléculas de receptor e de ligando e em como todos estes ensaios podem ser adaptados a métodos de rastreio em grande escala.

Métodos de apresentação

É muitas vezes desejável apresentar um determinado péptido ou polipéptido na superfície de uma

partícula proteica. Tais partículas podem servir como auxiliares de purificação para o péptido ou polipéptido (uma vez que as partículas portadoras do péptido ou polipéptido podem ser separadas dos contaminantes indesejáveis, através da sedimentação ou de outros métodos). Podem também servir como vacinas de partículas, a resposta imune ao péptido ou polipéptido apresentado na superfície sendo estimulada pela apresentação de partículas. A proteína p24 do retrotransposão de levedura e a proteína de superfície da hepatite B são exemplos de proteínas que se auto-associam em partículas. A fusão do péptido ou polipéptido com interesse com estas proteínas formadoras de partículas é uma forma reconhecida de apresentação do péptido ou polipéptido na superfície das partículas resultantes.

No entanto, os métodos de apresentação de partículas têm sido usados principalmente para identificar proteínas com propriedades desejáveis, como seja rendimentos de expressão elevados, características de ligação e/ou estabilidade. Estes métodos envolvem a criação de um conjunto diversificado ou "biblioteca" de proteínas ou polipéptidos expressos na superfície de partículas proteicas. Estas partículas possuem duas características-chaves, em primeiro lugar cada uma das partículas apresenta uma proteína ou polipéptido variante simples e em segundo lugar o material genético codificador da proteína ou polipéptido expresso está associado ao da partícula. Esta biblioteca é então sujeita a um ou mais ciclos de selecção.

Por exemplo, estes podem consistir no contacto de um ligando com uma biblioteca de partículas apresentadoras de receptores mutados e identificação de quais os receptores mutados se ligam com maior afinidade ao ligando. Uma vez completado o processo de selecção, o receptor ou receptores com as propriedades pretendidas pode ser isolado e o seu material genético pode ser amplificado de forma a permitir que os receptores sejam sequenciados.

Estes métodos de apresentação classificam-se em duas grandes categorias, apresentação *in vitro* e *in vivo*.

Todos os métodos de apresentação *in vivo* se baseiam num passo em que a biblioteca, geralmente codificada no ou com o ácido nucleico genético de uma partícula replicável, como seja um replicão plasmídico ou fágico, é usada para transformar células para permitir a expressão das proteínas ou polipéptidos. (Pluckthun (2001) Adv Protein Chem 55:367-403). Existe uma série de sistemas de replicação/hospedeiro que se demonstrou serem adequados para a apresentação *in vivo* da proteína ou polipéptidos. Estes incluem os seguintes:

Fago/células bacterianas

Plasmídeo/células CHO

Vectores baseados no plasmídeo 2 μ m de levedura/células de levedura

Baculovírus/células de insecto

Plasmídeo/células bacterianas

Os métodos de apresentação *in vivo* incluem os métodos de apresentação na superfície celular, em que um plasmídeo é introduzido na célula hospedeira codificadora de uma proteína de fusão consistindo na proteína ou polipéptido com interesse fundido com uma proteína ou polipéptido da superfície celular. A expressão desta proteína de fusão conduz à apresentação da proteína ou polipéptido com interesse na superfície da célula. As células que apresentam estas proteínas ou polipéptidos com interesse podem então ser sujeitas a um processo de selecção, como seja FACS, e os plasmídeos obtidos a partir da célula ou células seleccionadas podem ser isolados e sequenciados. Os sistemas de apresentação na superfície celular foram projectados para células de mamífero (Higuchi (1997) J Immunol. Methods 202: 193-204), células de levedura (Shusta (1999) J Mol Biol 292: 949-956) e células bacterianas (Sameulson (2002) J. Biotechnol 96 (2) 129-154).

Foram publicadas numerosas revisões das várias técnicas de apresentação *in vivo*. Por exemplo, (Hudson (2002) Expert Opin Biol Ther (2001) 1(5) 845-55) e (Schmitz (2000) (Supp A) S106-S112).

Os métodos de apresentação *in vitro* são baseados na utilização de ribossomas para traduzir bibliotecas de mRNA num arranjo diversificado de variantes de proteínas ou polipéptidos. A ligação entre as proteínas ou polipéptidos

formados e o mRNA codificador destas moléculas é mantido por um de dois métodos. A apresentação por ribossomas convencional utiliza mRNA, sequências que codificam uma curta sequência de um elemento de ligação (tipicamente 40-100 aminoácidos) e a proteína ou polipéptido a ser apresentado. A sequência do elemento de ligação confere espaço à proteína ou polipéptido apresentado para renaturar sem impedimento espacial pelo ribossoma. A sequência de mRNA não possui um codão de paragem, isto assegura que a proteína ou polipéptido expresso e o RNA se mantêm ligados à partícula de ribossoma. O método de apresentação de mRNA relacionado baseia-se na preparação de sequências de mRNA codificadoras da proteína ou polipéptido com interesse e em elementos de ligação de DNA portadores de um grupo de puromicina. Logo que os ribossomas atinjam a junção mRNA/DNA, a tradução é parada e a puromicina forma uma ligação covalente com o ribossoma. Para uma revisão recente destes dois métodos relacionados de apresentação *in vitro* ver (Amstutz (2001) Curr Opin Biotechnol 12: 400-405).

É particularmente preferida a técnica de apresentação fágica que se baseia na capacidade das partículas bacteriofágicas para expressarem um péptido ou polipéptido heterólogo fundido com as suas proteínas de superfície. (Smith (1985) Science 217:1315-1317). O procedimento está muito generalizado e bem compreendido na área da apresentação de monómeros de polipéptidos. No entanto, no caso de polipéptidos que na sua forma nativa se associam como dímeros, apenas a apresentação de anticorpos em fagos parece ter sido bem investigada.

Para a apresentação de polipéptidos monoméricos existem dois procedimentos principais:

Em primeiro lugar (Método A), através da inserção no vector (fagemídeo) de um DNA codificador do péptido ou polipéptido heterólogo fundido com o DNA codificador de uma proteína externa do bacteriófago. A expressão das partículas fágicas apresentadoras do péptido ou polipéptido heterólogo é então realizada através da transfecção de células bacterianas com o fagemídeo e depois infecção das células transformadas com um "fago auxiliar". O fago auxiliar actua como fonte das proteínas fágicas não codificadas pelo fagemídeo necessárias para produzir uma partícula fágica funcional.

Em segundo lugar (Método B), através da inserção de DNA codificador do péptido ou polipéptido heterólogo num genoma fágico completo fundido com o DNA codificador de uma proteína externa do bacteriófago. A expressão das partículas fágicas apresentadoras do péptido ou polipéptido heterólogo é então realizada através da infecção de células bacterianas com o genoma fágico. Este método tem a vantagem, relativamente ao primeiro método, de ser um processo de um único passo. No entanto, o tamanho da sequência de DNA heteróloga que pode ser encapsidada com êxito nas partículas fágicas resultantes é reduzida. M13, T7 e lambda são exemplos de fagos adequados para este método.

Uma variação do Método B envolve a adição de uma sequência de DNA codificadora de um domínio de ligação a nucleótidos ao DNA no genoma fágico codificador do péptido heterólogo a ser apresentado e, ainda, adição do correspondente local de ligação nucleotídico ao genoma fágico. Isto faz com que o péptido heterólogo fique directamente ligado ao genoma fágico. Este complexo péptido/genoma é então encapsidado numa partícula fágica que apresenta o péptido heterólogo. Este método está totalmente descrito em WO 99/11785.

As partículas fágicas podem então ser recuperadas e usadas para estudar as características de ligação do péptido ou polipéptido heterólogo. Uma vez isolado, o DNA do fagemídeo ou do fago pode ser recuperado a partir da partícula fágica apresentadora do péptido ou polipéptido e este DNA pode ser replicado através de PCR. O produto de PCR pode ser usado para sequenciar o péptido ou polipéptido heterólogo apresentado por uma determinada partícula fágica.

A apresentação fágica de anticorpos de cadeia simples, e seus fragmentos, tornou-se um método de rotina do estudo das características de ligação destes polipéptidos. Existem disponíveis numerosos livros que revêem as técnicas de apresentação fágica e a biologia dos bacteriófagos. (Ver, por exemplo, "Phage Display - A Laboratory Manual", Barbas *et al.*, (2001) Cold Spring Harbour Laboratory Press).

Um terceiro método de apresentação fágica (Método C) baseia-se no facto de os polipéptidos heterólogos possuindo um resíduo de cisteína numa localização pretendida poderem ser expressos, numa forma solúvel, por um fagemído ou genoma fágico e levados a associarem-se a uma proteína da superfície do fago modificada também tendo um resíduo de cisteína numa posição exposta na superfície, através da formação de uma ligação dissulfureto entre as duas cisteínas. WO 01/05950 descreve detalhadamente a utilização deste método de ligação alternativo para a expressão de péptidos derivados de anticorpos de cadeia simples.

Breve descrição do invento

Os TCRs nativos são heterodímeros que possuem domínios transmembranares longos, os quais são essenciais para manter a sua estabilidade como dímeros funcionais. Como discutido atrás, os TCRs são úteis nas suas formas solúveis para fins de pesquisa e terapêuticos, de forma que a apresentação da sua forma nativa insolúvel tem pouca utilidade. Por outro lado, as formas solúveis de TCRs são difíceis de projectar e, uma vez que a maioria dos métodos parece ter sido descrita apenas para péptidos e polipéptidos monoméricos, os métodos de apresentação adequados para os TCRs diméricos solúveis não foram investigados. Ainda, uma vez que a funcionalidade do TCR apresentado depende da associação correcta dos domínios

variáveis do dímero TCR, a apresentação com êxito de um TCR dimérico funcional não é fácil.

WO99/18129 contem a afirmação: "As construções de DNA codificadoras das proteínas de fusão sc-TCR podem ser usadas para preparar uma biblioteca de apresentação fágica de acordo com os métodos descritos no pedido de patente US pendente N° de Série 08/813781 entregue em 7 de Março, 1997, cuja descrição é aqui incluída como referência." mas de facto nenhuma descrição de tal apresentação está incluída neste pedido. Os inventores deste pedido de patente publicaram um artigo (Weidanz (1998) J Immunol Methods 221 59-76) que demonstra a apresentação de dois scTCRs murinos em partículas fágicas.

WO 01/62908 descreve métodos para a apresentação fágica de scTCRs e proteínas de fusão de scTCRs/Ig. No entanto, a funcionalidade (ligação específica a pMHC) das construções descritas não foi avaliada.

Finalmente, um método mediado por retrovírus para a apresentação de bibliotecas diversificadas de TCR na superfície de células T imaturas foi demonstrado para um TCR murino. A biblioteca de TCRs mutados apresentados na superfície das células T imaturas foi testada por citometria de fluxo usando tetrâmeros de pMHC e isto conduziu à identificação de variantes de TCR que eram específicas para o pMHC cognato ou para uma sua variante. (Helmut et al., (2000) PNAS 97(26) 14578-14583).

Este invento baseia-se, em parte, no facto de TCRs de cadeia simples e diméricos poderem ser expressos como fusões de superfície em partículas fágicas e disponibiliza partículas fágicas apresentadoras de construções scTCR e dTCR de análogos de alfa/beta e de análogos gama/delta. Tais TCRs apresentados por partículas fágicas são úteis para fins de purificação e rastreio, particularmente como uma biblioteca diversificada de TCRs apresentados por fagos para selecção por imobilização ("biopanning") no sentido de identificar TCRs com características desejáveis, tais como elevada afinidade para o complexo MHC-péptido alvo. Ainda relacionado, os scTCRs apresentados por fagos podem ser úteis para a identificação dos TCRs pretendidos, mas essa informação pode ser melhor aplicada na construção de TCRs diméricos análogos para utilização terapêutica. O invento também inclui TCRs de elevada afinidade identificáveis por estes métodos.

Descrição detalhada do invento

Num aspecto mais lato, o presente invento proporciona partículas fágicas apresentando na sua superfície um polipéptido TCR de cadeia simples (scTCR) ou o par de polipéptidos de TCR dimérico (dTCR), o referido polipéptido scTCR sendo constituído por

um primeiro segmento constituído por uma sequência do

domínio variável da cadeia α ou δ de TCR fundida com o extremo N de uma sequência extracelular do domínio constante da cadeia α de TCR,

um segundo segmento constituído por um domínio variável da cadeia β ou γ de TCR fundido com o extremo N de uma sequência extracelular do domínio constante da cadeia β de TCR,

uma sequência de um elemento de ligação que liga o extremo C do primeiro segmento ao extremo N do segundo segmento, ou vice versa, e

uma ligação dissulfureto entre o primeiro e segundo segmentos, a referida ponte dissulfureto sendo uma que não tem equivalente nos receptores de células T $\alpha\beta$ ou $\gamma\delta$ nativos;

o referido par de polipéptidos dTCR sendo constituído por

um primeiro polipéptido em que uma sequência do domínio variável da cadeia α ou δ de TCR está fundido com o extremo N de uma sequência extracelular do domínio constante da cadeia α de TCR e

um segundo polipéptido em que uma sequência do domínio variável da cadeia β ou γ de TCR está fundido com o extremo N de uma sequência correspondendo a uma sequência extracelular do domínio constante da cadeia β de TCR,

o primeiro e o segundo polipéptidos, estando ligados através de uma ligação dissulfureto que não tem equivalente nos receptores de células T $\alpha\beta$ ou $\gamma\delta$ nativos,

a referida ligação dissulfureto entre o primeiro e o segundo segmento do scTCR ou entre o primeiro e o segundo polipéptidos do dTCR, sendo entre cisteínas que substituem os seguintes pares de resíduos do exão 1 de TRAC*01 para a cadeia TCR α e TRBC 1*01 ou TRBC2*01 para a cadeia β de TCR:

Cadeia α de TCR	Cadeia β de TCR
Thr 48	Ser 57
Thr 45	Ser 77
Tyr 10	Ser 17
Thr 45	Asp 59
Ser 15	Glu 15

As sequências do domínio variável do par de polipéptidos dTCR ou do polipéptido scTCR estão mutuamente orientadas, substancialmente como nos TCRs nativos, e a ligação dissulfureto entre o primeiro e o segundo segmento do scTCR ou entre o primeiro e o segundo polipéptido do dTCR não tem equivalente nos receptores de células T nativas.

No caso dos scTCRs ou dTCRs α e β apresentados de acordo com o invento, o facto das sequências do domínio variável dos segmentos ou polipéptidos α ou β estarem

mutuamente orientados como nos receptores de células T $\alpha\beta$ nativos implica que os TCRs são funcionais e isto pode ser testado através da confirmação de que a molécula se liga ao ligando de TCR relevante (complexo pMHC, complexo CD1-antigénio, complexo superantigénio/pMHC) - caso se ligue, então está preenchido o requisito. As interacções com os complexos pMHC podem ser medidas usando um instrumento Biacore 3000™ ou Biacore 2000™. WO99/6120 proporciona descrições detalhadas dos métodos necessários para analisar a ligação de TCR aos complexos MHC-péptido. Estes métodos são igualmente aplicáveis ao estudo das interacções de TCR/CD1 e TCR/superantigénio. Para aplicar estes métodos ao estudo das interacções TCR/CD1, são necessárias formas solúveis de CD, a produção das quais está descrita em (Bauer (1997) Eur J Immunol 27 (6) 1366-1373). No caso dos TCRs $\gamma\delta$ do presente invento, os ligandos cognatos para estas moléculas são desconhecidos, assim podem ser empregues meios secundários de verificação da sua conformação, como seja reconhecimento por anticorpos. O anticorpo monoclonal MCCA991T (disponível na Serotec), específico para a região variável da cadeia δ , é um exemplo de um anticorpo adequado para este objectivo.

Os scTCRs ou dTCRs do presente invento podem ser apresentados em partículas fágicas, de preferência partículas fágicas filamentosas através, por exemplo, dos seguintes meios:

i) o extremo C de um membro do par de polipéptidos dTCR ou

o extremo C do polipéptido scTCR ou o extremo C de um elemento de ligação peptídico curto ligado ao extremo C de um deles, pode ser directamente ligado através de uma ligação peptídica a um resíduo exposto na superfície da partícula fágica. Por exemplo, o referido resíduo exposto na superfície está, de preferência, no extremo N do produto do gene II ou do gene VIII do bacteriófago; e

ii) o extremo C de um membro do par de polipéptidos dTCR, ou o extremo C do polipéptido scTCR, ou o extremo C de um curto elemento de ligação peptídico ligado ao extremo C de um deles, está ligado por uma ligação dissulfureto a um resíduo de cisteína exposto na partícula fágica através de um resíduo de cisteína introduzido. Por exemplo, o referido resíduo exposto na superfície está novamente, de preferência, no extremo N do produto do gene do gene III ou do gene VIII do bacteriófago.

Prefere-se o método (i) atrás descrito. No caso de um scTCR, o ácido nucleico codificador do TCR pode ser fundido com o ácido nucleico codificador do fago. No caso de um dTCR, o ácido nucleico codificador de uma cadeia do TCR pode ser fundido com o ácido nucleico codificador do fago e a segunda cadeia do par de polipéptidos TCR pode ser deixada a associar-se com a partícula expressa resultante apresentadora da primeira cadeia. A associação bifuncional adequada das duas cadeias é ajudada pela presença das cisteínas introduzidas no domínio constante das duas cadeias que são capazes de formar uma ligação dissulfureto entre cadeias, como mais detalhadamente discutido à frente.

Par de polipéptidos dTCR e polipéptido scTCR

As sequências extracelulares do domínio constante presentes nos scTCRs ou nos dTCRs, de preferência, correspondem às de um TCR humano, tal como as sequências do domínio variável. No entanto, a correspondência entre tais sequências não necessita de ser 1:1 ao nível dos aminoácidos. A truncagem N ou C e/ou a deleção e/ou substituição de aminoácidos relativamente às correspondentes sequências de TCR humanas é aceitável. Em particular, devido às sequências extracelulares do domínio constante presentes no primeiro e segundo segmentos não estarem directamente envolvidas nos contactos com o ligando a que scTCR ou dTCR se liga, podem ser mais curtos, podem conter substituições ou deleções relativamente às sequências extracelulares do domínio constante dos TCRs nativos.

A sequência extracelular do domínio constante presente num polipéptido do par dTCR, ou no primeiro segmento de um polipéptido scTCR, pode incluir uma sequência correspondente ao domínio constante extracelular Ig de uma cadeia TCR α e/ou a sequência extracelular do domínio constante presente num outro membro do par ou no segundo segmento pode incluir uma sequência correspondente ao domínio constante extracelular Ig de uma cadeia TCR β .

Numa realização do invento, um membro do par de polipéptidos dTCR, ou o primeiro segmento do polipéptido

scTCR, corresponde a substancialmente todo o domínio variável de uma cadeia TCR α fundido com o extremo N de substancialmente todo o domínio extracelular do domínio constante de uma cadeia TCR α ; e/ou o outro membro do par ou o segundo segmento corresponde a substancialmente todo o domínio variável de uma cadeia TCR β fundido com o extremo N de substancialmente todo o domínio extracelular do domínio constante de uma cadeia TCR β .

Numa outra realização, as sequências extracelulares do domínio constante presentes no par de polipéptidos dTCR ou no primeiro e no segundo segmento do polipéptido scTCR, correspondem aos domínios constantes das cadeias α e β de um TCR nativo truncado no seu extremo C, de forma que os resíduos de cisteína que formam a ligação dissulfureto nativa entre cadeias do TCR são excluídos. Como alternativa, os resíduos de cisteína podem ser substituídos por um outro resíduo de aminoácido, como seja serina ou alanina, de forma que a ligação dissulfureto nativa seja eliminada. Ainda, a cadeia de TCR β nativa possui um resíduo de cisteína desemparelhado e esse resíduo pode ser eliminado ou substituído por um resíduo que não seja cisteína na sequência β do scTCR do invento.

Numa realização particular do invento, as sequências do domínio variável das cadeias TCR α e β presentes no par de polipéptidos dTCR, ou no primeiro e no segundo segmentos do polipéptido scTCR, podem em conjunto corresponder ao domínio variável funcional de um primeiro TCR e

as sequências extracelulares do domínio constante das cadeias TCR α e β presentes no par de polipéptidos dTCR ou no primeiro e segundo segmentos do polipéptido scTCR podem corresponder aos de um segundo TCR, o primeiro e segundo TCRs sendo da mesma espécie. Assim, as sequências do domínio variável das cadeias α e β presentes no par de polipéptidos dTCR ou no primeiro e no segundo segmento do polipéptido scTCR, podem corresponder aos de um primeiro TCR humano e as sequências extracelulares do domínio constante das cadeias α e β podem corresponder aos de um segundo TCR humano. Por exemplo, as sequências extracelulares do domínio constante de scTCR A6 Tax podem ser usadas como estrutura à qual podem ser fundidos domínios variáveis α e β heterólogos.

Numa outra realização do invento, as sequências do domínio variável das cadeias TCR δ e γ presentes no par de polipéptidos dTCR ou no primeiro e no segundo segmento do polipéptido scTCR, respectivamente, podem em conjunto corresponder a um domínio variável funcional de um primeiro TCR, e as sequências extracelulares do domínio constante das cadeias TCR α e β presentes no polipéptido dTCR ou no primeiro e no segundo segmento do polipéptido scTCR respectivamente, podem corresponder aos de um segundo TCR, o primeiro e segundo TCRs sendo derivados da mesma espécie: Assim, as sequências do domínio variável das cadeias α e β presentes no par de polipéptidos dTCR ou no primeiro e no segundo segmentos do polipéptido scTCR podem corresponder aos de um primeiro TCR humano e as sequências extrace-

lulares do domínio constante das cadeias α e β podem corresponder aos de um segundo TCR humano. Por exemplo, as sequências extracelulares do domínio constante de scTCR A6 Tax podem ser usadas como uma estrutura à qual podem ser fundidos domínios variáveis α e β heterólogos.

Numa realização particular do invento, as sequências do domínio variável das cadeias α e β ou δ e γ de TCR presentes no par de polipéptidos dTCR ou no primeiro e no segundo segmento do polipéptido scTCR podem em conjunto corresponder ao domínio variável funcional de um primeiro TCR humano e as sequências extracelulares do domínio constante das cadeias α e β de TCR presentes no par de polipéptidos dTCR ou primeiro e segundo segmentos do polipéptido scTCR podem corresponder às do segundo TCR não humano. Assim, as sequências do domínio variável das cadeias α e β ou δ e γ presentes no par de polipéptidos dTCR ou nos primeiro e segundo segmentos do polipéptido scTCR podem corresponder aos de um primeiro TCR humano e as sequências extracelulares do domínio constante das cadeias α e β podem corresponder às do segundo TCR não humano. Por exemplo, as sequências extracelulares do domínio constante de TCR murino podem ser usadas como uma estrutura à qual podem ser fundidos domínios variáveis de TCR α e β humano heterólogo.

Elemento de ligação no polipéptido scTCR

Para as partículas fágicas apresentadoras de

scTCR do presente invento, uma sequência do elemento de ligação liga o primeiro e o segundo segmento de TCR, para formar um polipéptido de cadeia simples. A sequência do elemento de ligação pode, por exemplo, ter a fórmula -P-AA-P em que P é prolina e AA representa uma sequência de aminoácidos em que os aminoácidos são glicina e serina.

Para scTCR, apresentado pelas partículas fágicas do presente invento, se ligar a um ligando, o complexo MHC-péptido no caso dos TCRs $\alpha\beta$, o primeiro e segundo segmentos são emparelhados de forma que as sequências dos seus domínios variáveis fiquem orientadas para tal ligação. Assim, o elemento de ligação deverá ter tamanho suficiente para abranger a distância entre o extremo C do primeiro segmento e o extremo N do segundo segmento ou *vice versa*. Por outro lado, o comprimento excessivo do elemento de ligação deverá ser preferencialmente evitado, no caso do extremo do elemento de ligação presente no extremo N da sequência do domínio variável bloquear ou reduzir a ligação do scTCR ao ligando alvo.

Por exemplo, no caso das sequências extracelulares do domínio constante presentes no primeiro e segundo segmentos corresponderem aos domínios constantes das cadeias α e β de um TCR nativo truncado no extremo C, de forma que os resíduos que constituem a ligação dissulfureto nativa entre cadeias sejam excluídos e a sequência do elemento de ligação ligue o extremo C do primeiro segmento ao extremo N do segundo segmento, a sequência do elemento

de ligação pode consistir em, por exemplo, 26 a 41 aminoácidos, de preferência 29, 30, 31 ou 32 aminoácidos ou 33, 34, 35 ou 36 aminoácidos. Elementos de ligação particulares possuem a fórmula $-PGGG-(SGGGG)_5-P-$ e $-P-GGG-(S-GGGG)_6-P-$ em que P é prolina, G é glicina e S é serina.

Ligação dissulfureto inter-cadeias

Um princípio que caracteriza os dTCRs e scTCRs apresentados pelas partículas fágicas do presente invento, é a ligação dissulfureto introduzida especificamente entre as sequências extracelulares do domínio constante do par de polipéptidos dTCR ou entre o primeiro e segundo segmentos do polipéptido scTCR. Tal ligação não tem contraparte nos TCRs nativos, sendo entre cisteínas especificamente incorporadas nas sequências extracelulares do domínio constante do par de polipéptidos dTCR ou do primeiro e segundo segmentos do polipéptido scTCR. Nalguns casos, tanto uma ligação nativa como uma ligação dissulfureto não nativa podem ser desejáveis.

A ligação dissulfureto pode ser formada através da mutação de resíduos não cisteína no primeiro e segundo segmentos para cisteína e levando à formação da ligação entre os resíduos mutados. Os resíduos substituídos com cisteína são aqueles cujos carbonos β respectivos distam aproximadamente 6 Å (corrida 0,6) ou menos e, de preferência, na gama de 3,5 Å (0,35 nm) a 5,9 Å (0,59 nm) no TCR nativo. A ligação dissulfureto é entre os resíduos do domínio constante de imunoglobulina. Os locais em que

são introduzidas cisteínas para formar a ligação dissulfureto são os resíduos que se seguem no exão 1 de TRAC*01 para a cadeia α de TCR e de TRBC*01 ou TRBC2*01 para a cadeia β de TCR.

Cadeia α de TCR	Cadeia β de TCR	Separação dos carbonos β nativos (nm)
Thr 48	Ser 57	0,473
Thr 45	Ser 77	0,533
Tyr 10	Ser 17	0,359
Thr 45	Asp 59	0,560
Ser 15	Glu 15	0,59

Os motivos que se seguem nas respectivas cadeias TCR humanas podem ser usados para identificar o resíduo a ser mutado (o resíduo a sombreado é o resíduo para mutação de uma cisteína).

Cadeia α Thr 48:

DSDVYITDK[shaded]VLDMRSMDFK

(aminoácidos 39-58 do exão 1 do gene TRAC*01) (SEQ ID 1)

Cadeia α Thr 45:

QSKDSDVYI[shaded]DKTVLDMRSM

(aminoácidos 36-55 do exão 1 do gene TRAC*01) (SEQ ID 2)

Cadeia α Tyr 10:

DIQNPDPAV[shaded]QLRDSKSSDK

(aminoácidos 1-20 do exão 1 do gene TRAC*01) (SEQ ID 3)

Cadeia α Ser 15:

DPAVYQLRDSKSSDKSVCLF

(aminoácidos 6-25 do exão 1 do gene TRAC*01) (SEQ ID 4)

Cadeia β Ser 57:

NGKEVHSGVSTDPQPLKEQF

(aminoácidos 48-67 do exão 1 do gene TRBC1*01 & TRBC2*01)
(SEQ ID 5)

Cadeia β Ser 77:

ALNDSRYALSSRLRVSATFW

(aminoácidos 68-87 do exão 1 do gene TRBC1*01 & TRBC2*01)
(SEQ ID 6)

Cadeia β Ser 17:

PPEVAVFEPSEAEISHTQKA

(aminoácidos 8-27 do exão 1 do gene TRBC1*01 & TRBC2*01)
(SEQ ID 7)

Cadeia β Asp 59:

KEVHSGVSTDPQPLKEQPAL

(aminoácidos 50-69 do exão 1 do gene TRBC1*01 & TRBC2*01)
(SEQ ID 8)

Cadeia β Glu 15:

VFPPEVAVFEPSEAEISHTC

(aminoácidos 6-25 do exão 1 do gene TRBC1*01 & TRBC2*01)
(SEQ ID 9)

Noutras espécies, as cadeias TCR podem não ter uma região que possua 100% de identidade com os motivos atrás descritos. No entanto, os familiarizados com a matéria poderão usar os motivos atrás descritos para identificar a parte equivalente da cadeia de TCR α ou β e, portanto, o resíduo a ser mutado para cisteína. As técnicas de alinhamento podem ser usadas para este fim. Por exemplo, pode ser usado ClustalW, disponível no sítio do European Bioinformatics Institute (<http://www.ebi.ac.uk/index.html>) para comparar os motivos acima com uma sequência da cadeia TCR particular, de forma a localizar a parte relevante da sequência de TCR relativamente a mutações da sequência de TCR.

O presente invento inclui, dentro do seu âmbito, análogos de $\alpha\beta$ e de $\gamma\delta$ de scTCRs apresentados em partículas fágicas, assim como os de outros mamíferos, incluindo mas não estando limitados a murganho, rato, porco, cabra e ovelha. Como referido atrás, os familiarizados com a matéria serão capazes de determinar locais equivalentes aos locais humanos atrás descritos em que os resíduos cisteína podem ser introduzidos para formar uma ponte dissulfureto entre cadeias. Por exemplo, o que se segue mostra as sequências de aminoácidos dos domínios solúveis C α e C β , juntamente com motivos possuidores dos resíduos murinos equivalentes aos resíduos humanos atrás referidos que podem ser mutados para cisteína para formar uma ligação dissulfureto entre cadeias de TCR (em que os resíduos relevantes estão a sombreado):

Domínio solúvel de C α murino:

**PYIQNPEPAVYQLKDPRSQDSTLCLFTDFDSQINVPKTMESGTFITDKTVLDMK
AMDSKSNGAIAWSNQTSFTCQDIFKETNATYPSSDVP (SEQ ID 10)**

Domínio solúvel de C β murino:

**EDLRNVTPPKVSLFEPKAEIANKQKATLVCLARGFFPDHVELSWVNGREV
HSGVSTDPQAYKESNYSYCLSSRLRVSATFWHNPRNHFRQCQVQFHGLSEEDK
WPEGSPKPVTQNISAEAWGRAD (SEQ ID 11)**

Equivalente murino da cadeia α humana Thr 48:

ESGTFITDKTVLDMKAMDSK (SEQ ID 12)

Equivalente murino da cadeia α humana Thr 45:

KTMESGTFITDKTVLDMKAM (SEQ ID 13)

Equivalente murino da cadeia α humana Tyr 10:

YIQNPEPAVYQLKDPRSQDS (SEQ ID 14)

Equivalente murino da cadeia α humana Ser 15:

AVYQLKDPRSQDSTLCLFTD (SEQ ID 15)

Equivalente murino da cadeia β humana Ser 57:

NGREVHSGVSTDPQAYKESN. (SEQ ID 16)

Equivalente murino da cadeia β humana Ser 77:

KESNYSYCLSSRLRVSATFW (SEQ ID 17)

Equivalente murino da cadeia β humana Ser 17:

PPKV SLFEPKAEIANKQKA (SEQ ID 18)

Equivalente murino da cadeia β humana Asp 59:

REVHSGVSTDPQAYKESNYS (SEQ ID 19)

Equivalente murino da cadeia β humana Glu 15:

VTPPKVSLFEPKAEIANKQ (SEQ ID 20)

Como discutido atrás, os domínios extracelulares constantes scTCR A6 Tax podem ser usados como estrutura com a qual podem ser fundidos domínios variáveis heterólogos. Prefere-se que as sequências do domínio variável heterólogas sejam ligadas às sequências do domínio constante em qualquer ponto entre a ligação dissulfureto e os extremos N das sequências do domínio constante. No caso das sequências dos domínios constantes α e β de TCR A6 Tax, a ligação dissulfureto pode ser efectuada entre resíduos de cisteína introduzidos nos resíduos de aminoácidos 158 e 172, respectivamente. Assim, prefere-se que os pontos de ligação das sequências dos domínios variáveis das cadeias α e β heterólogas sejam entre os resíduos 159 e 173 e o extremo N das sequências dos domínios constantes α e β , respectivamente.

Apresentação de TCR

O método de apresentação de TCR *in vivo* para selecção por imobilização para identificar TCRs tendo as propriedades pretendidas, tais como afinidade elevada para um complexo péptido alvo-MHC, é a apresentação fágica.

Em primeiro lugar, foi construída uma biblioteca de DNA que codifica um arranjo diversificado de scTCRs ou dTCRs mutados. Esta biblioteca foi construída usando DNA codificador de um TCR nativo como matriz de amplificação. Existe uma série de métodos, conhecidos dos familiarizados

com a matéria, para a introdução das mutações pretendidas no DNA de TCR e, finalmente, na proteína TCR expressa. Meios adequados para introdução de mutações nas sequências de TCR podem ser PCR com introdução de erros (EP-TCR), técnicas mistura de DNA ("shuffling") e a utilização de estirpes bacterianas mutagénicas tais como XL-1-Red. É particularmente preferido que estas mutações sejam introduzidas no domínio definido dos TCRs. Por exemplo, as mutações no domínio variável, particularmente as regiões determinantes de complementaridade (CDRs) e/ou regiões estruturais são provavelmente os locais mais adequados para a introdução de mutações conduzindo à produção de uma biblioteca diversificada de TCRs para a produção de TCRs com melhores propriedades de ligação ao ligando. EP-PCR é um exemplo de um método através do qual tais mutações específicas de região podem ser introduzidas nos TCRs. Foram usadas sequências iniciadoras de ER-PCR complementares das sequências de DNA que delimitam a região a ser mutada para amplificar múltiplas cópias desta região do DNA de TCR, as quais possuem um nível controlável de mutações ao acaso. Estas sequências de DNA codificadoras de regiões mutadas são inseridas nas sequências de DNA, as quais codificam as secções não mutagenizadas do TCR, através da ligação ou de PCR sobreponível. O DNA codificador do TCR com a região mutada pode então ser ligado ao DNA codificador de um polipéptido heterólogo de forma a produzir uma proteína de fusão adequada para apresentação. No caso da apresentação fágica, o vector de expressão utilizado é um fagemídeo ou um genoma fágico em

que o DNA de TCR pode ser ligado ao DNA codificador de uma proteína de superfície, de preferência a proteína de superfície gIII ou gVIII. No caso de um scTCR tal ligação é realizada por apresentação fágica de qualquer péptido ou polipéptido monomérico. No caso de dTCRs, apenas uma das cadeias de TCR é ligada, como referido. A outra cadeia é codificada no ácido nucleico para co-expressão com fagemídeo e com o ácido nucleico do fago auxiliar, de forma que a segunda cadeia expressa encontre e se associe ao fago expresso com a primeira cadeia apresentada na superfície. Em ambos os casos, como discutido mais detalhadamente abaixo, as cisteínas adequadamente posicionadas nos domínios constantes são úteis para fazer com que os domínios variáveis do TCR adotem as suas posições funcionais, através da formação de uma ligação dissulfureto entre aquelas cisteínas.

Para a expressão, um vector de expressão compreendendo (a) um ácido nucleico codificador de uma cadeia de um par de polipéptidos dTCR e (b) a outra cadeia do par de polipéptidos dTCR fundida com uma sequência de ácidos nucleicos codificadora de uma proteína formadora de partículas ou uma proteína da superfície celular; ou ácido nucleico codificador de um polipéptido scTCR fundido com uma sequência de ácido nucleico codificadora de uma proteína formadora de partículas ou uma proteína da superfície celular, o par dTCR, ou uma composição compreendendo um primeiro vector compreendendo ácido nucleico (a) e um segundo vector compreendendo ácido

nucleico (b), são postos em contacto com as células hospedeiras capazes de causar a expressão do material genético codificado, em condições adequadas à transformação das referidas células. Tais vectores de expressão, sistemas de expressão compreendendo vectores fagemídeos ou de genoma fágico codificadores de dTCRs e scTCRs e células hospedeiras portadoras dos mesmos, constituem outros aspectos do presente invento. Numa realização preferida do invento, os vectores de fagemídeo ou de genoma fágico derivam de um fago filamentoso.

As células transformadas são então incubadas para permitir a expressão das partículas fágicas apresentadoras de TCR. Estas partículas podem então ser usadas para o rastreio ou em ensaios para identificar variantes de TCR com características específicas melhoradas. Quaisquer partículas que possuam características melhoradas podem ser então isoladas. O DNA codificador destes TCRs pode ser então amplificado por PCR e determinada a sua sequência.

É conhecido que níveis de expressão elevados de um polipéptido exógeno podem ser tóxicos para a célula hospedeira. Em tais casos, deve ser encontrada uma estirpe hospedeira que seja mais tolerante para o polipéptido exógeno ou os níveis de expressão na célula hospedeira devem ser limitados a um nível que seja tolerado. Por exemplo, (Beekwilder *et al.*, (1999) Gene 228 (1-2) 23-31) descreve que apenas as formas mutadas de um inibidor da protease de batata (PI2) que continha deleções ou codões de

paragem amber seriam seleccionadas com êxito de uma biblioteca fágica. No presente caso, uma observação no decurso do trabalho descrita nos Exemplos sugere que pode ser desejável limitar os níveis de expressão da proteína TCR apresentada em partículas do invento, pelo menos nalgumas estirpes de *E. coli*. Assim, a TCR A6 seleccionada no Exemplo 4, após ciclos repetidos de cultura, demonstrou-se que deriva de células em que o fagemídeo foi mutado relativamente ao introduzido no início. A mutação criou um codão de paragem "opal" na cadeia TCR β . Este codão é "ultrapassado sem reconhecimento" com uma frequência baixa pelos ribossomas da estirpe *E. coli* utilizada, resultando na inserção de um resíduo de triptofano neste local e num nível global muito reduzido da expressão da cadeia β de tamanho completo.

Existem várias estratégias para limitar os níveis de expressão de um polipéptido exógeno a partir de um determinado sistema de expressão num hospedeiro, o qual pode ser adequado para limitar os níveis de expressão de um scTCR ou de uma ou de ambas as cadeias TCR de um dTCR. Por exemplo:

Utilização de uma sequência de promotor fraco - O nível de expressão obtido para um determinado produto, como seja a cadeia α ou β de TCR, pode ser manipulado usando sequências de promotor de diversas forças. O promotor P_{RM} do fago lambda é um exemplo de um promotor fraco.

Locais de ligação a ribossomas mutados (RBS's) - A mutação de um ácido nucleico simples no RIBS associado a um produto de gene, como seja a cadeia α ou β de TCR, pode resultar num nível reduzido de expressão. Por exemplo, a mutação de uma sequência selvagem AGGA para AGGG.

"Codões de partida" mutados - A mutação de um ácido nucleico simples no codão de partida associado a um produto de gene, como seja a cadeia α ou β de TCR, pode também resultar num nível reduzido de expressão. Por exemplo, a mutação de um codão de partida selvagem AUG para GUG.

Mutações supressoras codificadoras de codões de paragem - Estas são inseridas nas regiões codificadoras da cadeia β de TCR. Exemplos incluem o codão de paragem "opal" (UGA), este codão de paragem que pode ser ignorado ("leaky") resulta na inserção, com baixa frequência, de um aminoácido triptofano e leitura do resto da sequência codificadora.

Modificação mediada por metabolitos da força do promotor - O nível de expressão de um produto de gene, como seja a cadeia α ou β de TCR, sob o controlo de determinados promotores pode ser regulado negativamente pela adição um metabolito relevante às células contendo o promotor. Por exemplo, as adições de glucose podem ser usadas para regular negativamente a expressão de um produto de gene sob o controlo de um promotor Lac.

Utilização de codões - Células bacterianas e, por exemplo, células de mamífero possuem diferentes "preferências" relacionadas com os codões que usam para codificar determinados aminoácidos. Por exemplo, as células bacterianas normalmente usam o codão CGU para codificar arginina, enquanto que as células eucarióticas geralmente usam AGA. É possível reduzir o nível de expressão de um produto, como seja a cadeia α ou β de TCR, através da utilização de sequências de DNA que possuam uma série de codões menos preferidos para o sistema de expressão a ser usado.

Os detalhes relacionados com os meios atrás referidos para regular negativamente a expressão de produtos de genes podem ser encontrados em Glass (1982) "Gene Function - *E. coli* and its heritable elements", Croom Helm) e (Razinoff (1980) "The operon" 2nd Edition, Cold Spring Harbor Laboratory).

É também conhecido que o fornecimento às culturas bacterianas de uma concentração relativamente elevada de um açúcar, como seja a sucrose, pode aumentar os níveis de expressão periplasmática de proteínas solúveis. (Ver por exemplo, (Sawyer *et al.*, (1994) Protein Engineering 7 (11) 1401-1406)).

Após expressão, o emparelhamento correcto das sequências do domínio variável do polipéptido scTCR é, de

preferência, auxiliado pela introdução de uma ligação dissulfureto no domínio constante extracelular da scTCR. Sem pretender estar limitado pela teoria, crê-se que a nova ligação dissulfureto proporcione estabilidade extra ao scTCR durante o processo de renaturação e facilite assim o emparelhamento correcto do primeiro e do segundo segmento.

Também, como referido atrás, para a apresentação fágica de TCR, um dos polipéptidos do par de dTCR é expresso como se fosse eventualmente apresentado como polipéptido monomérico no fago e o outro elemento do par de polipéptidos de dTCR é co-expresso na mesma célula hospedeira. Como a partícula fágica é auto-montável, os dois polipéptidos auto-associam-se para serem apresentados como um dímero no fago. Novamente, na realização preferida deste aspecto do invento, o enrolamento correcto durante a associação do par de polipéptidos é auxiliado por uma ligação dissulfureto entre as sequências constantes, como discutido atrás. Outros detalhes de um procedimento para a apresentação fágica de um dTCR tendo uma ligação dissulfureto entre cadeias surgem nos Exemplos apresentados.

Como alternativa, a o fago apresentador da primeira cadeia do TCR pode ser expresso em primeiro lugar e o segundo polipéptido da cadeia pode ser posto em contacto com o fago expresso num passo subsequente, para associação como um dTCR funcional na superfície fágica.

O método de apresentação de TCR, *in vitro*, preferido para selecção por imobilização biológica para identificar TCRs tendo as propriedades pretendidas, tais como elevada afinidade para um complexo péptido alvo-MHC é a apresentação ribossómica. Em primeiro lugar, foi construída uma biblioteca de DNA que codifica um arranjo diversificado de polipéptidos scTCRs ou dTCRs mutados usando as técnicas discutidas atrás. A biblioteca de DNA é então colocada em contacto com a RNA polimerase de forma a produzir uma biblioteca de mRNA complementar. Facultativamente, para as técnicas de apresentação de mRNA, as sequências de mRNA podem ser então ligadas a uma sequência de DNA compreendendo um local de ligação a puromicina. Estas construções genéticas são então colocadas em contacto com ribossomas *in vitro*, em condições que permitam a tradução do polipéptido scTCR ou do primeiro polipéptido do par de dTCR. No caso do dTCR, o segundo elemento do par dTCR é expresso separadamente e colocado em contacto com o polipéptido primeiro apresentado pelos ribossomas, para associação entre os dois, de preferência ajudado pela formação da ligação dissulfureto entre os domínios constantes. Como alternativa, o mRNA codificador de ambas as cadeias do TCR pode ser colocado em contacto com ribossomas *in vitro*, em condições que permitam a tradução das cadeias de TCR, de forma que se forme um ribossoma apresentador de um dTCR. Estes ribossomas apresentadores de scTCRs e de dTCR podem então ser usados para o rastreio ou em ensaios para identificar variantes de TCR com características específicas melhoradas. Quaisquer partículas que possuam características melhoradas nestes

estudos podem então ser isoladas. O mRNA codificador destes TCRs pode então ser convertido em sequências de DNA complementares usando transcriptase reversa. Este DNA pode depois ser amplificado por PCR e determinada a sua sequência.

Aspectos adicionais

Uma partícula fágica apresentadora de um scTCR ou de um dTCR (que preferencialmente é constituído por sequências constantes e variáveis correspondendo às sequências humanas) do presente invento pode ser proporcionada na forma substancialmente pura ou como preparação purificada ou isolada. Por exemplo, pode ser proporcionada numa forma que esteja substancialmente livre de outras proteínas.

Uma partícula fágica apresentadora de uma pluralidade de scTCRs ou dTCRs do presente invento pode ser proporcionada num complexo multivalente. Assim, o presente invento proporciona, num aspecto, um complexo de receptores de células T (TCR) multivalente, que compreende uma partícula fágica apresentadora de uma pluralidade de scTCRs ou de dTCRs como aqui descrito. Cada um da pluralidade dos referidos scTCRs ou dTCRs é, preferencialmente, idêntico.

Isolamento de variantes de TCR com características melhoradas

Um outro aspecto do invento é um método para a

identificação de TCRs com uma característica específica, o referido método compreendendo a submissão de uma biblioteca diversificada de TCRs apresentados em partículas fágicas a um processo de selecção que escolhe as referidas características e isolamento das partículas fágicas que apresentem um TCR tendo a referida característica e, facultativamente, a um processo de amplificação para multiplicar as partículas isoladas, e/ou processo de rastreio que mede a referida característica, identificação das partículas fágicas que apresentam um TCR com a característica pretendida e isolamento destas partículas fágicas e, facultativamente, a um processo de amplificação para multiplicar as partículas isoladas.

As sequências de DNA codificadoras dos TCRs variantes podem então ser obtidas e amplificadas por PCR para permitir que sejam determinadas as suas sequências. As características que podem ser melhoradas incluem, mas não estão limitadas a afinidade de ligação do ligando e estabilidade da construção.

Utilização do rastreio

As partículas fágicas apresentadoras de TCR do presente invento são susceptíveis de serem utilizadas nos métodos de rastreio destinados a identificar moduladores, incluindo inibidores, da sinapse imune celular mediada por TCR.

Como é do conhecimento dos familiarizados com a matéria, existe uma série de formatos de ensaio que proporcionam uma base adequada para rastreios da interacção proteína-proteína deste tipo.

Os sistemas de Ensaio Homogénio de Luminescência Amplificada pela Proximidade ("Amplified Luminescent Proximity Homogeneous Assay") tais como o AlphaScreen™, baseiam-se na utilização de esferas "Dadora" e "Aceitadora" que estão revestidas com uma camada de hidrogel à qual podem ser ligados receptores e ligandos. A interacção entre estas moléculas de receptor e de ligando conduz à proximidade entre as esferas. Quando estas esferas são expostas a luz laser, um foto-sensibilizador na esfera "Dadora" converte o oxigénio ambiente para um estado de singelete mais excitado. As moléculas de oxigénio no estado singelete difundem-se para reagirem com uma substância quimioluminescente na esfera "Aceitadora" que ainda activa os fluoróforos contidos na mesma esfera. Os fluoróforos, subsequentemente, emitem luz a 520-620 nm, esta sinaliza a ocorrência da interacção receptor-ligando. A presença de um inibidor da interacção receptor-ligando faz com que este sinal seja diminuído.

A ressonância de plasmões de superfície (SPR) é um ensaio óptico interfacial, em que um parceiro de ligação (normalmente o receptor) é imobilizado num "chip" (a superfície do sensor) e é detectada a ligação do outro parceiro de ligação (normalmente o ligando), que é solúvel

e que flui ao longo do chip. A ligação do ligando resulta num aumento na concentração da proteína perto da superfície do chip que causa uma alteração no índice de refração naquela região. A superfície do chip tem uma composição tal que a alteração no índice de refração pode ser detectada por ressonância de plasmões de superfície, um fenómeno óptico através do qual a luz, num determinado ângulo de incidência num filme de metal fino, produz um feixe reflectido de intensidade reduzida devido à excitação de ressonância de ondas de densidade oscilante da carga de superfície (plasmões de superfície). A ressonância é muito sensível a alterações no índice de refração no lado distal do filme de metal e é este sinal que é usado para detectar a ligação entre as proteínas imobilizada e solúvel. Estão comercializados sistemas que permitem a utilização conveniente da detecção SPR das interacções moleculares e análise de dados. São exemplos as máquinas Iasys™ (Fisons) e as máquinas Biacore™.

Outros ensaios ópticos interfaciais incluem fluorescência de reflexão interna total ("total internal reflectance fluorescence", TIRF), espelho ressonante (RM) e sensores de acoplagem com grelhas ópticas ("optical grating coupler sensor", GCS) e estão discutidos mais detalhadamente em Woodbury e Ventn (J. Chromatog. B. 725 113-137 (1999)). O ensaio de cintilação de proximidade (SPA) foi usado no rastreio de bibliotecas de compostos relativamente a inibidores de interacção de baixa afinidade entre CD28 e B7 (K_d provavelmente na região de 4 μ M (Van

der Merwe *et al.*, J. Exp. Med. 185:393-403 (1997), Jenh *et al.*, Anal Biochem 165(2) 287-93 (1998)). SPA é um ensaio radioactivo que utiliza a emissão de partículas beta por certos isótopos radioactivos com transferência de energia para um cintilador imobilizado na superfície indicadora. O percurso curto das partículas beta em solução assegura que a cintilação apenas ocorre quando as partículas beta são emitidas em estreita proximidade relativamente ao cintilador. Quando aplicado na detecção das interacções proteína-proteína, um parceiro de interacção é marcado com o isótopo radioactivo, enquanto o outro é ligado a esferas contendo cintilador ou usado para revestir uma superfície juntamente com o cintilador. Se o ensaio puder ser estabelecido de forma ideal, o isótopo radioactivo será colocado em contacto suficientemente estreito com o cintilador para que a emissão de fotões seja activada apenas quando ocorre a ligação entre as duas proteínas.

TCRs de elevada afinidade

O presente invento permite a identificação de TCRs mutados específicos para um determinado ligando de TCR, com maior afinidade para o referido ligando de TCR do que o TCR selvagem. Espera-se que estes TCRs de elevada afinidade sejam particularmente úteis para o diagnóstico e tratamento de doenças.

Como aqui é usado, o termo "TCR de elevada afinidade" refere-se a um scTCR ou dTCR mutado que interage

com um ligando de TCR específico e possui um K_d para o referido ligando de TCR inferior ao do TCR nativo correspondente, conforme medido pela Ressonância de Plasmões de Superfície ou tem uma constante de dissociação ("off-rate", K_{off}) para o referido ligando de TCR inferior à do correspondente TCR nativo conforme medido pela Ressonância de Plasmões de Superfície.

scTCRs ou dTCRs de elevada afinidade proporcionados pelo presente invento são, de preferência, mutados relativamente à TCR nativa em pelo menos uma região determinante de complementaridade e/ou regiões estruturais.

Num aspecto, o ligando de TCR para o qual um determinado TCR de elevada afinidade é específico é um complexo péptido-MHC (pMHC).

Num outro aspecto, o ligando de TCR para o qual um determinado TCR de elevada afinidade é específico é um tipo ou tipos de MHC.

Num outro aspecto, o ligando de TCR para o qual um determinado TCR de elevada afinidade é específico é o complexo HLA-A2 péptido tax (LLFYVYV) (SEQ ID 21).

Num outro aspecto o ligando de TCR para o qual um determinado TCR de elevada afinidade é específico é o complexo HLA-A2 péptido NY-ESO (SLLMITQC) (SEQ ID 22)

Um scTCR de elevada afinidade, ou uma ou ambas as cadeias de dTCR de elevada afinidade, pode ser marcado com um composto que confira uma imagem, por exemplo uma marca que é adequada para fins de diagnóstico. Tais TCRs de elevada afinidade marcados são úteis num método para detecção de um ligando de TCR seleccionado a partir de complexos CD1-antigénio, superantigénios bacterianos e complexos de MHC-péptido/superantigénio, método esse que compreende o contacto do ligando de TCR com um TCR de elevada afinidade (ou um complexo de TCR tetramérico de elevada afinidade) o qual é específico do ligando de TCR; e detecção da ligação ao ligando de TCR. Em complexos de TCR tetraméricos de elevada afinidade (formados, por exemplo, usando heterodímeros biotinilados) pode ser usada estreptavidina fluorescente (comercial) para se conseguir uma marca detectável. Um tetrâmero marcado com fluorescência é adequado para usar na análise por FACS, por exemplo para detectar células apresentadoras de antigénio portadoras do péptido para o qual é específico o TCR de elevada afinidade.

Uma outra forma em que os TCRs solúveis de elevada afinidade proporcionados pelo presente invento podem ser detectados é através da utilização de anticorpos específicos de TCR, em particular anticorpos monoclonais. Existem muitos anticorpos comerciais anti-TCR, como seja α F1 e β F1, os quais reconhecem os domínios constantes das cadeias α e β , respectivamente.

Um TCR de elevada afinidade (ou seu complexo multivalente) proporcionado pelo presente invento pode, em alternativa ou em adicionalmente, estar associado (e.g., ligado covalentemente ou de outra forma) a um agente terapêutico que pode ser, por exemplo, um grupo tóxico para usar na morte celular ou um agente imunoestimulador como seja uma interleucina ou uma citocina. Um complexo TCR de elevada afinidade multivalente pode ter maior capacidade de ligação a um ligando de TCR comparativamente com um heterodímero do receptor de células T selvagem multimérico ou de elevada afinidade. Assim, os complexos de TCR de elevada afinidade multivalentes são particularmente úteis para seguir ou atingir células apresentadoras de antigénios particulares *in vitro* ou *in vivo* e são também úteis como intermediários para a produção de complexos de TCR de elevada afinidade multivalentes tendo tais utilizações. O TCR de elevada afinidade ou o complexo de TCR de elevada afinidade multivalente pode assim ser proporcionado numa formulação farmacêuticamente aceitável para usar *in vivo*.

Um agente terapêutico pode ser entregue numa célula alvo através do contacto de potenciais células alvo com um TCR de elevada afinidade ou com um complexo de TCR de elevada afinidade multivalente em condições que permitam a ligação do TCR de elevada afinidade ou do complexo de TCR de elevada afinidade multivalente à célula alvo, o referido TCR de elevada afinidade ou o complexo de TCR de elevada afinidade multivalente sendo específico para o ligando de TCR e tendo o agente terapêutico associado a ele.

Em particular, o TCR solúvel de elevada afinidade ou o complexo de TCR de elevada afinidade multivalente pode ser usado na entrega de agentes terapêuticos a células que apresentem um antigénio particular. Isto será útil em muitas situações e, em particular, contra tumores. Um agente terapêutico poderá ser entregue de forma a exercer o seu efeito localmente mas não apenas sobre as células a que se liga. Assim, uma estratégia particular tem como alvo moléculas anti-tumorais ligadas a receptores de células T de elevada afinidade ou complexos de TCR de elevada afinidade multivalentes específicos para antígenos tumorais.

Muitos agentes terapêuticos poderão ser empregues para este uso, por exemplo compostos radioactivos, enzimas (perforina por exemplo) ou agentes quimioterapêuticos (por exemplo, cis-platina). Para assegurar que os efeitos tóxicos são exercidos no local pretendido, a toxina poderá estar dentro de um lipossoma ligado a estreptavidina de forma que o composto seja libertado lentamente. Isto evitará efeitos danosos durante o transporte no corpo e assegurará que a toxina tem um efeito máximo após ligação do TCR às células apresentadoras de antigénio relevantes.

Outros agentes terapêuticos adequados incluem:

- pequenas moléculas de agente citotóxico, *i.e.* compostos com a capacidade para matar células de mamífero

tendo um peso molecular inferior a 700 daltons. Tais compostos poderão também conter metais tóxicos capazes de ter um efeito citotóxico. Ainda, deve ser entendido que estas pequenas moléculas de agentes citotóxicos também incluem pró-fármacos, i.e. compostos que decaem ou são convertidos, em condições fisiológicas, para libertarem os agentes citotóxicos. Exemplos de tais agentes incluem cisplatina derivados de maitansina, raquelmicina, caliqueamicina, docetaxel, etopósido, gencitabina, ifosfamida, irinotecano, melfalan, mitoxantrona, sorfimer sodiofotofrina II, temozolmida, topotecano, trimetreato glucuronato, auristatina E vincristina e doxorubicina:

- citotoxinas peptídicas, i.e. proteínas ou fragmentos das mesmas, com capacidade para matar células de mamífero. Os exemplos incluem rícino, toxina da difteria, exotoxina A bacteriana de *Pseudomonas A*, DNase e RNase;

- radionuclidos, i.e. isótopos instáveis de elementos que decaem com a emissão simultânea de uma ou mais partículas α ou β , ou raios γ . Exemplos incluem iodo 131, rénio 186, índio 111, ítrio 90, bismuto 210 e 213, actínio 225 e astamina 213; podem ser usados agentes quelantes para facilitar a associação destes radionuclidos a TCRs de elevada afinidade ou seus multímeros;

- pró-fármacos, tais como enzimas pró-fármacos direccionadas por anticorpos;

- imuno-estimuladores, i.e. grupos que estimulam a resposta imune. Os exemplos incluem citocinas tais como IL-2, quimiocinas tais como IL-8, factor das plaquetas 4, proteína estimuladora do crescimento de melanoma, etc.,

anticorpos ou fragmentos dos mesmos, activadores do complemento, domínios proteicos xenogeneicos, domínios proteicos alogeneicos, domínios de proteínas virais/bacterianas e péptidos virais/bacterianos.

Os TCRs solúveis de elevada afinidade ou complexos multivalentes de TCR de elevada afinidade proporcionados pelo invento podem ser ligados a uma enzima capaz de converter uma pró-fármaco num fármaco. Isto permite que o pró-fármaco seja convertido no fármaco apenas no local onde é necessário (*i.e.*, dirigido pelo sTCR).

Uma variedade de tratamentos de doenças pode, potencialmente, ser aumentada pela localização do fármaco através da especificidade de TCRs solúveis de elevada afinidade. Por exemplo, espera-se que os TCRs A6 de elevada afinidade, específicos de HLA-A2-tax (LLFGYPVYV) (SEQ ID 21), aqui descritos possam ser usados em métodos para o diagnóstico e tratamento de HTLV-1 e que TCRs NY-ESO de elevada afinidade específico de HLA-A2-NY-ESO (SLLMITQC) (SEQ ID 22) aqui descritos possam ser usados em métodos para o diagnóstico e tratamento de cancro.

As doenças virais para as quais existem fármacos, *e.g.*, HIV, SIV, EBV, CMV, beneficiarão se o fármaco for libertado ou activado na vizinhança das células infectadas. Para o cancro, a localização na vizinhança de tumores ou metástases aumentará o efeito das toxinas ou imunoestimuladores. Nas doenças auto-imunes, os fármacos

imuno-supressores poderão ser libertados lentamente, tendo maior efeito local ao longo de mais tempo, ao mesmo tempo afectando minimamente a imunocapacidade global do indivíduo. Na prevenção da rejeição de enxertos, o efeito dos fármacos imuno-supressores poderão ser optimizados da mesma forma. Para a administração de vacinas, o antigénio da vacina poderá ser localizado na vizinhança de células apresentadoras de antigénio, aumentando assim a eficácia do antigénio. O método pode também ser aplicado para fins de imagiologia.

Os TCRs solúveis de elevada afinidade proporcionados pelo presente invento podem ser usados para modular a activação de células T através da ligação a ligandos de TCR específicos e assim inibir a activação de células T. As doenças auto-imunes que envolvem inflamação mediada por células T e/ou destruição de tecidos serão susceptíveis a esta abordagem, por exemplo diabetes tipo I. O conhecimento dos epitopos específicos apresentados pelo pMHC relevante é necessário para esta utilização.

Os TCRs de elevada afinidade terapêuticos, ou para imagiologia, proporcionados pelo invento serão geralmente fornecidos como parte de uma composição farmacêutica estéril que, normalmente, incluirá um veículo farmacêuticamente aceitável. Esta composição farmacêutica pode ser em qualquer forma adequada, (dependendo do método pretendido de administração ao doente). Pode ser administrado na forma de dosagem unitária, será geralmente

fornecido num recipiente selado e pode ser oferecido como parte de um kit. Tal kit, normalmente, incluirá (ainda que não necessariamente) instruções para utilização. Poderá incluir uma pluralidade das referidas formas de dosagem.

A composição farmacêutica pode ser adaptada para administração por uma via adequada, por exemplo parentérica, transdérmica ou através de inalação, de preferência uma via parentérica (incluindo subcutânea, intramuscular ou, mais de preferência intravenosa). Tais composições podem ser preparadas por qualquer método conhecido no campo farmacêutico, por exemplo através da mistura do ingrediente activo com um ou mais veículos ou excipientes em condições estéreis.

As dosagens das substâncias do presente invento podem variar entre limites largos, dependendo da doença ou perturbação a ser tratada, da idade e da condição do indivíduo a ser tratado, etc. e um médico determinará em último caso as dosagens adequadas a serem usadas.

Exemplos

O invento é ainda descrito nos exemplos que se seguem, os quais não limitam, de forma alguma, o âmbito do invento.

A seguir é feita referência aos desenhos juntos em que:

As Figuras 1a e 1b mostram, respectivamente, as sequências de ácido nucleico de cadeias TCR A6 α e β , mutadas de forma a introduzir um codão cisteína. O sombreado indica os codões cisteína introduzidos.

A Figura 2a mostra a sequência de aminoácidos extracelular da cadeia α de TCR A6, incluindo a mutação T₄₈→C (sublinhado) usada para produzir a nova ligação dissulfureto intercadeias e a Figura 2b mostra a sequência de aminoácidos extracelular da cadeia β de TCR A6, incluindo a mutação S₅₇→C (sublinhada) usada para produzir a nova ligação dissulfureto intercadeias.

A Figura 3 Descreve a clonagem das cadeias TCR α e β nos vectores fagemídeos. O diagrama descreve um vector de apresentação fágica. RSB é o local de ligação ao ribossoma. S1 e S2 são péptidos sinal para secreção de proteínas para o periplasma de *E. coli*. * indica o codão de paragem da tradução. Quer a cadeia α quer a cadeia β podem ser fundidas com a proteína externa do fago, no entanto neste diagrama apenas a cadeia β de TCR é fundida com a proteína externa do fago.

Figura 4 descreve detalhadamente a sequência de DNA do fagemídeo pEX746:A6.

Figura 5. Expressão bacteriana de partículas fágicas com fusões da proteína externa e TCR A6

heterodimérico em *E. coli*. As proteínas de fusão de TCR A6 heterodimérico::gIII são detectadas usando transferências Western. As partículas fágicas são preparadas a partir de *E. coli* XL-1-Blue e concentradas com PEG/NaCl. As amostras foram aplicadas em tampão de amostra redutor e não redutor. A pista 1 é a amostra do clone 7 contendo a sequência correcta e a pista 2 é a amostra do clone 14 contendo uma deleção no gene codificador da cadeia α . A proteína de fusão de TCR A6 heterodimérica::gIII foi detectada nos 125 kDa.

A Figura 6 ilustra a detecção por ELISA da actividade de ligação do complexo péptido pMHC de um TCR A6 heterodimérico apresentado no fago. O clone 7 liga-se especificamente ao complexo HLA A2-Tax. O clone 14 não se pode ligar a qualquer pMHC, uma vez que nenhum TCR está ligado às partículas fágicas.

A Figura 7a é uma ilustração esquemática da construção de apresentação nos ribossomas do DNA de TCR A6 de cadeia simples-C-kappa.

As Figuras 7b e 7c descrevem detalhadamente a sequência de DNA codificadora e a sequência de aminoácidos completas da construção de apresentação nos ribossomas de DNA de TCR A6 de cadeia simples-C-kappa codificada em pUC19, respectivamente.

A Figura 8 descreve detalhadamente a sequência de DNA de pUC19-T7.

A Figura 9 descreve detalhadamente a sequência de DNA da construção de apresentação nos ribossomas de DNA de TCR A6 de cadeia simples-C-kappa que foi clonada em pUC19-T7.

A Figura 10 é uma transferência Western que mostra a detecção de TCR A6 de cadeia simples-C-kappa traduzido *in vitro* usando os lisados de reticulócitos de coelho Ambion.

A Figura 11 RT-PCR do mRNA de TCR A6 de cadeia simples-C-kappa em esferas recuperadas das reacções de apresentação em ribossomas.

A Figura 12a descreve detalhadamente a sequência de DNA da cadeia β do clone 9 de TCR A6 mutada; o ácido nucleico mutado está indicado a cheio.

A Figura 12b descreve detalhadamente a sequência de aminoácidos da cadeia β do clone 9 de TCR A6 mutado, a posição correspondendo ao codão de paragem opal introduzido está indicada com um *.

A Figura 13 descreve detalhadamente a sequência de DNA da cadeia β do clone 49 de TCR A6 mutado; o ácido nucleico mutado está indicado a negrito. Uma vez que esta é uma mutação "silenciosa", através da mutação não foi

introduzida qualquer alteração na sequência de aminoácidos resultante.

A Figura 14a descreve detalhadamente a sequência de DNA da cadeia β do clone 134 de TCR A6 mutado; os ácidos nucleicos mutados estão indicados a negrito.

Figura 14b descreve detalhadamente a sequência de aminoácidos da cadeia β do clone 134 de TCR A6 mutado testada pelo ensaio BIAcore; os aminoácidos mutados estão indicados a negrito.

Figura 14c descreve detalhadamente a sequência de aminoácidos da cadeia β do clone 134 de TCR A6 mutado testada pelo ensaio de ELISA para fagos; os aminoácidos mutados estão indicados a negrito.

Figura 15 Dados de BIAcore para a ligação de TCR A6 clone 134 a HLA-A2 Tax e a HLA-A2 NY-ESO.

Figura 16 Dados de BIAcore usados para determinar T_{off} para a ligação de TCR A6 clone 134 a HLA-A2 Tax

As Figuras 17a e 17b mostram as sequências de DNA das cadeias α e β do TCR NY-ESO, respectivamente.

As Figuras 18a e 18b mostram as sequências de aminoácidos das cadeias α e β do TCR NY-ESO, respectivamente.

AS Figuras 19a e 19b mostram as sequências de DNA e de aminoácidos da cadeia β do TCR NY-ESO no fagemídeo pEX746:NY-ESO, respectivamente.

A Figura 20 mostra a ligação específica de partículas fágicas apresentadoras do TCR NY-ESO a HLA-A2-ESO num ensaio de ELISA para fagos.

A Figura 21 mostra a sequência de DNA da cadeia DR1 α que inclui os codões codificadores do péptido de dimerização Fos ligado ao extremo 3' da sequência DRA0101. O sombreado indica os codões Fos e os codões da marca de biotinilação estão indicados no texto a negrito.

A Figura 22 mostra a sequência de DNA do vector bi-cistrónico pAcAB3 usado para a expressão de complexos HLA classe II-péptido em células de insecto Sf9. O local de restrição BglII (AGATCT) usado para inserir a cadeia α de HLA e o local de restrição BamHI (GGATCC) usado para inserir a cadeia β de HLA estão indicados a sombreado.

A Figura 23 mostra a sequência de DNA da cadeia β de DR1 que inclui os codões codificadores do péptido de dimerização Jun ligado ao extremo 3' da sequência DRB0401 e os codões codificadores do péptido que se complexa com HLA ligado ao extremo 5' da sequência DRF0401. O sombreado indica os codões Jun e os codões do péptido Flu HA que se complexa com HLA estão sublinhados.

A Figura 24 mostra um registo BIAcore da ligação de TCR A6 clone 134 de elevada afinidade a células de fluxo revestidas como se segue:

Célula de fluxo 1 (FC 1) - Branco

Célula de fluxo 2 (FC 2) - HLA-A2 (LLGRNSFEV)

(SEQ ID 23)

Célula de fluxo 3 (FC 3) - HLA-A2 (KLVALGINAV)

(SEQ ID 24)

Célula de fluxo 4 (FC 4) - HLA-A2 (LLGDLFGV) **(SEQ**

ID 25)

A Figura 25 mostra um registo BIAcore da ligação do TCR A6 clone 134 de elevada afinidade a células de fluxo revestidas como se segue:

Célula de fluxo 1 (FC 1) - Branco

Célula de fluxo 2 (FC 2) - HLA-B8 (FLRGRAYGL)

(SEQ ID 26)

Célula de fluxo 3 (FC 3) - HLA-B27 (HECQAIRKK)

(SEQ ID 27)

Célula de fluxo 4 (FC 4) - HLA-Cw6 (YRSGIIAVV)

(SEQ ID 28)

A Figura 26 mostra um registo BIAcore da ligação do TCR A6 clone 134 de elevada afinidade a células de fluxo revestidas como se segue:

Célula de fluxo 1 (FC 1) - Branco

Célula de fluxo 2 (FC 2) - HLA-A24 (VYGFVRACL)

(SEQ ID 29)

Célula de fluxo 3 (FC 3) - HLA-A2 (ILAKFLHWL)

(SEQ ID 30)

Célula de fluxo 4 (FC 4) - HLA-A2 (LTLGEFLKL)

(SEQ ID 31)

A Figura 27 mostra um registo BIAcore da ligação do TCR A6 clone 134 de elevada afinidade a células de fluxo revestidas como se segue:

Célula de fluxo 1 (FC 1) - Branco

Célula de fluxo 2 (FC 2) - HLA-DR1 (PKYVKQNTLKL)

(SEQ ID 32)

Célula de fluxo 3 (FC 3) - HLA-A2 (GILGFVFTL)

(SEQ ID 33)

Célula de fluxo 4 (FC 4) - HLA-A2 (SLYNTVATL)

(SEQ ID 34)

A Figura 28 mostra um registo BIAcore da ligação do TCR A6 clone 134 de elevada afinidade a células de fluxo revestidas como se segue:

Célula de fluxo 1 (FC 1) - Branco

Célula de fluxo 4 (FC 4) - HLA-A2 (LLFGYPVYV)

(SEQ ID 21)

As Figuras 29a e 29b mostram os gráficos BIAcore da interacção entre TCR NY-ESO solúvel de elevada afinidade e HLA-A2 NY-ESO.

As Figuras 30a e 30b mostram gráficos BIAcore da interacção entre TCR NY-ESO solúvel selvagem e HLA-A2 NY-ESO.

As Figuras 31a e 31b mostram gráficos BIAcore da interacção entre TCR A6 solúvel mutante (Clone 1) e HLA-A2 Tax.

As Figuras 32a e 32b mostram os gráficos BIAcore da interacção entre TCR A6 solúvel mutante (Clone 111) e HLA-A2 Tax.

As Figuras 33a e 33b mostram os gráficos BIAcore da interacção entre TCR A6 solúvel mutante (Clone 89) e HLA-A2 Tax.

As Figuras 34a e 34b mostram os gráficos BIAcore da interacção entre TCR A6 solúvel mutante (contendo as mutações do Clone 71 e do Clone 134) e HLA-A2 Tax.

A Figura 35 mostra um gráfico BIAcore da interacção entre TCR A6 solúvel mutante (contendo as mutações do Clone 1 e β G102→A) e HLA-A2 Tax.

As Figuras 36a a 36c mostram os gráficos BIAcore da interacção entre TCR A6 solúvel mutante (contendo as mutações do Clone 89 e do Clone 134) e HLA-A2 Tax.

As Figuras 37a e 37b mostram os gráficos BIAcore da interacção entre TCR A6 solúvel mutante (contendo as mutações do Clone 71 e do Clone 89) e HLA-A2 Tax.

A Figura 38 descreve detalhadamente as sequências de aminoácidos do domínio variável da cadeia β dos seguintes clones de TCR A6:

38a - selvagem, 38b - Clone 134, 38c - Clone 89, 38d - Clone 1 e 38e - Clone 111

Os resíduos mutados estão apresentados a negrito, os resíduos entre parêntesis são resíduos alternativos que podem estar presentes num determinado local.

Figuras 39A e 39BI

Exemplo 1 - Desenho das sequências iniciadoras e mutagéneses das cadeias α e β para introduzir os resíduos de cisteína necessários para a formação de uma nova ligação dissulfureto entre cadeias

Para a mutação da treonina 48 de A6 Tax do exão 1 em TRAC*01 para cisteína, foram projectadas as sequências

iniciadoras que se seguem (a mutação está apresentada em minúsculas):

5'-C ACA GAC AAA tgT GTG CTA GAC AT (**SEQ ID 35**)

5'-AT GTC TAG CAC Aca TTT GTC TGT G (**SEQ ID 36**)

Para a mutação da serina 57 do exão 1 de A6 Tax em TRBC1*01 e TRBC2*01 para cisteína, foram desenhadas as seguintes sequências iniciadoras (a mutação está apresentada em minúsculas):

5'-C ACT GGG GTC Tgc aca gac cc (**SEQ ID 37**)

5'-GG GTC TGT Gca GAC CCC ACT G (**SEQ ID 38**)

Mutagénesse por PCR:

Os plasmídeos de expressão contendo os genes para a cadeia α ou β de TCR Tax A6 foram mutados usando as sequências iniciadoras da cadeia α ou da cadeia β , respectivamente, como se segue. 100 ng de plasmídeo foi misturado com 5 μ l de dNTPs 10 mM, 25 μ l de tampão Pfu 10X (Stratagene) e o volume final foi ajustado a 240 μ l com H₂O. 48 μ l desta mistura foi suplementado com sequências iniciadoras diluídas para dar uma concentração final de 0,2 μ M em 50 μ l de volume final de reacção. Após um passo inicial de desnaturação de 30 segundos a 95°C, a mistura de reacção foi sujeita a 15 ciclos de desnaturação (95°C, 30 seg.), emparelhamento (55°C, 60 seg) e elongação (73°C, 8 min) num aparelho de PCR Hybaid PCR Express (New England

Biolabs). 10 μ l da reacção digerida foi usada para transformar bactérias XL1-Blue competentes e crescidas durante 18 horas a 37°C. Uma colónia isolada foi picada e crescida durante a noite em 5 ml de TYP + ampicilina (16 g/l de Bacto-Tryptona, 16 g/l de Extracto de levedura, 5 g/l de NaCl, 2,5 g/l de K₂HPO₄, 100 mg/l de ampicilina). O DNA do plasmídeo foi purificado numa coluna de minipreparação Qiagen de acordo com as instruções do fabricante e a sequência foi verificada por sequenciação automática no serviço de sequenciação do Departamento de Bioquímica, Universidade de Oxford. As sequências do ácido nucleico e de aminoácidos mutados estão apresentadas, respectivamente, nas Figuras 1a e 2a para a cadeia α e Figuras 1b e 2b para a cadeia β .

Exemplo 2 - Construção de vectores de apresentação fágica e clonagem das cadeias TCR A6 α e β nos vectores fagemídeos

Para apresentar um TCR A6 heterodimérico contendo uma ligação dissulfureto não nativa entre cadeias em partículas fágicas filamentosas, foram construídos vectores fagemídeos para a expressão de proteínas de fusão compreendendo o TCR A6 heterodimérico contendo uma ligação dissulfureto não nativa entre cadeias com uma proteína externa do fago. Estes vectores possuem uma origem de pUC19, uma origem M13, um gene *bla* (resistência à ampicilina⁹, o promotor/operador Lac e um local de ligação

a CAP. O desenho destes vectores está descrito na Figura 3, a qual descreve vectores codificadores de ambas as proteínas de fusão, cadeia β de TCR A6-gp3 ou cadeia β de TCR A6-gp8, para além da cadeia α de TCR A6 solúvel. Os vectores de expressão contendo as sequências de DNA das cadeias α e β de TCR A6 mutadas incluindo os resíduos de cisteína adicionais necessários à formação de uma nova ligação dissulfureto entre cadeias, preparados no Exemplo 1 e mostrados nas figuras 1a e 1b foram usados como fonte das cadeias α e β de TCR A6 para a produção de um fagemídeo codificador deste TCR. A sequência de DNA completa da construção do fagemídeo (pEX746) utilizado está apresentada na Figura 4.

Os métodos de clonagem molecular para a construção dos vectores estão descritos em "Molecular Cloning: A laboratory manual", de J. Sambrook and D.W. Russell.

As sequências iniciadoras apresentadas na lista da Tabela 1 foram usadas para a construção dos vectores. Um exemplo do programa de PCR é 1 ciclo de 94°C durante 2 minutos, seguido de 25 ciclos de 94°C durante 5 segundos, 53°C durante 5 segundos e 72°C durante 90 segundos, seguido de 1 ciclo de 72°C durante 10 minutos e, depois, mantido a 4°C. A DNA polimerase Taq de alta fidelidade Expand foi adquirida à Roche.

Tabela 1. Sequências iniciadoras usadas na construção dos vectores de apresentação fágica de TCR A6

Nome da sequência iniciadora	Sequência 5' para 3'
YOL1	TAATAATACGTATAATAATATTCTATTTCAAG GAGACAGTC (SEQ ID 39)
YOL2	CAATCCAGCGGCTGCCGTAGGCAATAGGTATT TCATTATGACTGTCTCCTTGAAATAG (SEQ ID 40)
YOL3	CtaCGGCAGCCGCTGGATTGTTATTACTCGCG GCCCAGCCGCCATGGCCcag (SEQ ID 41)
YOL4	GTTCTGCTCCACTTCCTTCTGGGCCATGGCCG GCTGGGCCG (SEQ ID 42)
YOL5	CAGAAGGAAGTGGAGCAGAAC (SEQ ID 43)
YOL6	CTTCTTAAAGAATTCTTAATTAACCTAGGTTA TTAGGAACFTTCTGGGCTGGGGAAG (SEQ ID 44)
YOL7	GTTAATTAAGAATTCTTTAAGAAGGAGATATA CATATGAAAAAATTATTATTTCGCAATTC (SEQ ID 45)
YOL8	CGCGCTGTGAGAATAGAAAGGAACAATAAG GAATTGCGAATAATAATTTTTTCATATG (SEQ ID 46)
YOL9	CTTCTATTCTCACAGCGCGCAGGCTGGTGTG ACTCAGAC (SEQ ID 47)
YOL10	ATGATGTCTAGATGCGGCCGCGTCTGCTCTAC CCCAGGCCTC (SEQ ID 48)
YOL11	GCATCTAGACATCATCACCATCATCACTAGAC TGTTGAAAGTTGTTTAGCAAAAC (SEQ ID 49)
YOL12	CTAGAGGGTACCTTATTAAGACTCCTTATTAC GCAGTATG (SEQ ID 50)

Exemplo 3 - Expressão bacteriana de fusões da proteína externa e TCR A6 heterodimérico em E. coli

De forma a validar a construção preparada no Exemplo 2, prepararam-se partículas fágicas apresentadoras de TCR A6 heterodimérico, contendo uma ligação dissulfureto não nativa entre cadeias, usando os métodos anteriormente descritos para a geração de partículas fágicas apresentadoras de anticorpo scFvs (Li et al., 2000, Journal of Immunological Methods 236:133-146) com as modificações que se seguem. Células *E. coli* XL-1-Blue contendo o fagemídeo pEX746:A6 (*i.e.* o fagemídeo codificador da cadeia α de TCR A6 solúvel e da cadeia β de TCR A6 fundida com a proteína gIII fágica produzido como descrito no Exemplo 2) foram usados para inocular 5 ml de Lbatg (meio líquido Lennox L contendo 100 $\mu\text{g/ml}$ de ampicilina, 12,5 $\mu\text{g/ml}$ de tetraciclina e 2% de glucose) e depois a cultura foi incubada, com agitação, a 37°C durante a noite (16 horas). 50 μl da cultura crescida durante a noite foi usado para inocular 5 ml de TYPatg (TYP é 16 g/l de extracto de levedura, 5 g/l de NaCl e 2,5 g/l de K_2HPO_4) e em seguida a cultura foi incubada com agitação a 37°C até uma $\text{DO}_{600\text{nm}} = 0,8$. O fago auxiliar M13 K07 foi adicionado à cultura para a concentração final de 5×10^9 pfu/ml. A cultura foi então incubada a 37°C, em repouso durante trinta minutos e depois com agitação a 200 rpm durante mais 30 minutos. O meio da cultura atrás referida foi então mudado para TYPak (TYP contendo 100 $\mu\text{g/ml}$ de ampicilina, 25 $\mu\text{g/ml}$ de canamicina),

a cultura foi depois incubada a 25°C com agitação a 250 rpm durante 36 a 48 horas. A cultura foi então centrifugada a 4°C durante 30 minutos a 4000 rpm. O sobrenadante foi filtrado através de um filtro de seringa de 0,45 µm e guardado a 4°C para posterior concentração ou análise.

A proteína de fusão da proteína externa filamentosa e TCR A6 heterodimérico contendo uma ligação dissulfureto entre cadeias não nativa foi detectada no sobrenadante através de transferências Western. Aproximadamente 10^{11} cfu partículas fágicas foram aplicadas em cada uma das pistas de um gel de SDS-PAGE em tampão de aplicação redutor e não redutor. As proteínas separadas foram testadas com anticorpos primários usando o mAb anti-gIII de M13, seguido de um anticorpo secundário conjugado com peroxidase de rábano silvestre (HRP). A actividade de HRP foi então detectada com o kit de substrato Opti-4CN da Bio-Rad (Figura 5). Estes dados indicaram que TCR A6 ligado por dissulfureto do clone 1 está fundido com a proteína externa do fago filamentoso, proteína gIII.

Exemplo 4 - Detecção de TCR A6 heterodimérico funcional com uma ligação dissulfureto não nativa entre cadeias nas partículas fágicas filamentosas

A presença de TCR A6 funcional (ligação a HLA-A2-tax), apresentado nas partículas fágicas, foi detectada usando um método de ELISA para fagos.

ELISA para TCR-fago

A ligação das partículas fágicas apresentadoras de TCR A6 a péptido-MHC imobilizado em ELISA foi detectada com anti-soros de coelho anti-fd (Sigma) seguido de mAb anti-coelho conjugado com fosfatase alcalina (AP) (Sigma). Os locais de ligação não específica de proteínas nas placas podem ser bloqueados com 2% MPBS ou 3% BSA-PBS.

Materiais e reagentes

1. Tampão de revestimento, PBS
2. PBS: 138 mM NaCl, 2,7 mM KCl 10 mM Na₂HPO₄ 2 mM KH₂PO₄
3. MPBS, 3% Marvel-PBS
4. PBS-Tween: PBS, 0,1% Tween-20
5. Solução de substrato, Sigma FAST pNPP, Cat# N2770

Método

1. Lavar com PBS, duas vezes, das placas revestidas com NeutrAvidin.

2. Adicionar 25 μ l de biotina-HLA-A2 Tax ou biotina-HLA-A2 NYESO em PBS numa concentração de 10 μ g/ml e incubar à temperatura ambiente durante 30 a 60 min.

3. Lavar os alvéolos duas vezes com PBS.

4. Adicionar 300 μ l de 3% Marvel-PBS e incubar à temperatura ambiente durante 1 hr. Misturar a suspensão de TCR-fago com 1 volume de 3% Marvel-PBS e incubar à temperatura ambiente.

5. Lavar os alvéolos duas vezes com PBS.

6. Adicionar 25 μ l da mistura de fago-TCR/Marvel-PBS, incubar em gelo durante 1 hr.

7. Lavar os alvéolos três vezes com PBS-Tween arrefecido em gelo e três vezes com PBS arrefecido em gelo.

8. Adicionar 25 μ l de anticorpo de coelho anti-fd arrefecido em gelo diluído a 1:1000 em Marvel-PBS e incubar em gelo durante 1 hr.

9. Lavar os alvéolos três vezes com PBS-Tween arrefecido em gelo e três vezes com PBS arrefecido em gelo.

10. Adicionar 25 μ l de conjugado mAb anti-coelho-

Ap diluído a 1:50000 em Marvel-PBS e incubar em gelo durante 1 hr.

11. Lavar os alvéolos três vezes com PBS Tween arrefecido em gelo e três vezes com PBS arrefecido em gelo.

12. Adicionar 150 μ l de substrato de fosfatase alcalina a cada alvéolo e ler o sinal a 405 nm.

Os resultados apresentados na Figura 6 indicam que o clone 1 produziu uma partícula fágica apresentadora de um TCR A6 que se pode ligar especificamente ao seu pMHC cognato (HLA-A2 Tax).

A análise da sequência de DNA deste TCR A6 apresentado revelou a presença de um codão de paragem opal na cadeia TCR β não presente na sequência correspondente da construção do vector de expressão do Exemplo 2. Este codão é ignorado com baixa frequência pelos ribossomas da estirpe de *E. coli* utilizada, resultando na inserção de um resíduo triptofano neste local e num nível global muito mais reduzido de expressão da cadeia β de tamanho completo. A partir desta observação foi inferido que apenas as células que expressam esta sequência TCR A6 mutada sobreviveram os ciclos de cultura do Exemplo 3 e que portanto os elevados níveis de TCR A6 que se previu serem expressos pelo vector de expressão original foram tóxicos para as células hospedeiras.

Exemplo 5 - Apresentação por ribossomas de TCR de cadeia simples (scTCR)

Construção de vectores de apresentação de scTCR por ribossomas para usar na geração de matrizes de PCR para apresentação por ribossomas.

As construções de apresentação por ribossomas foram clonadas no DNA do plasmídeo pUC19, facilmente disponível, de forma a gerar uma matriz de DNA para PCR sem erros e estável, a partir da qual seja possível realizar subsequentes experiências de apresentação por ribossomas. A construção de vectores foi efectuada em dois passos de forma a evitar a utilização de grandes sequências iniciadoras oligonucleotídicas (com os seus problemas de erros associados). A construção final de apresentação por ribossomas do DNA de scTCR-C-kappa está apresentada numa forma esquemática na figura 7a e as sequências de DNA e da proteína estão apresentadas na Figura 7b. Esta construção pode ser removida de pUC19 como produto de digestão dupla com PstI/EcoRI.

Os métodos de clonagem molecular para construção dos vectores estão descritos em "Molecular cloning: A Laboratory manual", de J. Sambrook e D.W. Russell. As sequências iniciadoras apresentadas na Tabela 2 foram usadas na construção dos vectores. O programa de PCR utilizado foi como se segue - 1 ciclo de 94°C durante 2 minutos, seguido de 25 ciclos de 94°C durante 30 segundos,

55°C durante 20 segundos e 72°C durante 120 segundos, seguido de 1 ciclo de 72°C durante 5 minutos e depois mantido a 4°C. A DNA polimerase Pfu foi adquirida à Stratagene. As sequências iniciadoras oligonucleotídicas usadas estão descritas na Tabela 2.

Construção de pUC19-T7 - Passo 1

A construção de pUC19-T7 está descrita abaixo, a construção resulta num vector pUC19 contendo uma região do promotor de T7 seguida de uma curta região espaçadora e da sequência óptima eucariótica Kozak. Esta é uma parte essencial da construção apresentada pelo ribossoma, uma vez que é necessária para a iniciação da transcrição de qualquer sequência ligada em lisados de reticulócitos de coelho. As sequências para apresentação por ribossomas, tais como a A6scTCR-C-kappa, podem ser ligadas ao vector pUC19-T7 entre os locais de restrição NcoI e EcoRI.

Quantidades equimolares da sequência iniciadora Rev-link e For-link foram emparelhadas através do aquecimento a 94°C durante 10 minutos e arrefecimento lento da reacção até à temperatura ambiente. Isto resulta na formação de um complexo de DNA de cadeia dupla que pode ser observado abaixo.

5' AGCTGCAGCTAATACGACTCACTATAGGAACAGGCCACCATGG
CGTCCATTATGCTGAGTGATATCCTTGTCCGGTGGTACCCTAG 3'
(SEQ ID 51)

A região 5' possui um extremo coesivo saliente complementar de um local de restrição HindIII, enquanto que o extremo 3' possui um extremo coesivo que é complementar de um local de restrição BamHI.

Os oligonucleótidos emparelhados foram ligados a pUC19 duplamente digerido com HindIII/BamHI, o qual tinha sido purificado por electroforese em gel de agarose, recuperado e ainda purificado com um kit Qiagen para extracção do gel. As ligações foram usadas para transformar *E. coli* XL1-BLUE. Os clones pUC19-T7 individuais foram sequenciados para confirmar a presença da sequência correcta. A sequência está apresentada na Figura 8.

Construção do vector A6scTCR-C-kappa - Passo 2

A construção da sequência de DNA de cadeia simples A6scTCR-C-kappa requer a geração de três fragmentos de PCR que devem ser montados num fragmento A6scTCR-C-kappa. Os fragmentos consistem em (a.) A região variável da cadeia TCR A6 alfa flanqueada por um local NcoI na região 5' e uma segmento do elemento de ligação Glicina Serina na região 3' flanqueado por um local de restrição BamHI. Este produto foi gerado através de um PCR convencional do vector pEX202 com as sequências iniciadoras 45 e 50 (ver Tabela 2). Fragmento (b.) A região variável e a região constante de TCR A6 beta flanqueadas por um local de restrição BamHI na

região 5' seguido de um segmento do elemento de ligação Glicina Serina. Este produto foi gerado através de um PCR convencional do vector pEX207 com as sequências iniciadoras 72 e 73 (Ver Tabela 2). Fragmento (c.) Porção de uma região C-kappa humana gerada por um PCR convencional do vector p147 com as sequências iniciadoras 61-60 (Ver Tabela 2). Todos os produtos de PCR foram corridos num gel de 1,6% de agarose e as bandas de DNA do tamanho correcto cortadas e purificadas usando o kit Qiagen de extracção do gel. Os fragmentos (b.) e (c.) foram fundidos por um PCR de sobreponível convencional através da complementaridade nas suas sequências iniciadoras 73 e 61 (Ver Tabela 2). Os produtos de PCR foram corridos num gel de 1,6% de agarose TBE e as bandas de DNA do tamanho correcto removidas e purificadas usando o kit Qiagen de extracção do gel. Este fragmento é designado (d.). O fragmento (a.) foi duplamente digerido com NcoI e BamHI, enquanto o fragmento (d.) foi duplamente digerido com BamHI e EcoRI. pUC19-T7 foi duplamente digerido com NcoI e EcoRI. Todos os produtos de DNA digeridos foram corridos num gel de 1,2% de agarose TBE e as bandas de DNA do tamanho correcto foram removidas e purificadas usando o kit Qiagen de extracção do gel. O pUC19-T7 digerido, fragmentos (a.) e (d.) foram ligados e usados para transformar *E. coli* XL1-BLUE. Os transformantes foram sequenciados para confirmar a sequência correcta. A sequência da construção de apresentação pelo ribossoma de A6scTCR-C-kappa, clonada em pUC19, está apresentada na Figura 9 flanqueada pelos locais PstI e EcoRI.

Tabela 2

Oligonucleótidos usados (adquiridos à MWG).	
Rev-Link	5'GATCCCATGGTGGCCTGTTCCCTATAGTGAGTCGTATTAGCTGC (SEQ ID 52)
For-Link	5'AGCTGCAGCTAATACGACTCACTATAGGAACAGGCCACCATGG (SEQ ID 53)
45-A6	5'CCACCATGGGCCAGAAGGAAGTGGAGCAGAACTC (SEQ ID 54)
7a6-beta (RTPCR(a))	5'CGAGAGCCCGTAGAACTGGACTTG (SEQ ID 55)
49-A6- BamHI-F	5'GTGGATCCGGCGGTGGCGGGTCCAACGCTGGTGTCA CTCAGACCCC (SEQ ID 56)
50-A6- BamHI-R	5'CCGGATCCACCTCCGCCTGAACCGCCTCCACCGGTGACCACAAC CTGGGTCCCTG (SEQ ID 57)
60-kappa- rev-EcoRI	5' CTGAGAATTCTTATGACTCTCCGCGGTTGAAGCTC (SEQ ID 58)
61-Betac- kappa-for 1	5' TGACGAATTCTGACTCTCCGCGGTTGAAGCTC (SEQ ID 59)
71 primer T7	5' AGCTGCAGCTAATACGACTCACTATAGG (SEQ ID 60)
72 A6-beta	5'GGCCACCATGGGCAACGCTGGTGTCACTCAGACCCC (SEQ ID 61)
73-A6- cons-rev	5'TGAACCGCCTCCACCGTCTGCTCTACCCCAGGCCTCGGCG (SEQ ID 62)
75-kappa- rev	5' TGA CTCTCCGCGGTTGAAGCTC (SEQ ID 63)

Demonstração da produção de sc A6 TRC-C-kappa por transcrição e tradução in vitro.

Preparação do produto de PCR sc A6 TCR-C-kappa para transcrição e tradução in vitro.

Descrevemos aqui a síntese de sc A6 TCR-C-kappa

através de transcrição e tradução *in vitro* na presença de lisina biotinilada e sua subsequente detecção em transferências Western e detecção com estreptavidina marcada com fosfatase alcalina.

O produto de PCR sc A6 TCR-C-kappa foi preparado numa reacção de PCR convencional usando o vector sc A6 TCR-C-kappa como matriz e as sequências iniciadoras de PCR 71 e 60. A sequência iniciadora 60 possui um codão de paragem para permitir a libertação da scTCR do ribossoma. A polimerase Pfu (Stratagene) foi usada para aumentar a fidelidade durante a síntese de PCR. Os produtos de PCR foram corridos num gel de 1,6% de agarose TBE e as bandas do tamanho correcto foram removidas do gel e purificadas usando o kit Qiagen de extracção do gel. As reacções de transcrição e tradução foram realizadas usando o kit de transcrição e tradução associadas Ambion PROTEINscript II Cat 1280-1287 com 300 ng do produto de PCR atrás descrito. Foram efectuadas três reacções de transcrição tradução de acordo com o protocolo do fabricante. A única modificação foi a adição de lisina biotinilada do sistema de detecção de tradução não radioactivo Transcend™ Non-Radioactive Translation Detection System.

Reacção 1 sc A6 TCR-C-kappa, 300 ng com 2 µl de lisina biotinilada

Reacção 2 sc A6 TCR-C-kappa, 300 ng sem 2 µl de lisina biotinilada

Reacção 3 sem DNA, controlo com 2 µl de lisina biotinilada.

Dois microlitros de cada uma das reacções foram corridos num gel de SDS-PAGE de gradiente 4-20% Novex (Invitrogen). Ainda, foi também corrido uma série de diluições de um TCR controlo biotinilado. O gel foi transferido e as proteínas detectadas com estreptavidina fosfatase alcalina e subsequentemente revelado colorimetricamente com Western Blue® Stabilized Substrate para fosfatase alcalina, como descrito no protocolo do sistema de detecção da tradução não radioactivo Transcend™. A transferência Western está apresentada na Figura 10.

No controlo sem DNA e na reacção de A6 sc TCR-C-kappa na ausência de lisina biotinilada, conforme esperado, não pode ser observada uma banda com o tamanho aproximado do tamanho correcto, enquanto na reacção A6scTCR-C-kappa na presença de lisina biotinilada pode ser observada uma banda com aproximadamente o tamanho correcto. Isto demonstra a síntese da scA6 TCR-C-kappa através da transcrição tradução *in vitro*.

Preparação do produto de PCR para apresentação por ribossomas de sc A6 TCR-C-kappa

O produto de PCR sc A6 TCR-C-kappa foi preparado numa reacção de PCR convencional usando o vector A6scTCR-C-kappa como matriz e as sequências iniciadoras de PCR 71 e 75 (Ver Tabela 2). A sequência iniciadora 75 não possui um codão de paragem. A polimerase Pfu (Stratagene) foi usada para aumentar a fidelidade durante a síntese de PCR. Os

produtos de PCR foram corridos num gel de 1,6% de TBE-agarose e as bandas de DNA do tamanho correcto foram removidas do gel e purificadas com o kit Qiagen de extracção do gel.

Processo de apresentação por ribossomas

Transcrição e tradução de sc A6 TCR-C-kappa

As reacções de transcrição/tradução foram realizadas com o kit de transcrição tradução Ambion PROTEINscript II (Cat N° 1280-1287).

Reacções de transcrição

Efectuaram-se as reacções de transcrição que se seguem em tubos não aderentes Ambion de 0,5 ml (Cat N° 12350).

Conteúdo	Tubo 1 (A6 normal)	Tubo 2 (Controlo)
Água	4,53 µl	5,7 µl
Matriz (produto de PCR)	Produto de PCR Sc A6 TCR-C-kappa 1,17 µl (300 ng)	Sem DNA
Mix de transcrição 5X	2 µl	2 µl
Mistura de enzimas	2 µl	2 µl
Supersasin _{inibidor} de RNase ambion	0,3 µl	0,3 µl
Volume final	10 µl	10 µl

Os tubos foram incubados a 30°C, durante 60 min, num bloco de PCR sem a tampa de aquecimento.

Reacções de tradução

As reacções de tradução que se seguem foram efectuadas em tubos não aderentes de 0,5 ml Ambion.

Conteúdo	1 (A6 normal)	2 (Controlo)
Lisado de reticulócitos	105 µl	105 µl
Mg-Acetato 25 mM	3 µl	3 µl
Mistura de tradução	7,5 µl	7,5 µl
Metionina	7,5 µl	7,5 µl
Água	18 µl	18 µl
Supersasin _{inibidor de RNase ambion}	0,3 µl	0,3µl
Reacção de transcrição	6 µl do tubo 1 atrás	6 µl do tubo 2 atrás

Cada um dos tubos contém o suficiente para 3 x 50 µl selecções. O conteúdo dos tubos foi misturado e incubado a 30°C, durante 60 min, num bloco de PCR sem a tampa de aquecimento. Após 30 min, adicionou-se 3 Unidades de DNase RQ1 sem RNase (Promega), para destruir a matriz de DNA original, ao tubo 1 e 3 Unidades de DNase RQ1 sem RNase (Promega) ao tubo 2. Após 60 min, 18 µl de solução de Heparina foram adicionados à reacção de tradução 2 e 18 µl

de solução de Heparina foram adicionados à reacção de tradução 1. As amostras foram guardadas em gelo até serem usadas na selecção com esferas revestidas com HLA.

Revestimento das esferas magnéticas.

20 µl de partículas magnéticas com estreptavidina ressuspensas (Roche Cat: N° 1641778) foram transferidas para um tubo eppendorf de 1,5 ml, estéril, sem RNases. As esferas foram imobilizadas com um separador de partículas magnéticas (Roche Cat. N° 1641794) e o sobrenadante foi removido. As esferas foram então lavadas com 100 µl de PBS 1X sem RNase (10X PBS Ambion Cat. N° 9624, H₂O Ambion Cat. 9930) as esferas foram imobilizadas e o sobrenadante foi removido. Fez-se um total de 3 lavagens com PBS.

As esferas foram ressuspensas em 20 µl de PBS eo conteúdo dividido por dois tubos (10 µl cada). Um tubo seria usado para produzir esferas bloqueadas controlo e os outros tubos para produzir esferas revestidas com HLA-A2-Tax.

Às esferas do tubo controlo adicionou-se 80 µl de SA/Biotina e misturou-se. A solução de BSA/Biotina foi preparada como se segue. 10 µl de uma solução de base Tris 0,2M e biotina 0,1M foram adicionados a 990 µl de PBS 0,1% BSA (Ambion Ultrapure Cat. N° 2616). Igualmente, 20 µl de solução de Heparina (138 mg/ml de Heparina (Sigma H-3393)

em PBS 1X) foram adicionados e a solução misturada. As esferas foram incubadas à temperatura ambiente durante 1 hora com mistura intermitente. As esferas foram então lavadas três vezes com 100 µl de PBS e foram ressuspensas em 10 µl de PBS, 0,1% BSA.

As esferas revestidas com HLA-Tax foram preparadas como se segue. 40 µl de HLA-A2-Tax (1,15 mg/ml preparado como descrito em W099/60120) foram adicionados aos 10 µl de esferas e misturados. As esferas foram incubadas à temperatura ambiente durante 15 min e depois adicionou-se 20 µl de BSA 50 mg/ml Ambion Cat 2616 e 20 µl de solução de heparina (ver atrás) e misturou-se. As esferas foram incubadas durante mais 45 min e depois adicionou-se 20 µl da solução de BSA/Biotina. As esferas foram então lavadas três vezes com 100 µl de PBS e foram ressuspensas em 10 µl de PBS, 0,1% BSA.

Seleção com esferas magnéticas

A reacção de tradução de sc A6 TCR foi dividida em três alíquotas de 50 µl e cada uma das alíquotas recebeu 2 µl das seguintes esferas:

Controlo (sem HLA)

HLA-A2-Tax

HLA-A2-Tax mais 10 µg de scA6 TCR solúvel

Realizou-se uma reacção de controlo da tradução e dividiu-se em três alíquotas de 50 µl e cada uma das alíquotas recebeu 2 µl das seguintes esferas:

Controlo (sem HLA)

HLA-A2-Tax

HLA-A2-Tax mais 10 µg de sc A6 TCR solúvel

Isto deu um total de seis tubos. Os tubos foram incubados num bloco de PCR a 5°C, durante 60 minutos, com mistura intermitente. As esferas foram então lavadas três vezes com 100 µl de tampão arrefecido em gelo (PBS, Mg-acetato 5 mM, 0,2% Tween 20 (Sigma, sem RNase). Cada uma das alíquotas de esferas foi então ressuspensa em 50 µl de tampão de digestão com DNase RQ1 1X contendo 1 µl (40 U) de Supersasin e 1 µl (1U) de DNase RQ1. As esferas foram incubadas num bloco de PCR durante 30 min a 30°C. As esferas foram então lavadas três vezes com 100 µl de tampão arrefecido em gelo (PBS, Mg-acetato 5 mM, 0,2% Tween 20) e uma vez com H₂O arrefecida em gelo. As esferas foram ressuspensas em 10 µl de H₂O sem RNase. As esferas foram então congeladas até serem usadas em RT-PCR.

RT-PCR de mRNA de sc A6 TCR-C-kappa em esferas recuperadas das reacções de apresentação de ribossomas.

As reacções de RT-PCR nas esferas foram realizadas usando o kit Titan One Tube RT-PCR cat 1855476 como descrito no protocolo do fabricante. Dois microlitros

das esferas foram adicionados a cada uma das reacções de RT-PCR juntamente com as sequências iniciadoras 45 e 7 e 0,3 µl de inibidor de RNase Supersin. Para cada reacção de RT-PCR realizou-se apenas uma segunda reacção de PCR que diferiu apenas no facto de não estar presente transcriptase reversa, mas apenas polimerase de alta fidelidade da Roche. Esta segunda reacção serviu como controlo para a contaminação com DNA. Ainda, realizou-se um controlo positivo de RT-PCR usando 1 ng do vector sc A6 TCR-C-kappa.

As reacções foram sujeitas a ciclos térmicos como se segue. Um passo de RT-PCR foi realizado por incubação das amostras a 50°C, durante 30 min, seguido da inactivação da transcriptase reversa por incubação a 94°C, durante 3 min, num bloco de PCR. As reacções foram sujeitas a ciclos de PCR, num total de 38 ciclos:

94°C 30 segundos

55°C 20 segundos

68°C 130 segundos.

A reacção de PCR foi terminada por incubação a 72°C durante 4 minutos.

Foi tido cuidado durante os passos de apresentação por ribossomas para evitar a contaminação com RNases. As reacções de RT-PCR e PCR foram corridas num gel de 1,6% de agarose TBE, que pode ser observado na Figura

11. A análise do gel mostra que não existe contaminação e que todos os produtos de PCR são derivados do mRNA. A banda de DNA do tamanho correcto na pista 2 demonstra que os ribossomas que apresentam sc A6 TCR-C-kappa foram seleccionados das esferas revestidas com HLA-A2-Tax. A pista 3 mostra que podemos inibir esta selecção específica de sc A6 TCR-C-kappa apresentado por ribossomas pela adição de scA6 TCR solúvel. A redução significativa na intensidade da banda na pista 3 relativamente à amostra não inibida na pista 2 demonstra isto. Não se pode mostrar ligação de scA6 TCR-C-kappa apresentada por ribossomas ao controlo de esferas não revestidas com HLA.

Exemplo 6 - Análise das sequências dos clones de TCR A6 apresentados em partículas fágicas e métodos para melhorar as características da apresentação.

Após a construção de vectores para a apresentação de TCR A6 em fagos por PCR e clonagem molecular, os clones bacterianos que podem produzir partículas fágicas apresentadoras de TCR A6 foram testados por ELISA de fagos como descrito no Exemplo 4. Foram identificados três clones diferentes que deram ligação específica a HLA-A2-tax no ensaio de ligação de ELISA. Estes clones continham todos mutações no DNA de TCR A6 selvagem ou nas sequências reguladoras associadas, as quais estão descritas na tabela que se segue:

Clones funcionais do rastreio de TCR A6 apresentado em
fagos

Nome	Característica	
Clone 7	O terceiro local de ligação ao ribossoma, que está situado antes do gene $v\beta$, está mutado de <u>AAGGAGA</u> para AAGGGGA.	
Clone 9	Um codão opal foi introduzido em CDR3 de $v\beta$.	Sequência completa de DNA e de aminoácidos nas Figuras 12a & 12b
Clone 49	Um codão amber foi introduzido em FR1 de $v\beta$. Esta mutação introduz uma mutação silenciosa que não afecta a sequência de aminoácidos resultante	Sequência completa de DNA na Figuras 13a

Estes clones continham todos mutações que provavelmente causam uma redução nos níveis de expressão da cadeia β de TCR A6. Foi inferido que os clones com baixo nível de expressão foram seleccionados relativamente aos clones de elevado nível de expressão como resultado de toxicidade celular causado pelos níveis elevados de expressão de TCR.

Exemplo 7 Mutagénese das regiões CDR3 de TCR A6

Nas regiões CDR3 de TCR A6 foram introduzidas mutações para estudar a possibilidade de gerar mutantes de elevada afinidade.

O PCR sobreponível foi usado para modificar a

O PCR foi usado para introduzir mutações para maturação da afinidade. O clone 9 de TCR A6 (incorporação de um codão opal introduzido na sequência CDR3 da cadeia β) foi usado como fonte de DNA matriz e as cadeias de TCR foram amplificadas com as sequências iniciadoras das mutações (detalhado na tabela que se segue) e YOL22 5'CATTTTCAGGGATAGCAAGC3' (SEQ ID 68) (cadeia β) ou YOL13 5'TCACACAGGAAACAGCTATG3' (SEQ ID 69) (cadeias α).

Sequências iniciadoras para introdução de mutações em CDR3
da cadeia β e da cadeia α

Nome da sequência iniciadora	Sequências 5' a 3'
YOL59	TGTGCCTCGAGGNNKNNKNNKNNKNNKNNKNNK NKKCGACCAGAGCAGTACTTCG (SEQ ID 70)
YOL60	TGTGCCTCGAGGCCGNKNKNKNKNKNKNKNKN NKNNKCCAGAGCAGTACTTCGGGC (SEQ ID 71)
YOL61	TGTGCCTCGAGGCCGNKNKNKNKNKNKNKNKN NKNNKCGACCAGAGCAGTACTTCG (SEQ ID 72)

(continuação)

Nome da sequência iniciadora	Sequências 5' a 3'
YOL62	TGTGCCTCGAGGCCGNNKNNKNNKNNKNNKNNKNNKGGAGGGCGACCAGAGCAG (SEQ ID 73)
YOL63	TGTGCCTCGAGGCCGGGANNKNNKNNKNNKNNKNNKNNKNNKGGGCGACCAGAGCAGTAC (SEQ ID 74)
YOL68	TGTGCCTCGAGGNNKNNKNNKNNKNNKNNKNNKNNKNNKCCAGAGCAGTACTTCGggc (SEQ ID 75)
YOL69	TGTGCCTCGAGGNNKNNKNNKNNKNNKNNKNNKNNKNNKNGAGCAGTACTTCGggccg (SEQ ID 76)
YOL70	TGTGCCTCGAGGNNKNNKNNKNNKNNKNNKNNKNNKNNKNCAGTACTTCGggccgggc (SEQ ID 77)
YOL71	TGTGCCTCGAGGccgNNKNNKNNKNNKNNKNNKNNKggCGACCAGAGCAGTACTTCG (SEQ ID 78)
YOL58	AAACTGAAGCTTMNNMNNMNNMNNMNNMNNNT GTAACGGCACAGAGGTAG (SEQ ID 79)
YOL72	AAACTGAAGCTTMNNMNNgctgtcMNNNT GTAACGGCACAGAGGTAG (SEQ ID 80)
YOL73	AAACTGAAGCTTMNNMNNMNNMNNgctgtcM NNTGTAACGGCACAGAGGTAG (SEQ ID 81)
YOL74	AAACTGAAGCTTMNNMNNMNNgctgtcMNNNA ACGGCACAGAGGTAG (SEQ ID 82)

Os fragmentos da cadeia α foram digeridos com NcoI e HindIII, novamente purificados usando um kit Qiagen, e o vector foi preparado através da digestão do clone 9 com NcoI e HindIII, seguido de purificação em gel usando um kit Qiagen. Os fragmentos da cadeia β foram digeridos com XhoI e NotI e novamente purificados usando um kit Qiagen e o

vector foi preparado por digestão do clone 9 com XhoI e NotI, seguido de purificação em gel usando um kit Qiagen. Os insertos purificados e os vectores, numa proporção molar de 3:1, foram misturados com tampão ligase de T4, ligase de T4 e água sem nuclease. As ligações foram realizadas num banho-maria a 16°C, durante a noite. Para cada biblioteca de mutações, um total de 0,5 a 1 µg dos produtos ligados purificados foram electroporados em *E. coli* TG1 numa proporção de 0,2 µg de DNA por 40 µl de células competentes para electroporação (Stratagene), seguindo os protocolos proporcionados pelo fabricante. Após electroporação, as células foram ressuspensas imediatamente em 960 µl de meio SOC a 37°C e semeadas numa placa de cultura de tecidos de 244 mm x 240 mm com YTE (15 g de Bacto-Agar, 8 g de NaCl, 10 g de Triptona, 5 g de Extracto de levedura em 1 litro) suplementado com 100 µg/ml de ampicilina e 2% de glucose. A placa foi incubada a 30°C durante a noite. As células foram então raspadas das placas com 5 ml de DYT (16 g de Triptona, 10 g de extracto de levedura e 5 g de NaCl em 1 litro, autoclavado a 125°C durante 15 minutos) suplementado com 15% de glicerol.

Para preparar partículas fágicas apresentadoras de TCR A6, 500 ml de DYTag (DYT contendo 100 µg/ml de ampicilina e 2% de glucose) foi inoculado com 500 a 1000 µl dos stocks de bibliotecas. A cultura foi crescida até ser atingida uma DO (600 nm) de 0,5. 100 ml da cultura foi infectada com fago auxiliar (M13 K07 (Invitrogen)) ou HUPER PHAGE (Progen Biotechnik, GmbH 69123 Heidelberg) e incubada a 37°C num banho-Maria durante 30 minutos. O meio foi

substituído por 100 ml de DYTak (DYT contendo 100 µg/ml de ampicilina e 25 µg/ml de canamicina). A cultura foi então incubada com agitação a 300 rpm e 25°C, durante 20 a 36 horas.

Exemplos 8 - Isolamento de mutantes TCR A6 de elevada afinidade

O isolamento dos mutantes TCR A6 de elevada afinidade foi realizado usando dois métodos diferentes.

O primeiro método envolve a selecção de partículas fágicas apresentadoras de TCRs A6 mutantes, capazes de se ligarem ao complexo HLA-A2 Tax, usando imunotubos Maxisorp (Invitrogen). Os imunotubos foram revestidos com 1 a 2 ml de estreptavidina a 10 µg/ml em PBS, durante a noite, à temperatura ambiente. Os tubos foram lavados duas vezes com PBS e depois 1 ml do complexo HLA-A2 Tax a 5 µg/ml em PBS foi adicionado e incubado à temperatura ambiente durante 30 minutos. O resto do protocolo para selecção de ligandos de elevada afinidade foi descrito anteriormente (Li et al. (2000) Journal of Immunological Methods 236: 133-146), excepto para as modificações que se seguem. A selecção foi realizada durante três a quatro ciclos. As concentrações do complexo de HLA-A2 Tax biotinilado foram de 5 µg/ml para o primeiro ciclo de selecção, 0,5 µg/ml para o segundo, 0,05 µg/ml para o terceiro e 0,005 µg/ml para o quarto ciclo de selecção. O fago auxiliar M13 K07 foi usado nos ciclos um e dois e o hiper-fago foi usado em ciclos subsequentes para selecção.

O segundo método utilizado foi a selecção de partículas fágicas apresentadoras de TCRs mutantes capazes de se ligarem ao complexo HLA-A2 Tax em solução. As esferas paramagnéticas revestidas com estreptavidina (DynaL M280) foram pré-lavadas de acordo com os protocolos do fabricante. As partículas fágicas apresentadoras de TCR A6 mutado, numa concentração de 10^{12} a 10^{13} cfu, foram pré-misturadas com o complexo HLA-A2 Tax biotinilado em concentrações de 2×10^{-8} , 2×10^{-9} M, 2×10^{-10} M e 2×10^{-11} M respectivamente para o primeiro, segundo, terceiro e quarto ciclo de selecção. A mistura das partículas fágicas apresentadoras de TCR A6 e o complexo HLA-A2 Tax foram incubados durante uma hora, à temperatura ambiente, com rotação suave e as partículas fágicas apresentadoras de TCR A6 ligadas ao complexo HLA-A Tax biotinilado foram capturadas usando 200 μ l (ciclo 1) ou 50 μ l (ciclo 2, 3 e 4) de esferas magnéticas M280 revestidas com estreptavidina. Após captura das partículas fágicas, as esferas foram lavadas num total de dez vezes (três vezes em PBS Tween 20, duas vezes em PBS Tween 20 contendo 2% de leite em pó desnatado, duas vezes em PBS, uma vez em PBS contendo 2% de leite em pó desnatado e duas vezes em PBS) usando um concentrador de partículas magnéticas Dynal. Após a lavagem final, as esferas foram ressuspensas em 1 ml de trietilamina 100 mM pH 11,5 preparada de fresco e incubadas durante 5 a 10 minutos à temperatura ambiente com rotação suave. As partículas fágicas eluidas das esferas foram neutralizadas imediatamente com 300 μ l de Tris-HCl 1M pH 7,0. Metade do eluato foi usada para infectar 10 ml de *E.*

coli TG1 a um DO (600 nm) = 0,5, preparadas de fresco, para a amplificação das partículas fágicas seleccionadas de acordo com os métodos anteriormente descritos (Li, *et al.*, (2000) *Journal of Immunological Methods* 236:133-146).

Após o terceiro ou quarto ciclo de selecção, 95 colónias foram picadas das placas e usadas para inocular 100 µl de DYTag numa placa de microtitulação de 96 alvéolos. A cultura foi incubada a 37°C com agitação durante a noite. 100 µl de DYTag foi então sub-inoculada com 2 a 5 µl das culturas crescidas durante a noite e incubadas a 37°C com agitação durante 2 a 3 horas, ou até a cultura ficar turva. Para infectar as células com o fago auxiliar, a cultura foi infectada com 25 µl de DYTag contendo 5×10^9 pfu de fagos auxiliares e incubada a 37°C durante 30 minutos. O meio foi substituído por DYTak. As placas foram incubadas a 25°C durante 20 a 36 horas com agitação a 300 rpm. As células foram precipitadas por centrifugação a 3000 g durante 10 minutos a 4°C. Os sobrenadantes foram usados para testar mutantes de TCR A6 de elevada afinidade através de ELISA competitivo para fagos.

Alvéolos Nunc-Immuno Maxisorp revestidos com estreptavidina foram lavados duas vezes com PBS. 25 µl do complexo HLA-A2 Tax biotinilado, 5 µg/ml, foram adicionados a cada alvéolo e estes foram incubados à temperatura ambiente durante 30 a 60 minutos, seguido de duas lavagens com PBS. Os locais de ligação não específica a proteínas nos alvéolos foram bloqueados pela adição de 300 µl de

leite desnatado a 3% em PBS, seguido de incubação à temperatura ambiente durante 2 horas. Para preparar as partículas fágicas apresentadoras de TCR A6 heterodimérico, as partículas fágicas foram misturadas com 3% de leite desnatado em PBS contendo HLA-A2-Tax 0, 20 e 200 nM, seguido de incubação à temperatura ambiente durante 1 hora. Os fagos foram adicionados aos alvéolos revestidos com HLA-A2-Tax e incubados à temperatura ambiente durante 1 hora, seguido de 3 lavagens com PBS contendo 0,1% Tween 20 e depois 3 lavagens com PBS. As partículas fágicas apresentadoras de TCR ligadas foram detectadas com um anticorpo anti-fd (Sigma), como descrito no Exemplo 4.

Vários mutantes TCR A6 de elevada afinidade putativos foram identificados e as sequências CDR3 estão apresentadas nas duas tabelas que se seguem juntamente com as sequências selvagens correspondentes. Os codões de paragem amber (X) foram encontrados em todos os mutantes da cadeia β e para um mutante da cadeia α .

Mutantes da cadeia β de TCR A6

clone	Sequência CDR3
selva- gem	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGAGCAGTAG (SEQ ID 83) A S R P G L A G G R P E Q Y (SEQ ID 84)
134	GCCTCGAGGCCGGGGCTGATGAGTGCGTAGCCAGAGCAGTAC (SEQ ID 85) A S R P G L M S A X P E Q Y (SEQ ID 86)
86	GCCTCGAGGCCGGGGCTGAGGTGCGCTAGCCAGAGCAGTAC (SEQ ID 87) A S R P G L R S A X P E Q Y (SEQ ID 88)
87	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGAGGCGTAG (SEQ ID 89) A S R P G L A G G R P E A X (SEQ ID 90)
89	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGAGGATTAG (SEQ ID 91) A S R P G L A G G R P E D X (SEQ ID 92)

(continuação)

clone	Sequência CDR3
85	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGATCAGTAG (SEQ ID 93) A S R P G L A G G R P D Q X (SEQ ID 94)
83	GCCTCGAGGCCGGGTCTGTAGGCTGGGCGACCAGAGCAGTAC (SEQ ID 95) A S R P G L X A G R P E Q Y (SEQ ID 96)
1	GCCTCGAGGCCGGGGCTGGTTCCGGGGCGACCAGAGCAGTAG (SEQ ID 97) A S R P G L V P G R P E Q X (SEQ ID 98)
2	GCCTCGAGGCCGGGGCTTGTGTCTGCTTAGCCAGAGCAGTAC (SEQ ID 99) A S R P G L V S A X P E Q Y (SEQ ID 100)
111	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCACATCCGTAG (SEQ ID 101) A S R P G L A G G R P H P X (SEQ ID 102)
125	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGATGCGTAG (SEQ ID 103) A S R P G L A G G R P D A X (SEQ ID 104)
133	GCCTCGAGGCCGGGTCTGATTAGTGCTTAGCCAGAGCAGTAC (SEQ ID 105) A S R P G L I S A X P E Q Y (SEQ ID 106)

Mutantes da cadeia α de TCR A6

clone	Sequência CDR3
selvagem	GCCGTTACAACCTGACAGCTGGGGGAAGCTTCAG (SEQ ID 107) A V T T D S W G K L Q (SEQ ID - 108)
149	GCCGTTACAACCTGACAGCTGGGGGCCGCTTCAG (SEQ ID 109) A V T T D S W G P L Q (SEQ ID 110)
65	GCCGTTACAACCTGACAGCTGGGGGAAGATGCAG (SEQ ID 111) A V T T D S W G K M Q (SEQ ID 112)
66	GCCGTTACAACCTGACAGCTGGGGGAAGTTGCAT (SEQ ID 113) A V T T D S W G K L H (SEQ ID 114)
153	GCCGTTACAACCTGACAGCTGGGGGTAGCTTCAT (SEQ ID 115) A V T T D S W G X L H (SEQ ID 116)
71	GCCGTTACAACCTGACAGCTGGGGGGAGCTTCAT (SEQ ID 117) A V T T D S W G E L H (SEQ ID 118)
70	GCCGTTACAACCTGACAGCTGGGGGAGGCTGCAT (SEQ ID 119) A V T T D S W G R L H (SEQ ID 120)
121	GCCGTTACAACCTGACAGCTGGGGGCAGCTTCAT (SEQ ID 121) A V T T D S W G Q L H (SEQ ID 122)
117	GCCGTTACAACCTGACAGCTGGGGGCAGCTTCAT (SEQ ID 121) A V T T D S W G Q L H (SEQ ID 122)
72	GCCGTTACAACCTGACAGCTGGGGGAAGGTGAAT (SEQ ID 125) A V T T D S W G K V N (SEQ ID 126)
150	GCCGTTACAACCTGACAGCTGGGGGAAGCTTCTG (SEQ ID 127) A V T T D S W G K L L (SEQ ID 128)

Exemplo 9 - Produção de TCR A6 heterodimérico solúvel com ligação dissulfureto, não nativa, entre as regiões constantes com mutações em CDR3

O DNA de fagemídeo codificador de mutantes de TCR A6 de elevada afinidade identificados no Exemplo 8 foi isolado a partir de células de *E. coli* relevantes usando um kit Mini-Prep (Qiagen, UK).

A amplificação por PCR, usando o DNA de fagemídeo como alvo e as sequências iniciadoras que se seguem, foi usada para amplificar as sequências de DNA das cadeias α e β de TCR.

Sequência iniciadora directa da cadeia alfa de TCR A6

ggaattc atcgatg cagaaggaagtggagcag (SEQ ID 129)

(O local de restrição ClaI está sublinhado)

Sequência iniciadora reversa universal da cadeia alfa de TCR

gtacacggccgccgggttcttgatatac (SEQ ID 130)

(o local de restrição EagI está sublinhado)

Sequência iniciadora directa da cadeia beta A6

tctctccattaatgaatgctggtgtcactcagacccc (SEQ ID 131)

(o local de restrição AseI está sublinhado)

Sequência iniciadora reversa universal da cadeia beta de TCR

tagaaaccggtggccaggcacaccagtgtggc (SEQ ID 132)

(o local de restrição AgeI está sublinhado)

No caso da cadeia β de TCR, foi efectuado um outro PCR de reparação para substituir o codão de paragem amber na região CDR3 com um codão codificador de ácido glutâmico. Quando um codão de paragem amber é ignorado em *E. coli*, um resíduo de glutamina é normalmente introduzido em lugar da tradução ser parada. Assim, quando o TCR contendo o codão amber é apresentado na superfície dos fagos, possui um resíduo de glutamina nesta posição. No entanto, quando o gene da cadeia β de TCR foi transferido para o plasmídeo de expressão, um resíduo de ácido glutâmico foi usado como alternativa a glutamina. As sequências iniciadoras usadas para este PCR de reparação foram as seguintes.

YOL124 CTGCTCTGGTTCGGCACTC (SEQ ID 133)

YOL125 GAGTGCGGAACCAGAGCAG (SEQ ID 134)

A sequência do DNA da cadeia β de TCR A6 solúvel mutada foi verificada por sequenciação automática (ver Figura 14a para a sequência do DNA da cadeia β de TCR A6 e 14b para a respectiva sequência de aminoácidos codificada). A Figura 14c mostra a sequência de aminoácidos da cadeia β de TCR A6 sem a substituição da glutamina por ácido glutâmico, *i.e.* a sequência que estava presente no clone 134 conforme isolado por ELISA de fagos.

Estas sequências de DNA de TCR A6 α e β foram então usadas para produzir um TCR A6 solúvel como descrito em WO 03/020763. Resumidamente, as duas cadeias foram expressas como corpos de inclusão em culturas de *E. coli* separadas. Os corpos de inclusão foram então isolados, desnaturados e renaturados em conjunto *in vitro*.

Exemplo 10 - Caracterização por ressonância de plasmões de superfície BIAcore de um TCR A6 de elevada afinidade que se liga a HLA-A2 Tax.

Um biosensor de ressonância de plasmões de superfície (BIAcore 3000™) foi usado para analisar a ligação do clone 134 de elevada afinidade (Ver Figuras 15a & 15b para as sequências completas de DNA e de aminoácidos da cadeia β de TCR mutada, respectivamente) ao ligando HLA-A2 Tax. Isto foi facilitado através da produção de complexos pMHC (descrito abaixo) que foram imobilizados a uma superfície de ligação revestida com estreptavidina numa forma semi-orientada, permitindo a testagem eficiente da ligação de um receptor de células T solúvel a um máximo de quatro pMHC diferentes (imobilizados em células de fluxo separadas) simultaneamente. A injeção manual do complexo HLA permite que o nível preciso das moléculas classe I imobilizadas seja facilmente manipulado.

Os complexos HLA-A2 classe I tax biotinilados foram renaturados *in vitro* a partir de corpos de inclusão

expressos em bactérias contendo as proteínas das subunidades constituintes e o péptido sintético, seguido de purificação e biotinylação enzimática *in vitro* (O'Callaghan *et al.* (1999) Anal. Biochem. 266:9-15). A cadeia pesada de HLA foi expressa com uma marca de biotinylação C-terminal que substitui os domínios transmembranar e citoplasmático da proteína numa construção adequada. Foram obtidos níveis de expressão dos corpos de inclusão de ≈ 75 mg/litro de cultura bacteriana. A cadeia leve de HLA ou a $\beta 2$ -microglobulina foi igualmente expressa como corpos de inclusão em *E. coli* a partir de uma construção adequada, num nível de ≈ 500 mg/litro de cultura bacteriana.

As células de *E. coli* foram lisadas e os corpos de inclusão foram purificados até aproximadamente 80% de pureza. A proteína dos corpos de inclusão foi desnaturada em guanidina-HCl 6M Tris 50 mM pH 8,1, NaCl 100 mM, DTT 10 mM, EDTA 10 mM e foi renaturada numa concentração de 30 mg/litro de cadeia pesada, 30 mg/litro de $\beta 2m$ em L-arginina-HCl 0,4M, Tris 100 mM, pH 8,1, cistamina 3,7 mM, cisteamina mM, 4 mg/ml de péptido (e.g., tax 11-19), através da adição de um único pulso de proteína desnaturada no tampão de renaturação a $<5^{\circ}\text{C}$. A renaturação decorreu até estar completa, a 4°C , durante pelo menos 1 hora.

O tampão foi mudado por diálise em 10 volumes de Tris 10 mM pH 8,1. Foram necessárias duas mudanças de tampão para reduzir suficientemente a força iónica. A solução proteica foi então filtrada através de um filtro de

acetato de celulose de 1,5 μ m e aplicada numa permuta aniónica POROS 50HQ (volume de 8 ml). A proteína foi eluída com um gradiente linear de NaCl 0-500 mM. O complexo HLA-A2-péptido eluiu a aproximadamente NaCl 250 mM e as fracções do pico foram colhidas, adicionado uma mistura de inibidores de protease (Calbiochem) e as fracções arrefecidas em gelo.

Mudou-se o tampão dos complexos HLA-A2 marcados por biotinilação para Tris 10 mM pH 8,1, NaCl 5 mM, usando uma coluna para remoção rápida de sal Pharmacia equilibrada no mesmo tampão. Imediatamente após a eluição, as fracções contendo proteína foram arrefecidas em gelo e adicionada a mistura de inibidores de proteases (Calbiochem). Os reagentes de biotinilação foram então adicionados: biotina 1 mM, ATP 5 mM (tamponado para pH 8), MgCl₂ 7,5 mM e 5 μ g/ml da enzima BirA (purificada de acordo com O'Callaghan *et al.* (1999) *Anal. Biochem.* 266: 9-15). A mistura foi então deixada a incubar à temperatura ambiente durante a noite.

Os complexos HLA-A2 biotinilados foram purificados por cromatografia de filtração em gel. Uma coluna Pharmacia Superdex 75 HR 10/30 foi pré-equilibrada com PBS filtrado, aplicou-se 1 ml da mistura de reacção de biotinilação e a coluna foi eluída com PBS a 0,5 ml/min. Os complexos HLA-A2 biotinilados eluíram como um único pico a aproximadamente 15 ml. As fracções contendo proteína foram

reunidas, arrefecidas em gelo e adicionada a mistura de inibidores de proteases. A concentração proteica foi determinada usando um ensaio de ligação de Coomassie (PerBio) e alíquotas dos complexos de HLA-A2 biotinilados foram guardadas congeladas a -20°C. A estreptavidina foi imobilizada por métodos convencionais de acoplagem de amina.

As interacções entre TCR A6 Tax de elevada afinidade contendo uma nova ligação entre cadeias e o complexo HLA-A2 Tax ou uma combinação irrelevante HLA-A2 NY-ESO, cuja produção está descrita atrás, foram analisadas num biosensor de ressonância de plasmões de superfície (SPR) BIAcore 3000™. SPR mede alterações no índice de refração expressas em unidades de resposta (RU) perto da superfície de sensor dentro de uma pequena célula de fluxo, um princípio que pode ser usado para detectar interacções de receptor e ligando e para analisar a sua afinidade e parâmetros cinéticos. As células de fluxo sonda foram preparadas através da imobilização dos complexos peptídicos individuais HLA-A2 em células de fluxo separadas através da ligação entre a biotina ligada de forma cruzada a β 2m e estreptavidina, que tinham sido ligadas de forma cruzada à superfície activada das células de fluxo. O ensaio foi então realizado através da passagem de scTCR sobre as superfícies das diferentes células de fluxo, a um fluxo constante, medindo a resposta SPR ao fazê-lo. Inicialmente, a especificidade da interacção foi verificada através da

passagem de TCR A6 solúvel num fluxo constante a $5 \mu\text{l min}^{-1}$ sobre quatro superfícies diferentes; uma revestida com ≈ 1000 RU do complexo HLA-A2 Tax, a segunda revestida com ≈ 1000 RU do complexo HLA-A2 NY-ESO e duas células de fluxo brancas revestidas apenas com estreptavidina (ver Figura 15).

O aumento da afinidade do TCR A6 solúvel mutado tornou difíceis os cálculos de K_d para a interação deste grupo com o complexo HLA-A2 Tax. No entanto, a semi-vida ($t_{1/2}$) para a interação foi calculada como sendo 51,6 minutos (ver Figura 16), que se compara com uma $t_{1/2}$ da interação selvagem de 7,2 segundos.

Exemplo 11 - Produção de vector codificador de um TCR NY-ESO solúvel com uma nova ligação dissulfureto.

A cadeia β do TCR A6 solúvel preparada no Exemplo 1 possui na sequência nativa um local de restrição BglII (AAGCTT) adequado para usar como um local de ligação.

A mutagénese por PCR foi realizada como detalhado abaixo para introduzir um local de restrição BamHI (GGATCC) na cadeia α de TCR A6 solúvel, 5' relativamente ao novo codão cisteína. A sequência descrita na Figura 2a foi usada como matriz para esta mutagénese. Foram usadas as sequências iniciadoras que se seguem:

|BamHI |

5' -ATATCCAGAACCCgGATCCTGCCGTGTA-3' (SEQ ID 135)

5'-TACACGGCAGGAATCCGGGTTCTGGATAT-3' (SEQ ID 136)

100 ng de plasmídeo foram misturados com 5 µl de dNTP 10 mM, 25 µl de tampão Pfu 10X (Stratagene), 10 unidades de polimerase Pfu (Stratagene) e o volume final foi ajustado a 240 µl com H₂O. 48 µl desta mistura foram suplementados com sequências iniciadoras diluídas para dar uma concentração final de 0,2 µM em 50 µl de volume final de reacção. Após um passo de desnaturação inicial de 300 segundos a 95°C, a mistura de reacção foi sujeita a 15 ciclos de desnaturação (95°C, 30 seg.), emparelhamento (55°C, 60 seg.) e elongação (73°C, 8 min) num aparelho de PCR Hybaid PCR Express. O produto foi então digerido durante 5 horas, a 37°C, com 10 unidades da enzima de restrição DpnI (New England Biolabs). 10 µl da reacção de digestão foram usados para transformar bactérias competentes XL1-Blue e estas crescidas durante 18 horas a 37°C. Uma colónia isolada foi picada e crescida durante a noite em 5 ml de TYP + ampicilina (16 g/l de Bactotripton, 16 g/l de extracto de levedura, 5 g/l de NaCl, 2,5 g/l K₂HPO₄, 100 mg/l de ampicilina). O DNA de plasmídeo foi purificado numa coluna mini-prep Qiagen de acordo com as instruções do fabricante e a sequência foi verificada por sequenciação automática na secção de sequenciação do Departamento de Bioquímica, Universidade de Oxford. O cDNA

codificador do TCR NY-ESO foi isolado a partir de células T de acordo com técnicas conhecidas. O cDNA codificador de TCR NY-ESO foi produzido por tratamento do mRNA com transcriptase reversa.

Para produzir vectores codificadores de um TCR NY-ESO solúvel contendo uma nova ligação dissulfureto, os plasmídeos TCR A6 contendo os locais BamHI na cadeia α e BglIII na cadeia β foram usados como matrizes. Foram usadas as sequências iniciadoras que se seguem:

| NdeI |

5'-GGAGATATACATATGCAGGAGGTGACACAG -3' (SEQ ID 137)

5'-TACACGGCAGGATCCGGGTTCTGGATATT-3' (SEQ ID 138)

| BamHI |

| NdeI |

5'-GGAGATATACATATGGGTGTCACTCAGACC-3' (SEQ ID 139)

5'-CCCAAGCTTAGTCTGTCTTACCCAGGCCTCGGC -3' (SEQ ID 140)

| BglIII |

As construções das cadeias α e β de TCR NY-ESO foram obtidas por clonagem de PCR como se segue. As reacções de PCR foram efectuadas usando as sequências iniciadoras mostradas acima e as matrizes contendo as cadeias de TCR NY-ESO nativas. Os produtos de PCR foram sujeitos a digestão com enzimas de restrição e clonados em pGMT7 para se obter plasmídeos de expressão. A sequência dos insertos dos plasmídeos foram confirmadas por sequen-

iação automática de DNA. As Figuras 17a e 17b mostram a sequência de DNA das cadeias α e β , respectivamente, e as Figuras 18a e 18b mostram as sequências de aminoácidos resultantes.

Exemplo 12 - Construção de vectores de apresentação fágica e clonagem de DNA codificador das cadeias α e β de TCR NY-ESO nos vectores fagemídeos.

O DNA codificador das cadeias α e β de TCR NY-ESO com novos codões de cisteína para facilitar a formação de uma ligação dissulfureto, não nativa, entre cadeias, produzidos como descrito no Exemplo 11 foram incluídos no vector fagemídeo pEX746 como se segue.

O DNA codificador das duas cadeias TCR NY-ESO foram sujeitos, individualmente, a PCR de forma a introduzir locais de clonagem compatíveis com o vector fagemídeo pEX746 (contendo DNA codificador do clone 7 de TCR A6) usando as seguintes sequências iniciadoras:

Para a cadeia alfa de TCR NY-ESO

TRAV21

GCCGGCCATGGCCCAAACAGGAGGTGACGCAGATTCT(SEQ ID 141)

YOL6

**CTTCTTAAAGAATTCTTAATTAACTTAGGTTATTAGGAACTTTCTGGGCTG
GGGAAG (SEQ ID 142)**

Para a cadeia beta de TCR NY-ESO

TRBV6-1/2/3/5/6/7/8/9 TCACAGCGCGCAGGCTGGTGTCACTCAGACCCCAAA (SEQ ID 143)
RT1 CGAGAGCCCGTAGAACTGGACTTG (SEQ ID 144)

Os métodos de clonagem molecular para a construção dos vectores estão descritos em "Molecular cloning: A laboratory manual", de J. Sambrooke D.W. Russell. As sequências iniciadoras apresentadas na Tabela 1 foram usadas para a construção dos vectores. Um exemplo do programa de PCR é 1 ciclo de 94°C durante 2 minutos, seguido de 25 ciclos de 94°C durante 5 segundos, 53°C durante 5 segundos e 72°C durante 90 segundos, seguido de 1 ciclo de 72°C durante 10 minutos e depois mantido a 4°C. A DNA polimerase Taq Expand HighFidelity foi adquirida à Roche.

O DNA codificador do clone 7 da cadeia β de TCR A6 foi removido de pEX746 por digestão com enzimas de restrição BssHII e BglII. O correspondente DNA de PCR digerido, codificador da cadeia β de NY-ESO, foi então substituído no fagemídeo através de ligação. A sequência do produto de clonagem foi verificada por sequenciação automática.

De forma semelhante, o DNA codificador do clone 7 da cadeia α de TCR A6 foi removido de pEX746 através da digestão com as enzimas de restrição NcoI e AvrII. O correspondente DNA de PCR digerido, codificador da cadeia α de NY-ESO, foi então substituído no fagemídeo já

contendo o DNA codificador da cadeia β de TCR NY-ESO através de ligação. A sequência do produto de clonagem foi verificada por sequenciação automática.

As Figuras 19a e 19b descrevem de forma detahada, respectivamente a sequência de DNA e de aminoácidos da cadeia α e β de TCR NY-ESO, assim como a sequência releante incluída no fagemídeo (pEX746:NY-ESO). A sequência que precede o local NcoI é a mesma de pEX746.

Exemplo 13 - Expressão bacteriana de fusões da proteína externa e de TCR NY-ESO heterodimérico em E. coli.

As partículas fágicas apresentadoras de TCR NY-ESO heterodimérico, contendo uma ligação dissulfureto não nativa entre cadeias, foram preparadas usando métodos descritos anteriormente para a geração de partículas fágicas apresentadoras de anticorpos scFvs (Li et al., 2000, Journal of Immunological Methods 236: 133-146) com as modificações que se seguem. Células *E. coli* TG1 contendo o fagemídeo pEX746:NY-ESO (*i.e.* o fagemídeo codificador da cadeia α de TCR NY-ESO solúvel e uma cadeia β de TCR NY-ESO fundida com a proteína fágica gIII, produzido como descrito no Exemplo 12) foram usadas para inocular 10 ml de TY 2x (contendo 100 μ g/ml de ampicilina e 2% de glucose) e depois a cultura foi incubada com agitação, a 37°C, durante a noite (16 horas). 50 μ l da cultura crescida durante a noite foram usados para inocular 10 ml de TY 2x (contendo 100

µg/ml de ampicilina e 2% de glucose) e a cultura foi incubada com agitação a 37°C até uma $DO_{600\text{ nm}} = 0,8$. O fago auxiliar HYPERPHAGE foi adicionado à cultura para uma concentração final de 5×10^9 pfu/ml. A cultura foi então incubada a 37°C sem agitação, durante trinta minutos, e depois com agitação, a 200 rpm, durante mais 30 minutos. O meio da cultura atrás referida foi então aumentado para 50 ml com TY 2x (contendo 100 µg/ml de ampicilina e 25 µg/ml de canamicina), a cultura foi incubada a 25°C com agitação a 250 rpm durante 36 a 48 horas. A cultura foi depois centrifugada a 4°C, durante 30 minutos, a 4000 rpm. O sobrenadante foi filtrado através de uma seringa com filtro de 0,45 µm e guardado a 4°C para ser mais concentrado. O sobrenadante foi então concentrado por precipitação com PEG e ressuspenso em PBS a 10% do volume original guardado.

Exemplo 14 - Detecção de TCR NY-ESO heteroimérico funcional com uma ligação dissulfureto, não nativa, entre cadeias em partículas fágicas filamentosas.

A presença de TCR NY-ESO funcional (ligação a HLA-A2-NY-ESO), apresentado nas partículas fágicas da suspensão concentrada preparada no Exemplo 13, foi detectada usando os métodos de ELISA para fagos descritos no Exemplo 4. A Figura 20 mostra a ligação específica de partículas fágicas apresentadoras de TCR NY-ESO a HLA-A2-NY-ESO num ensaio de ELISA para fagos.

Exemplo 15 - Construção de plasmídeos para a expressão celular dos genes HLA-DRA.

As sequências de DNA codificadoras da fracção extracelular das cadeias HLA-DRA foram amplificadas a partir de cDNA isolado a partir do sangue de um indivíduo saudável, usando a reacção em cadeia da polimerase (PCR), com pares de sequências iniciadoras de DNA sintético que foram projectadas para incluir um local de restrição BglII.

A mutagénese por PCR foi então usada para adicionar DNA codificador do fecho de leucinas Fos ao extremo 3' da sequência amplificada.

As manipulações de DNA e clonagem descritas atrás foram realizadas como descrito em Sambrook, J et al., (1989). "Molecular Cloning - A laboratory Manual". Second edition. Cold Spring Harbor Laboratory Press, USA.

A Figura 21 proporciona a sequência de DNA da cadeia β de HLA-DR pronta para inserção no vector de expressão bi-cistrónico. Esta figura indica a posição dos codões codificadoras do péptido fecho de leucinas Fos e da marca de biotinilação.

A numeração de aminoácidos baseia-se na sequência de murganho (Kabat, 1991, "Sequences of Proteins of Immunological Interest", 5th edition, US Dept of Health & Human Services, Public Health Service, NIH, Bethesda, MD 1-1137).

Esta sequência de DNA foi então inserida num vector baculovírus bicistrónico pAcAB3 (Ver Figura 22 para a sequência deste vector) juntamente com DNA codificador da correspondente cadeia β de HLA Classe II em células de insecto Sf9. Este vector pode ser usado para expressar qualquer complexo HLA Classe II-péptido em células de insecto.

Exemplo 16. Construção de plasmídeos para a expressão celular de HLA-DRB selvagem e genes mutantes.

Sequências de DNA codificadoras da fracção extracelular das cadeias HLA-DRB foram amplificadas a partir de cDNA isolado de sangue de um ser humano saudável, usando a reacção em cadeia da polimerase (PCR) com pares de sequências iniciadoras de DNA sintético que foram projectadas para incluírem um local de restrição BamHI.

A mutagénese por PCR foi então usada para adicionar DNA codificador do fecho de leucinas Jun ao extremo 3' da sequência amplificada e o DNA codificador do péptido Flu HA carregado pela molécula HLA-DR1 ao extremo 5' da sequência.

As manipulações de DNA e clonagem descritas atrás foram efectuadas como descrito em Sambrook, J et al., (1989). "Molecular Cloning - A laboratory Manual". Second edition. Cold Spring Harbor Laboratory Press, USA.

A Figura 23 apresenta a sequência de DNA da cadeia β de HLA-DR pronta para inserção no vector de expressão bicistrónico. Esta figura indica a posição dos codões codificadores do péptido fecho de leucinas Jun e o péptido Flu HA.

A numeração de aminoácidos baseia-se na sequência de murganho (Kabat, 1991, "Sequences of proteins of Immunological Interest", 5th edition, US Dept of Health & Human Services, Public Health Service, NIH, Bethesda, MD 1-1137).

Esta sequência de DNA foi então inserida num vector bicistrónico de baculovírus pAcAB3 (Ver Figura 22 para a sequência deste vector) juntamente com DNA codificador da cadeia α Classe II correspondente para expressão em células Sf9. Este vector pode ser usado para expressar qualquer complexo HLA Classe II-péptido em células de insecto.

Exemplo 17 - Expressão e renaturação de complexos HLA-DR1 Classe II-Flu HA

A expressão de MHC Classe II foi realizada usando os vectores de expressão bicistrónicos, produzidos como descrito nos Exemplos 15 e 16, contendo as cadeias α e β de HLA-DR1 Classe II e o péptido Flu HA. Os métodos de expressão e de purificação usados estão descritos em

(Gauthier (1998) PNAS USA 95 p11828-118333). Resumidamente, HLA-DR1 solúvel foi expresso no sistema de baculovírus através da substituição das regiões transmembranares hidrofóbicas e segmentos citoplasmáticos das cadeias α e β DR com domínios de dimerização fecho de leucina dos factores de transcrição Fos e Jun. Na construção de expressão, a sequência necessária MHC Classe II carregado com Flu HA foi ligada covalentemente ao extremo N da cadeia β DR madura e a cadeia α DR possui uma sequência de biotinylation para facilitar a formação de ligandos bifuncionais usando a metodologia de multimerização biotina/estreptavidina. A proteína recombinante foi secretada pelas células Sf9 infectadas com o baculovírus recombinante e purificada por cromatografia de afinidade. A proteína foi ainda purificada por HPLC de permuta aniónica.

Exemplo 18 - Construção de moléculas solúveis de péptido-HLA Classe I

Para investigar melhor a especificidade do clone 134 de TCR A6 de elevada afinidade, foram produzidas as seguintes moléculas solúveis de péptido-HLA Classe I:

HLA-A2 - péptido (LLGRNSFEV) (**SEQ ID 23**)

HLA-A2 - péptido (KLVALGINAV) (**SEQ ID 24**)

HLA-A2 - péptido (LLGDLFGV) (**SEQ ID 25**)

HLA-B8 - péptido (FLRGRAYGL) (**SEQ ID 26**)

HLA-B27 - péptido (HRCQAIRKK) (**SEQ ID 27**)

HLA-Cw6 - péptido (YRSGIIAVV) (**SEQ ID 28**)

HLA-A24 - péptido (VYGFVRACL) (**SEQ ID 29**)

HLA-A2 - péptido (ILAKFLHWL) (**SEQ ID 30**)

HLA-A2 - péptido (LTLGEFLKL) (**SEQ ID 31**)

HLA-A2 - péptido (GILGFVFTL) (**SEQ ID 33**)

HLA-A2 - péptido (SLYNTVATL) (**SEQ ID 34**)

Estes péptidos-HLAs solúveis foram produzidos usando os métodos descritos no Exemplo 10.

Exemplo 19 - Medição por ressonância de plasmões de superfície BIAcore da especificidade da ligação de elevada afinidade do Clone 134 de TCR A6 ao péptido-HLA.

Um biosensor de ressonância de plasmões de superfície (BIAcore 3000™) foi usado para analisar a especificidade da ligação do clone 134 de TCR A6 de elevada afinidade. (Ver Figuras 15a & 15b para as sequências completas de DNA e de aminoácidos, respectivamente, da cadeia β de TCR mutada). Isto foi efectuado usando os complexos HLA-DR1 Classe II-péptido produzido como descrito nos Exemplos 15-17 e HLA Classe I-péptido apresentado no Exemplo 18, produzido usando os métodos descritos detalhadamente no Exemplo 10. A tabela que se segue apresenta os complexos péptido-HLA utilizados:

1. HLA-A2 - péptido (LLGRNSFEV) (**SEQ ID 23**)
2. HLA-A2 - péptido (KLVALGINAV) (**SEQ ID 24**)
3. HLA-A2 - péptido (LLGDLFGV) (**SEQ ID 25**)
4. HLA-B8 - péptido (FLRGRAYGL) (**SEQ ID 26**)

5. HLA-B27 - péptido (HRCQAIRKK) **(SEQ ID 27)**
6. HLA-Cw6 - péptido (YRSGIIAVV) **(SEQ ID 28)**
7. HLA-A24 - péptido (VYGFVRACL) **(SEQ ID 29)**
8. HLA-A2 - péptido (ILAKFLHWL) **(SEQ ID 30)**
9. HLA-A2 - péptido (LTLGEFLKL) **(SEQ ID 31)**
10. HLA-DR1 - péptido (PKYVKQNTLKLA) **(SEQ ID 32)**
11. HLA-A2 - péptido (GILGFVFTL) **(SEQ ID 33)**
12. HLA-A2 - péptido (SLYNTVATL) **(SEQ ID 34)**

Os péptidos-HLAs acima foram imobilizados em superfícies revestidas com estreptavidina nas células de fluxo de um BIAcore 3000™ numa forma semi-orientada.

O BIAcore 3000™ permite testar a ligação do receptor de células T solúvel a até quatro pMHC diferentes (imobilizados em células de fluxo separadas) em simultâneo. Para esta experiência, três HLAs-péptidos diferentes foram imobilizados nas células de fluxo 2-4 e a célula de fluxo 1 foi deixada como branco como controlo. A injeção manual dos complexos HLA-péptido permitiu que fosse manipulado o nível preciso das moléculas imobilizadas.

Após avaliada a capacidade do clone 134 de TCR A6 de elevada afinidade para se ligar aos 3 primeiros complexos de HLA-péptido na lista acima, os três seguintes foram imobilizados nestas células de fluxo directamente por cima dos anteriores. Este processo foi continuado até a ligação do clone 134 de TCR A6 de elevada afinidade a todos os 12 complexos de HLA-péptido ter sido avaliada.

Dez injeções de 5 µl do clone 134 de TCR A6 de elevada afinidade foram passadas por cada uma das células a 5 µl/min, em concentrações que variam entre 4,1 ng/ml a 2,1 mg/ml. (Ver Figuras 24-28).

Como controlo final, o clone 134 de TCR A6 de elevada afinidade foi passado numa célula de fluxo contendo HLA-A2 Tax (LLFGYPVYV) (SEQ ID 21) imobilizado, o ligando cognato deste TCR.

A ligação específica do clone 134 de TCR A6 de elevada afinidade foi apenas observada com o seu ligando cognato. (HLA-A2 Tax (LLFGYPVYV) (SEQ ID 21)). Estes dados ainda demonstram a especificidade do clone 134 de TCR A6 de elevada afinidade. (Ver Figuras 24-28).

Exemplo 20 - Mutagénesse das regiões CDR3 de TCR NY-ESO

As regiões CDR3 de TCR NY-ESO foram consideradas para a introdução de mutações para estudar a possibilidade da geração de mutantes de elevada afinidade. Isto foi conseguido usando as sequências iniciadoras de PCR específicas de TCR NY-ESO em combinação com métodos substancialmente iguais aos descritos detalhadamente no Exemplo 7.

Exemplo 21 - Isolamento de mutantes de TCR A6 de elevada afinidade

O isolamento de mutantes TCR NY-ESO de elevada afinidade foi realizado usando o primeiro dos dois métodos descritos no Exemplo 8.

Foi identificado um único mutante TCR NY-ESO de elevada afinidade.

Exemplo 22 - Produção de TCR NY-ESO solúvel heterodimérico de elevada afinidade, com ligação dissulfureto não nativa entre as regiões constantes, com mutações nas regiões variáveis.

O DNA de fagemídeo, codificador do mutante de TCR NY-ESO de elevada afinidade, identificado no Exemplo 21 foi isolado a partir de células de *E. coli* relevantes usando um kit Mini-Prep (Qiagen, UK)

A amplificação por PCR usando o DNA de fagemídeo como alvo e as sequências iniciadoras que se seguem foi usada para amplificar a sequência de DNA da região variável da cadeia β de TCR NY-ESO solúvel mutada.

Sequência iniciadora directa da cadeia beta NY-ESO
tctctcattaatgaatgctggtgtcactcagacccc (SEQ ID 145)
(o local de restrição AseI está sublinhado)

Sequência iniciadora reversa universal da cadeia beta
tagaaaccggtggccaggcacaccagtgtgg (SEQ ID 146)
(o local de restrição AgeI está sublinhado)

O produto de PCR foi então digerido com AgeI/AseI e clonado em pEX821 (produzido como descrito no Exemplo 11) cortado com Nde/AgeI.

A sequência de DNA da cadeia β de TCR NY-ESO mutada amplificada como descrito atrás, e a cadeia α de TCR NY-ESO produzida como descrito no Exemplo 11, foram então usadas para produzir um TCR NY-ESO de elevada afinidade solúvel como descrito em WO 03/020763. Resumidamente, as duas cadeias foram expressas como corpos de inclusão em culturas separadas de *E. coli*. Os corpos de inclusão foram então isolados, desnaturados e renaturados em conjunto *in vitro*.

Exemplo 23 - Caracterização por ressonância de plasmões de superfície BIAcore da ligação de um TCR NY-ESO de elevada afinidade a HLA-A2 NY-ESO.

Um biosensor de ressonância de plasmões de superfície (BIAcore 3000™) foi usado para analisar a ligação do TCR NY-ESO de elevada afinidade ao ligando HLA-A2 NY-ESO. Isto foi facilitado pela produção de complexos pMHC (como descrito no Exemplo 10) que foram imobilizados numa superfície de ligação revestida com estreptavidina numa forma semi-orientada, permitindo a testagem eficiente

da ligação de um receptor de células T solúvel a até quatro pMHC diferentes (imobilizados em células de fluxo separadas) simultaneamente. A injeção manual do complexo de HLA permite que seja facilmente manipulável o nível preciso de moléculas classe I.

As interações entre o TCR NY-ESO de elevada afinidade contendo uma nova ligação entre cadeias e o complexo HLA-A2 NY-ESO ou uma combinação HLA-A2 Tax irrelevante, a produção do qual está descrita descrita no Exemplo 10, foram analisadas num bio-sensor de ressonância de plasmões de superfície (SPR) Biacore 3000™, novamente como descrito no Exemplo 10.

O Kd para a interação do NY-ESO de elevada afinidade solúvel com o HLA-A2 NY-ESO foi calculado como sendo 4,1 μM , (Ver Figuras 29a e 29b) que se compara com um kd de 15,7 μM para a interação selvagem. (Ver Figuras 30a e 30b).

Exemplo 24 - Produção e testagem de outros TCRs A6 de elevada afinidade

TCRs solúveis contendo as mutações que seguem correspondendo às identificadas nos clones 89, 1, 111 e 71 (ver Exemplo 8) foram produzidos usando os métodos detalhados no Exemplo 9. A ligação destes TCRs solúveis a HLA-A2 Tax foi então avaliada com o ensaio BIAcore descrito detalhadamente no Exemplo 10.

Mutantes da cadeia 3 de TCR A6

Clone	Sequência CDR3
selvagem	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGAGCAGTAG (SEQ ID 83) A S R P G L A G G R P E Q Y (SEQ ID 84)
89	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCAGAGGATTAG (SEQ ID 91) A S R P G L A G G R P E D X (SEQ ID 92)
1	GCCTCGAGGCCGGGGCTGGTTCCGGGGCGACCAGAGCAGTAG (SEQ ID 97) A S R P G L V P G R P E Q X (SEQ ID 98)
111	GCCTCGAGGCCGGGACTAGCGGGAGGGCGACCACATCCGTAG (SEQ ID 101) A S R P G L A G G R P H P X (SEQ ID 102)

Mutante da cadeia α de TCR A6

Clone	Sequência CDR3
selvagem	GCCGTTACAACTGACAGCTGGGGGAAGCTTCAG (SEQ ID 107) A V T T D S W G K L Q (SEQ ID 108)
71	GCCGTTACAACTGACAGCTGGGGGGAGCTTCAT (SEQ ID 117) A V T T D S W G E L H (SEQ ID 118)

Mutações combinadas usadas como base para a produção de TCRs A6 solúveis mutados:

Mutações do Clone 89 + mutações do Clone 134

Mutações do Clone 71 + mutações do Clone 134

Mutações do Clone 71 + mutações do Clone 89

Clone 1 + mutação β G102→A

Resultados:

A tabela que se segue compara a afinidade de HLA-A2 Tax para os TCRs A6 mutados solúveis acima com a obtida usando um TCR A6 solúvel contendo regiões variáveis não mutadas. Note-se que a afinidade dos mutantes com mais alta afinidade é expressa como semi-vida para a interação.

($T_{1/2}$) Estes TCRs A6 mutantes solúveis apresentaram maior afinidade para HLA-A2 Tax do que TCR A6 solúvel não mutado, conforme demonstrado pelo seu k_d mais baixo ou $T_{1/2}$ mais elevada para a interacção.

As Figuras 31-37 mostram os perfis de BIAcore usados para calcular a afinidade para HLA-A2 Tax destes TCRs mutados solúveis. As Figuras 38a-e mostram a sequência de aminoácidos das cadeias TCR A6 mutadas.

TCR A6	K_d (nM)	$T_{1/2}$ (secs)
Selvagem	1,9	7
Clone 1		810
Clone 89	0,41	
Clone 111	1,18	
Clone 71	1,37	
Mutações do Clone 89 + clone 134		114 (fase 1) 4500 (fase 2)
Mutações do Clone 71 + clone 134		882
Mutações do Clone 71 + clone 89	0,35	
Clone 1 + mutação $\beta G102 \rightarrow A$		738

Exemplo 25 - Coloração de células usando tetrameros de monómeros de TCR A6 de elevada afinidade

As células apresentadoras de antígeno T2 foram

incubadas com $\beta 2m$ (3 $\mu g/ml$) pulsada com o péptido Tax numa gama de concentrações (10^{-5} - $10^{-9}M$), durante 90 minutos, a $37^{\circ}C$. Os controlos, usando também células T2 incubadas com $\beta 2m$ (3 $\mu g/ml$), foram pulsados com o péptido Flu $10^{-5}M$ ou incubados sem péptido (não pulsados). Após o pulso, as células foram lavadas em RPMI sem soro e 2×10^5 células foram incubadas com o tetrâmero de TCR A6 Clone 134 de elevada afinidade ligado a estreptavidina marcado com ficoeritrina (PE) (Molecular Probes, The Netherlands) (10 $\mu g/ml$) ou com os monómeros de TCR A6 Clone 134 de elevada afinidade marcados com Alexa 488 (Molecular Probes, The Netherlands) durante 10 minutos, à temperatura ambiente. Após lavagem das células, a ligação dos tetrâmeros e monómeros de TCR marcados foi avaliada por citometria de fluxo usando um FACSVantage SE (Bencton Dickinson).

Resultados

Como ilustrado na Figura 39a a coloração específica de células T2 pelos tetrâmeros de TCR A6 de elevada afinidade foi observada para concentrações do péptido Tax até $10^{-9}M$.

Como ilustrado na Figura 39b a coloração específica de células T2 pelos monómeros de TCR A6 de elevada afinidade foi observada para concentrações do péptido Tax até $10^{-8}M$.

Organização requerente

Rua: 57C Milton Park
Cidade: Abingdon
Estado: Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: martin.green@avidex.com

<110> Nome da Organização: Avidex Ltd

Requerente Individual

Rua: Avidex Ltd, 57C Milton Park
Cidade: Abingdon
Estado: Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: bent.jakobsen@avidex.com

<110> Último Nome: Jakobsen
<110> Primeiro Nome: Bent
<110> Inicial do Meio: K
<110> sufixo: Dr

Requerente Individual

Rua: Avidex Ltd, 57C Milton Park
Cidade: Abingdon
Estado: Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: torben.andersen@avidex.com

<110> Último Nome: Andersen
<110> Primeiro Nome: Torben
<110> Inicial do Meio: B
<110> sufixo: Dr

Requerente Individual

Rua: Avidex Ltd, 57C Milton Park
Cidade: Abingdon
Estado: Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: peter.molloy@avidex.com

<110> Último Nome: Molloy
<110> Primeiro Nome: Peter
<110> Inicial do Meio: E
<110> sufixo: Dr

Requerente Individual

Rua: Avidex Ltd, 57C Milton Park
Cidade: Abingdon
Estado:Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: jonathan.boulter@avidex.com

<110> Último Nome: Boulter
<110> Primeiro Nome: Jonathan
<110> Inicial do Meio: M
<110> sufixo: Dr

Requerente Individual

Rua: Avidex Ltd, 57C Milton Park
Cidade: Abingdon
Estado:Oxfordshire
País: Inglaterra
Código Postal: OX14 4RX
Número de Telefone: +44 1235 438600
Número de Fax: +44 1235 438601
Endereço de correio electrónico: yi.li@avidex.com

<110> Último Nome: Li
<110> Primeiro Nome: Yi
<110> Inicial do Meio:
<110> sufixo: Dr

Projecto do Pedido de Patente

<120> Título: substâncias
<130> Referência do Ficheiro do Pedido de Patente: pasta N° 19
<140> Presente Número de Pedido de Patente: GB 0226227.7
<141> Data da Entrega: 2002-11-09

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: GB0226227.7
<151> Data de Entrega Anterior: 2002-11-09

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: GB 0301814.0
<151> Data de Entrega Anterior: 2003-01-25

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: GB 0304067.2
<151> Data de Entrega Anterior: 2003-02-22

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: GB 0311397.4

<151> Data de Entrega Anterior: 2003-05-16

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: GB 0316356.5

<151> Data de Entrega Anterior: 2003-07-11

Pedidos de Patente Anteriores

<150> Número do Pedido de Patente Anterior: US 60/436 323

<151> Data de Entrega Anterior: 2003-04-16

Sequência

<213> nome do organismo: sequência sintética

<400> segmento de pré-sequência:

DSDVYITDKT VLDMRSMDFK

20

<212> Tipo: PRT

<211> Comprimento: 20

Nome da sequência: SEQ ID 1

Descrição da sequência:

Sequência

<213> nome do organismo: sequência sintética

<400> segmento de pré-sequência:

QSKDSDVYIT DKTVLDMRSM

20

<212> Tipo: PRT

<211> Comprimento: 20

Nome da sequência: SEQ ID 2

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

DIQNPDPAVY QLRDSKSSDK

20

<212> Tipo: PRT

<211> Comprimento: 20

Nome da sequência: SEQ ID 3

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

DPAVYQLRDS KSSDKSVCLF

20

<212> Tipo: PRT

<211> Comprimento: 20

Nome da sequência: SEQ ID 4

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:
NGKEVHSGVS TDPQPLKEQP 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 5

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ALNDSRYALS SRLRVSATFW 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 6

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
PPEVAVFEPS EAEISHTQKA 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 7

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
KEVHSGVSTD PQPLKEQPAL 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 8

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
VFPPEVAVFE PSEAEISHTQ 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 9

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

PYIQNPPEAV YQLKDPKSDQ ŠTLCLFTDFD SQINVPKIME SGTFITDKTV LDMKAMDSKS 60

NGAIAWSNQT SFTCQDIFKE TNATYPSSDV P 91

<212> Tipo: PRT
<211> Comprimento: 91

Nome da sequência: SEQ ID 10
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

EDLRNVTPPK VSLFEPKAE IANKQKATLV CLARGFFPDH VELSWVNGR EVHSGVSTDP	60
QAYKESNYSY CLSSRLRVSF TFVHNPRNHF RCQVQFHGLS EEDKWPEGSP KPVTONISAE	120
AWGRAD	126

<212> Tipo: PRT
<211> Comprimento: 126

Nome da sequência: SEQ ID 11
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

ESGTFITDKT VLDMKAMDSK 20

<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 12
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

KTMESGTFIT DKTVLDMKAM 20

<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 13
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

YIQNPEPAVY QLKDPRSQDS 20

<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 14
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

AVYQLKDPRS QDSTLCLFTD 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 15
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
NGREVHSGVS TDPQAYKESN 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 16
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
KESNYSYCLS SRLRVSATFW 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 17
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
PPKVSLFEPS KAEIANKQKA 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 18
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
REVHSGVSTD PQAYKESNYS 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 19
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
VTPPKVSLFE PSKAEIANKQ 20
<212> Tipo: PRT
<211> Comprimento: 20

Nome da sequência: SEQ ID 20

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

LLFGYPVYV

9

<212> Tipo: PRT

<211> Comprimento: 9

Nome da sequência: SEQ ID 21

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

SLLMITQC

8

<212> Tipo: PRT

<211> Comprimento: 8

Nome da sequência: SEQ ID 22

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

LLGRNSFEV

9

<212> Tipo: PRT

<211> Comprimento: 9

Nome da sequência: SEQ ID 23

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

KLVALGINAV

10

<212> Tipo: PRT

<211> Comprimento: 10

Nome da sequência: SEQ ID 24

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

LLGDLFGV

8

<212> Tipo: PRT

<211> Comprimento: 8

Nome da sequência: SEQ ID 25

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

FLRGRAYGL 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 26
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
HRCQAIRKK 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 27
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
YRSGIIAVV 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 28
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
VYGFVRACL 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 29
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ILAKFLHWL 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 30
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
LTLGEFLKL 9
<212> Tipo: PRT
<211> Comprimento: 9

Nome da sequência: SEQ ID 31

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
PKYVKQNTLK LA
<212> Tipo: PRT
<211> Comprimento: 12

12

Nome da sequência: SEQ ID 32
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
GILGFVFTL
<212> Tipo: PRT
<211> Comprimento: 9

9

Nome da sequência: SEQ ID 33
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
SLYNTVATL
<212> Tipo: PRT
<211> Comprimento: 9

9

Nome da sequência: SEQ ID 34
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
cacagacaaa tgtgtgctag acat
<212> Tipo: DNA
<211> Comprimento: 24

24

Nome da sequência: SEQ ID 35
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 35

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
atgtctagca cacatttgctc tgtg
<212> Tipo: DNA
<211> Comprimento: 24

24

Nome da sequência: SEQ ID 36
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 36

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

cagtgggggtc tgcacagacc c

21

<212> Tipo: DNA

<211> Comprimento: 21

Nome da sequência: SEQ ID 37

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 37

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gggtctgtgc agacccact

21

<212> Tipo: DNA

<211> Comprimento: 21

Nome da sequência: SEQ ID 38

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 38

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

taataatacg tataataata ttctatttca aggagacagt c

41

<212> Tipo: DNA

<211> Comprimento: 41

Nome da sequência: SEQ ID 39

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 39

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

caatccagcg gctgccgtag gcaataggta tttcattatg actgtctcct tgaaatag

58

<212> Tipo: DNA

<211> Comprimento: 58

Nome da sequência: SEQ ID 40

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 40

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ctacggcagc cgctggattg ttattactcg cggccagcc ggccatggcc cag

53

<212> Tipo: DNA

<211> Comprimento: 53

Nome da sequência: SEQ ID 41
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 41

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gttctgctcc acttccttct gggccatggc cggctgggcc g 41
<212> Tipo: DNA
<211> Comprimento: 41

Nome da sequência: SEQ ID 42
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 42

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
cagaaggaag tggagcagaa c 21
<212> Tipo: DNA
<211> Comprimento: 21

Nome da sequência: SEQ ID 43
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 43

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
cttcttaaag aattcttaat taacctaggt tattaggaac tttctgggct ggggaag 57
<212> Tipo: DNA
<211> Comprimento: 57

Nome da sequência: SEQ ID 44
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 44

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gttaattaag aattctttaa gaaggagata tacatatgaa aaaattatta ttcgcaattc 60
<212> Tipo: DNA
<211> Comprimento: 21

Nome da sequência: SEQ ID 45
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 45

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
cgcgctgtga gaatagaaag gaacaactaa aggaattgcg aataataatt ttttcatatg 60
<212> Tipo: DNA
<211> Comprimento: 60

Nome da sequência: SEQ ID 46
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 46

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ctttctattc tcacagcgcg caggctggtg tcactcagac 40
<212> Tipo: DNA
<211> Comprimento: 40

Nome da sequência: SEQ ID 47
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 47

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
atgatgtcta gatgcggccg cgtcgctct accccaggcc tc 42
<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 48
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 48

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gcatctagac atcatcacca tcatactag actgttgaaa gttgttttagc aaaac 55
<212> Tipo: DNA
<211> Comprimento: 55

Nome da sequência: SEQ ID 49
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 49

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ctagagggtg ccttattaag actccttatt acgcagtatg 40
<212> Tipo: DNA
<211> Comprimento: 40

Nome da sequência: SEQ ID 50

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 50

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

agctgcagct aatacgactc actataggaa caggccacca tggcgtcgat tatgctgagt 60

gatatccttg tccggtggta ccctag 86

<211> Comprimento: 86

Nome da sequência: SEQ ID 51

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 51

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gatcccatgg tggcctgttc ctatagtgcg tcgtattagc tgc 43

<212> Tipo: DNA

<211> Comprimento: 43

Nome da sequência: SEQ ID 52

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 52

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

agctgcagct aatacgactc actataggaa caggccacca tgg 43

<212> Tipo: DNA

<211> Comprimento: 43

Nome da sequência: SEQ ID 53

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 53

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ccaccatggg ccagaaggaa gtggagcaga actc 34

<212> Tipo: DNA

<211> Comprimento: 34

Nome da sequência: SEQ ID 54

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 54

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:
cgagagcccg tagaactgga cttg 24
<212> Tipo: DNA
<211> Comprimento: 24

Nome da sequência: SEQ ID 55

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 55

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gtggatccgg cgggtggcggg tcgaacgctg gtgtcactca gacccc 46
<212> Tipo: DNA
<211> Comprimento: 46

Nome da sequência: SEQ ID 56

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 56

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ccggatccac ctccgctga accgctcca ccggtgacca caacctgggt ccctg 55
<212> Tipo: DNA
<211> Comprimento: 55

Nome da sequência: SEQ ID 57

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 57

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ctgagaattc ttatgactct ccgcggttga agctc 35
<212> Tipo: DNA
<211> Comprimento: 35

Nome da sequência: SEQ ID 58

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 58

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
tgacgaattc tgactctccg cggttgaagc tc 32
<212> Tipo: DNA
<211> Comprimento: 32

Nome da sequência: SEQ ID 59

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 59

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgacgaattc tgactctccg cggttgaagc tc

32

<212> Tipo: DNA

<211> Comprimento: 32

Nome da sequência: SEQ ID 59

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 59

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

agctgcagct aatacgactc actatagg

28

<212> Tipo: DNA

<211> Comprimento: 28

Nome da sequência: SEQ ID 60

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 60

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ggccaccatg ggcaacgctg gtgtcactca gacccc

36

<212> Tipo: DNA

<211> Comprimento: 36

Nome da sequência: SEQ ID 61

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 61

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgaaccgcct ccaccgtctg ctctacccca ggcctcggcg

40

<212> Tipo: DNA

<211> Comprimento: 40

Nome da sequência: SEQ ID 62

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 62

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgactctccg cggttgaagc tc

22

<212> Tipo: DNA
<211> Comprimento: 22

Nome da sequência: SEQ ID 63
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 63

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
cagctggggg aagcttcagt ttggagcag
<212> Tipo: DNA
<211> Comprimento: 29

29

Nome da sequência: SEQ ID 64
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 64

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ctgctccaaa ctgaagcttc ccccagctg
<212> Tipo: DNA
<211> Comprimento: 29

29

Nome da sequência: SEQ ID 65
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 65

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gtacttctgt gcctcgaggc cgggactag
<212> Tipo: DNA
<211> Comprimento: 29

29

Nome da sequência: SEQ ID 66
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 66

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ctagtcccg cctcgaggca cagaagtac
<212> Tipo: DNA
<211> Comprimento: 29

29

Nome da sequência: SEQ ID 67
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 67

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

cattttcagg gatagcaagc

20

<212> Tipo: DNA

<211> Comprimento: 20

Nome da sequência: SEQ ID 68

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 68

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tcacacagga aacagctatg tcacacagga aacagctatg

40

<212> Tipo: DNA

<211> Comprimento: 40

Nome da sequência: SEQ ID 69

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 69

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggnknknkn knknknknkn cgaccagagc agtacttcg

49

<212> Tipo: DNA

<211> Comprimento: 49

Nome da sequência: SEQ ID 70

Descrição da sequência:

Característica

Sequência: SEQ ID 70:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 49

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 70

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggccgnknkn knknknknkn nnkccagagc agtacttcgg gc

52

<212> Tipo: DNA

<211> Comprimento: 52

Nome da sequência: SEQ ID 71

Descrição da sequência:

Característica

Sequência: SEQ ID 71:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 71:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 52

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 71

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggccggnknkn knnknknknkn nnkcgaccag agcagtactt cg 52

<212> Tipo: DNA

<211> Comprimento: 52

Nome da sequência: SEQ ID 72

Descrição da sequência:

Característica

Sequência: SEQ ID 72:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 52

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 72

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggccggnknkn knnknknknkn nnkggagggc gaccagagca g 51

<212> Tipo: DNA

<211> Comprimento: 51

Nome da sequência: SEQ ID 73

Descrição da sequência:

Característica

Sequência: SEQ ID 73:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 51

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 73

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggccgggann knknknknkn nnknkgggc gaccagagca gtac 54

<212> Tipo: DNA

<211> Comprimento: 54

Nome da sequência: SEQ ID 74

Descrição da sequência:

Característica

Sequência: SEQ ID 74:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 74:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 54

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 73:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 51

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 74

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggnnknnknn knnknnknnk ccagagcagt acttcgggc

49

<212> Tipo: DNA

<211> Comprimento: 49

Nome da sequência: SEQ ID 75

Descrição da sequência:

Característica

Sequência: SEQ ID 75:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 75:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 49

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 73:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 51

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 75

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tgtgcctcga ggnnknnknn knnknnknnk gagcagtact tcggggccg

48

<212> Tipo: DNA

<211> Comprimento: 48

Nome da sequência: SEQ ID 76

Descrição da sequência:

Característica

Sequência: SEQ ID 76:

<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 48

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 76

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
tgtgcctcga gggnknknkn knnknknknk cagtaccttcg ggccgggc 48
<212> Tipo: DNA
<211> Comprimento: 48

Nome da sequência: SEQ ID 77

Descrição da sequência:

Característica

Sequência: SEQ ID 77:

<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 48

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 77

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
tgtgcctcga ggccgnknkn knnknknkgg cgaccagagc agtacttcg 49
<212> Tipo: DNA
<211> Comprimento: 49

Nome da sequência: SEQ ID 78

Descrição da sequência:

Característica

Sequência: SEQ ID 78:

<221> Característica chave: "misc_feature"
<222> localização a partir de:
<222> Localização até:

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 78:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 49

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 78

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

aaactgaagc ttmnnmnnmn nmnnmntgt aacggcacag aggtag 46

<212> Tipo: DNA

<211> Comprimento: 46

Nome da sequência: SEQ ID 79

Descrição da sequência:

Característica

Sequência: SEQ ID 79:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 46

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 79

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

aaactgaagc ttmnnmnnngc tgctmnnmtgt aacggcacag aggtag 46

<212> Tipo: DNA

<211> Comprimento: 46

Nome da sequência: SEQ ID 80

Descrição da sequência:

Característica

Sequência: SEQ ID 80:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 46

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 80

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:
aaactgaagc ttmnnmnmnn ngctgtcmnn tgtaacggca cagaggtag 49
<212> Tipo: DNA
<211> Comprimento: 49

Nome da sequência: SEQ ID 81

Descrição da sequência:

Característica

Sequência: SEQ ID 81:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 81:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 49

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 81

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

aaactgaagc ttmnnmnnngc tgtcmnnaac ggcacagagg tag 43

<212> Tipo: DNA

<211> Comprimento: 43

Nome da sequência: SEQ ID 82

Descrição da sequência:

Característica

Sequência: SEQ ID 81:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 43

Outra informação: m é a ou c; n é a ou t ou g ou g; k é g
ou t

Liga sequência codificadora: Não

Codões por encomenda

Nome da sequência: SEQ ID 82

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cgggactagc gggagggcga ccagagcagt ag 42
<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 83
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 83

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ASRPGLAGGR PEQY 14
<212> Tipo: PRT
<211> Comprimento: 14

Nome da sequência: SEQ ID 84
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gcctcgaggc cggggctgat gagtgcgtag ccagagcagt ac 42
<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 85
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 85

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ASRPGLMSAX PEQY 14
<212> Tipo: PRT
<211> Comprimento: 14

Nome da sequência: SEQ ID 86
Descrição da sequência:

Característica

Sequência: SEQ ID 86:

<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber
Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gcctcgaggc cggggctgag gtcggcgtag ccagagcagt ac 42

<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 87
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 87

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ASRPGLRSAX PEQY
<212> Tipo: PRT
<211> Comprimento: 14

14

Nome da sequência: SEQ ID 88
Descrição da sequência:

Característica

Sequência: SEQ ID 88:
<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber
Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gcctcgaggc cgggactagc gggagggcga ccagaggcgt ag
<212> Tipo: DNA
<211> Comprimento: 42

42

Nome da sequência: SEQ ID 89
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 89

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ASRPGLAGGR PEAX
<212> Tipo: PRT
<211> Comprimento: 14

14

Nome da sequência: SEQ ID 90
Descrição da sequência:

Característica

Sequência: SEQ ID 90:
<221> Característica chave: "misc_feature"
<222> localização a partir de:
<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 90:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cgggactagc gggagggcga ccagaggatt ag 42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 91

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 91

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLAGGR PEDX 14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 92

Descrição da sequência:

Característica

Sequência: SEQ ID 92:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cgggactagc gggagggcga ccagatcagt ag 42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 93

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 93

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLAGGR PDQX

14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 94

Descrição da sequência:

Característica

Sequência: SEQ ID 94:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 94:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 94:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cgggtctgta ggctgggcga ccagagcagt ac

42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 95

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 95

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLXAGR PEQY

14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 96

Descrição da sequência:

Característica

Sequência: SEQ ID 96:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 96:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cggggctggt tccggggcga ccagagcagt ag

42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 97

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 97

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLVPGR PEQX

14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 98

Descrição da sequência:

Característica

Sequência: SEQ ID 98:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 98:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cggggcttgt gtctgcttag ccagagcagt ac

42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 99

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 99

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLVSAX PEQY

14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 100

Descrição da sequência:

Característica

Sequência: SEQ ID 100:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:
gcctcgaggc cgggactagc gggagggcga ccacatccgt ag 42
<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 101

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 101

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLAGGR PHPX 14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 102

Descrição da sequência:

Característica

Sequência: SEQ ID 102:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gcctcgaggc cgggactagc gggagggcga ccagatgcgt ag 42

<212> Tipo: DNA

<211> Comprimento: 42

Nome da sequência: SEQ ID 103

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 103

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ASRPGLAGGR PDAX 14

<212> Tipo: PRT

<211> Comprimento: 14

Nome da sequência: SEQ ID 104

Descrição da sequência:

Característica

Sequência: SEQ ID 102:

<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gcctcgaggc cgggtctgat tagtgcttag ccagagcagt ac 42
<212> Tipo: DNA
<211> Comprimento: 42

Nome da sequência: SEQ ID 105

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 105

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ASRPGLISAX PEQY 14
<212> Tipo: PRT
<211> Comprimento: 14

Nome da sequência: SEQ ID 106

Descrição da sequência:

Característica

Sequência: SEQ ID 106:

<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 14

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gccgttaca ctgacagctg ggggaagctt cag 33
<212> Tipo: DNA
<211> Comprimento: 33

Nome da sequência: SEQ ID 107

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 107

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

AVTTDSWGKL Q 11
<212> Tipo: PRT
<211> Comprimento: 14

Nome da sequência: SEQ ID 108
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gccgttaca ctgacagctg ggggccgctt cag 33
<212> Tipo: DNA
<211> Comprimento: 33

Nome da sequência: SEQ ID 109
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 109

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
AVTTDSWGKL Q 11
<212> Tipo: PRT
<211> Comprimento: 11

Nome da sequência: SEQ ID 110
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gccgttaca ctgacagctg ggggagatg cag 33
<212> Tipo: DNA
<211> Comprimento: 33

Nome da sequência: SEQ ID 111
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 111

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
AVTTDSWGKM Q 11
<212> Tipo: PRT
<211> Comprimento: 11

Nome da sequência: SEQ ID 112
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gccgttaca ctgacagctg ggggaagttg cat 33
<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 113

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 113

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTTDSWGKL H

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 114

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggtagctt cat

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 115

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 115

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTTDSWGXL H

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 116

Descrição da sequência:

Característica

Sequência: SEQ ID 116:

<221> Característica chave: "misc_feature"

<222> localização a partir de:

<222> Localização até:

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Característica

Sequência: SEQ ID 116:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 11

Outra informação: x significa um resíduo de aminoácidos
codificado por um codão amber

Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggaggctg cat

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 119

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 119

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTTDSWGRL H

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 120

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggcagctt cat

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 121

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 115

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTTDSWGQL H

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 122

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggaaggtt cat

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 123

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 123

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTIDSWGKV H

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 124

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggaaggtg aat

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 125

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 125

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTIDSWGKV N

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 126

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccgttaca ctgacagctg ggggaagctt ctg

33

<212> Tipo: DNA

<211> Comprimento: 33

Nome da sequência: SEQ ID 127

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 127

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AVTIDSWGKL L

11

<212> Tipo: PRT

<211> Comprimento: 11

Nome da sequência: SEQ ID 128
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
ggaattcatc gatgcagaag gaagtggagc ag 32
<212> Tipo: DNA
<211> Comprimento: 32

Nome da sequência: SEQ ID 129
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 129
Sequência
<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
gtacacggcc gggtcagggt tctggatata c 31
<212> Tipo: DNA
<211> Comprimento: 31

Nome da sequência: SEQ ID 130
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 130
Sequência
<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
tctctcatta atgaatgctg gtgtcactca gacccc 36
<212> Tipo: DNA
<211> Comprimento: 36

Nome da sequência: SEQ ID 131
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 131
Sequência
<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:
tagaaaccgg tggccaggca caccagtgtg gc 32
<212> Tipo: DNA
<211> Comprimento: 32

Nome da sequência: SEQ ID 132
Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 132
Sequência
<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

ctgctctggt tccgcactc 19

<212> Tipo: DNA

<211> Comprimento: 19

Nome da sequência: SEQ ID 133

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 133

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gagtgcggaa ccagagcag

19

<212> Tipo: DNA

<211> Comprimento: 19

Nome da sequência: SEQ ID 134

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 134

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

atatccagaa cccggatcct gccgtgta

28

<212> Tipo: DNA

<211> Comprimento: 28

Nome da sequência: SEQ ID 135

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 135

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tacacggcag gaatccgggt tctggatat

29

<212> Tipo: DNA

<211> Comprimento: 29

Nome da sequência: SEQ ID 136

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 136

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ggagatatac atatgcagga ggtgacacag

30

<212> Tipo: DNA

<211> Comprimento: 30

Nome da sequência: SEQ ID 137

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 137

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tacacggcag gatccgggtt ctggatatt

29

<212> Tipo: DNA

<211> Comprimento: 29

Nome da sequência: SEQ ID 138

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 138

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ggagatatac atatgggtgt cactcagacc

30

<212> Tipo: DNA

<211> Comprimento: 30

Nome da sequência: SEQ ID 139

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 139

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

cccaagctta gtctgctcta ccccaggcct cggc

34

<212> Tipo: DNA

<211> Comprimento: 34

Nome da sequência: SEQ ID 140

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 140

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gccggccatg gccaaacagg aggtgacgca gattcct

37

<212> Tipo: DNA

<211> Comprimento: 37

Nome da sequência: SEQ ID 141

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 141

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

cttcttaaag aattcttaat taacctaggt tattaggaac tttctgggct ggggaag

57

<212> Tipo: DNA

<211> Comprimento: 57

Nome da sequência: SEQ ID 142

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 142

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tcacagcgcg caggctggtg tcactcagac cccaaa

36

<212> Tipo: DNA

<211> Comprimento: 36

Nome da sequência: SEQ ID 143

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 143

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

cgagagcccg tagaactgga cttg

24

<212> Tipo: DNA

<211> Comprimento: 24

Nome da sequência: SEQ ID 144

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 144

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tctctcatta atgaatgctg gtgtcactca gacccc

36

<212> Tipo: DNA

<211> Comprimento: 36

Nome da sequência: SEQ ID 145

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 145

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

tagaaaccgg tggccaggca caccagtgtg g

31

<212> Tipo: DNA

<211> Comprimento: 31

Nome da sequência: SEQ ID 146

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 146

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

PE1558643

- 162 -

```

atgcagaagg aagtggagca gaactctgga cccctcagtg ttccagaggg agccattgcc      60
tctctcaact gcacttacag tgaccgaggt tcccagtcct tcttctggta cagacaatat      120
tctgggaaaa gccctgagtt gataatgtcc atttactcca atggtgacaa agaagatgga      180
aggtttacag cacagctcaa taaagccagc cagtatgttt ctctgctcat cagagactcc      240
cagcccagtg attcagccac ctacctctgt gccgttacaa ctgacagctg ggggaaattg      300
cagtttgag cagggaccca ggttggtggt accccagata tccagaacct tgacctgcc      360
gtgtaccagc tgagagactc taaatccagt gacaagtctg tctgcctatt caccgatttt      420
gattctcaaa caaatgtgtc acaaagtaag gattctgatg tgtatatcac agacaaatgt      480
gtgctagaca tgaggtctat ggacttcaag agcaacagtg ctgtggcctg gagcaacaaa      540
tctgactttg catgtgcaaa cgccctcaac aacagcatta ttccagaaga caccttcttc      600
cccagcccag aaagttccta a                                     621

```

<212> Tipo: DNA

<211> Comprimento: 621

Nome da sequência: SEQ ID 147

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 147

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

atgaacgctg gtgtcactca gacccccaaa ttccaggtcc tgaagacagg acagagcatg      60
acactgcagt gtgcccagga tatgaaccat gaatacatgt cctggatcgc acaagaccca      120
ggcatggggc tgaggctgat tcattactca gttgggtgctg gtatcactga ccaaggagaa      180
gtccccaatg gctacaatgt ctccagatca accacagagg atttcccgtc caggctgctg      240
tcggctgctc cctcccagac atctgtgtac ttctgtgcca gcaggccggg actagcggga      300
gggcgaccag agcagtactt cgggcccggc accagggtca cggtcacaga ggacctgaaa      360
aacgtgttcc cacccgaggt cgtgtgttt gagccatcag aagcagagat ctcccacacc      420
caaaaggcca cactggtgtg cctggccaca ggcttctacc ccgaccacgt ggagctgagc      480
tgggtgggtg atgggaagga ggtgcacagt ggggtctgca cagaccgca gcccccaag      540
gagcagcccg cctcaatga ctccagatac gctctgagca gccgcctgag ggtctcggcc      600
accttctggc aggacccccg caaccacttc cgctgtcaag tccagttcta cgggctctcg      660
gagaatgacg agtggaccca ggatagggcc aaaccgctca cccagatcgt cagcgccgag      720
gcctggggta gagcagacta a                                     741

```

<212> Tipo: DNA

<211> Comprimento: 7411

Nome da sequência: SEQ ID 148

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 148

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MQKEVEQNSG PLSVPEGAIA SLNCTYSDRG SQSFFWYRQY SGKSPELIMS IYSNGDKEDG	60
RFTAQLNKAS QYVSLLRDS QPSDSATYLC AVTTDSWGKL QFGAGTQVVV TPDIQNPDP	120
VYQLRDSKSS DKSVCFLTDF DSQTNSVQSK DSDVYITDKC VLDMRSNDFK SNSAVAWSNK	180
SDFACANAFN NSIIPEDTFF PSPESS	206

<212> Tipo: PRT

<211> Comprimento: 206

Nome da sequência: SEQ ID 149

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE	60
VPNGYNVSRS TTEDFPLRLI SAAPSQTSVY FCASRPLAG GRPEQYFGPG TRLTVTEDLK	120
NVFPPEVAVF EPSEAEISHT QKATLVCLAT GFYPOHVELS WWVNGKEVHS GVCTDPOPLK	180
EQPALNDSRY ALSSRLRVSA TFWQDPRNHF RCQVQFYGLS ENDEWTQDRA KPVQTIVSAE	240
AWGRAD	246

<212> Tipo: PRT

<211> Comprimento: 246

Nome da sequência: SEQ ID 150

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

gttaactacg tcaggtggca cttttcgggg aaatgtgagc ggaacccta tttgtttatt	60
tttctaaata cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca	120
ataatattga aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattccctt	180
ttttcgggca ttttgccttc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga	240
tgctgaagat cagttgggtg cagcagtggtg ttacatcgaa ctggatctca acagcggtaa	300
gatccttgag agttttcgcc ccgaagaacg ttctccaatg atgagcactt ttaaagtctt	360
gctatgtggc gcggtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgcat	420
acactattct cagaatgact tgggtgagta ctaccagtc acagaaaagc atcttacgga	480
tggcatgaca gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc	540
caacttactt ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcaaacat	600
gggggatcat gtaactcgcc ttgatcggtg ggaaccggag ctgaatgaag ccataccaaa	660
cgacgagcgt gacaccacga tgccgttagc aatggcaaca acgttgcgca aactattaac	720

tggcgaacta cttactctag cttcccggca acaattaata gactggatgg aggcggataa	780
agttgcagga ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc	840
tggagccggt gagcgtgggt ctgcgggtat cattgcagca ctggggccag atggttaagcc	900
ctcccgatc gtagtattct acacgacggg gagtcaggca actatggatg aacgaaatag	960
acagatcgct gagatagggtg cctcactgat taagcattgg taactgtcag accaagttta	1020
ctcatatata ctttagattg atttaccctg gttgataatc agaaaagccc caaaaacagg	1080
aagattgtat aagcaaatat tttaaattgta aacgttaata ttttgttaaa attcgcgtta	1140
aatttttgtt aaatcagctc attttttaac caataggccg aaatcggcaa aatcccttat	1200
aaatcaaaag aatagcccga gatagggttg agtgttggtc cagtttggaa caagagtcca	1260
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc	1320
ccactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta	1380
aatcgggaacc ctaaaggag ccccgattt agagcttgac ggggaaagcg aacgtggcga	1440
gaaagggaagg gaagaaagcg aaaggagcgg gcgctagggc gctggcaagt gtagcggcca	1500
cgctgcgcgt aaccaccaca cccgccgcgc ttaatgcgc gctacagggc gcgtaaaagg	1560
atctaggtda agatcctttt tgataatctc atgaccaaaa tcccttaacy tgagttttcg	1620
ttccactgag cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt	1680
ctgcgcgtaa tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg	1740
ccgatcaag agctaccaac tctttttccg aaggtaactg gcttcagcag agcgcagata	1800
ccaaatactg ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca	1860
ccgcctacat acctcgctct gctaatcctg ttaccagtgg ctgctgccag tggcgataag	1920
tcgtgtctta ccgggttggg ctcaagacga tagttaccgg ataaggcgca gcggtcgggc	1980
tgaacggggg gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga	2040
tacctacagc gtgagctatg agaaagcgcc acgcttcccg aaggagaaaa ggcggacagg	2100
tatccggtaa gcggcagggg cgggaacagga gagcgcacga gggagcttcc agggggaaac	2160
gcctggtatc tttatagtcc tgtcgggttt cgccacctct gacttgagcg tcgatttttg	2220
tgatgctcgt cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg	2280
ttcctggcct tttgctggcc ttttctcac atgtaatgtg agttagctca ctcattagcc	2340
acccagget ttacacttta tgcttccggc tcgtatgttg tgtggaattg tgagcggata	2400
acaatttcac acaggaaaca gctatgacca tgattacgcc aagctacgta taataatatt	2460
ctatttcaag gagacagtca taatgaaata cctattgcct acggcagccg ctggattgtt	2520
attactcgcg gccagccgg ccatggccca gaagggaagt gagcagaact ctggaccctt	2580
cagtgttcca gagggagcca ttgcctctct caactgcact tacagtgacc gaggttccca	2640
gtccttcttc tggtagagac aatattctgg gaaaagccct gagttgataa tgtccatata	2700
ctccaatggg gacaaagaag atggaagggt tacagcacag ctcaataaag ccagccagta	2760

tgtttctctg	ctcatcagag	actcccagcc	cagtgaattca	gccacctacc	tctgtgccgt	2820
tacaactgac	agctggggga	aattgcagtt	tggagcaggg	accaggttg	tggtcacccc	2880
agatatecag	aaccctgacc	ctgccgtgta	ccagctgaga	gactctaaat	ccagtgacaa	2940
gtctgtctgc	ctattcaccg	attttgatcc	tcaaacaat	gtgtcacaaa	gtaaggattc	3000
tgatgtgtat	atcacagaca	aatgtgtgct	agacatgagg	tctatggact	tcaagagcaa	3060
cagtgtctgt	gcctggagca	acaaatctga	ctttgcatgt	gcaaacgcct	tcaacaacag	3120
cattattcca	gaagacacct	tcttccccag	cccagaaagt	tcctaataac	ctagggttaat	3180
taagaattct	ttaagaagga	gatatacata	tgaaaaaatt	attattcgca	attcctttag	3240
ttgttccctt	ctattctcac	agcgcgcagg	ctggtgtcac	tcagacccca	aaattccagg	3300
tctgaagac	aggacagagc	atgacactgc	agtgtgcccc	ggatatgaac	catgaatata	3360
tgtcctggta	tcgacaagac	ccaggcatgg	ggctgaggct	gattcattac	tcagttgggt	3420
ctggtatcac	tgaccaagga	gaagtcccca	atggctacaa	tgtctccaga	tcaaccacag	3480
aggatttccc	gtcagggtcg	ctgtcggctg	ctccctccca	gacatctgtg	tacttctgtg	3540
ccagcaggcc	gggactagcg	ggagggcgac	cagagcagta	cttcgggccc	ggcaccaggc	3600
tcacggtcac	agaggacctg	aaaaacgtgt	tcccaccgga	ggtcgtctgt	tttgagccat	3660
cagaagcaga	gatctcccac	acccaaaagg	ccacactggt	gtgcctggcc	acaggcttct	3720
accccgacca	cgtggagctg	agctgggtgg	tgaatgggaa	ggaggtgcac	agtgggtctt	3780
gcacagaccc	gcagccccct	aaggagcagc	ccgcccctca	tgactccaga	tacgctctga	3840
gcagccgect	gagggctctg	gccaccttct	ggcaggaccc	ccgcaaccac	ttccgctgtc	3900
aagtcacagt	ctacgggctc	tcggagaatg	acgagtggac	ccaggatagg	gccaaacccg	3960
tcacccagat	cgtcagcgcc	gaggcctggg	gtagagcaga	cgcgcccgca	tctagacatc	4020
atcaccatca	tcactagact	gttgaaagtt	gttttagcaa	accccatata	gaaaattcat	4080
ttactaacgt	ctggaaagac	gacaaaactt	tagatcgtta	cgctaactat	gagggttgtc	4140
tgrggaatgc	tacagggcgt	gtagtttgta	ctggtgacga	aactcagtgt	tacggtacat	4200
gggttccctat	tgggcttgct	atccctgaaa	atgaggggtg	tggctctgag	ggtggcggtt	4260
ctgaggggtg	cggttctgag	ggtggcggtg	ctaaacctcc	tgagtacggt	gatacaccta	4320
ttccggggcta	tacttatatc	aaccctctcg	acggcactta	tccgcctggt	actgagcaaa	4380
accccgctaa	tcctaatact	tctcttgagg	agtctcagcc	tcttaatact	ttcatgtttc	4440
agaataatag	gttccgaaat	aggcaggggg	cattaactgt	ttatacgggc	actgttactc	4500
aaggcactga	ccccgttaaa	acttattacc	agtacactcc	tgtatcatca	aaagccatgt	4560
atgacgctta	ctggaacggg	aaattcagag	actgcgcttt	ccattctggc	tttaatgagg	4620
atccattcgt	ttgtgaatat	caaggccaat	cgtctgacct	gcctcaacct	cctgtcaatg	4680
ctggcgggcg	ctctggtggt	ggttctggtg	gcggctctga	gggtgggtgg	tctgaggggtg	4740
gcggttctga	gggtggcggc	tctgaggagg	gcggttccgg	tgggtggctct	ggttccgggtg	4800

```

atTTTgatta tGaaaagatg gCaaacgcta ataagggggc tatgaccgaa aatgccgatg 4860
aaaacgcgct acagtctgac gctaaaggca aacttgattc tgcgctact gattacggtg 4920
ctgctatcga tggtrtcatt ggtgacgttt cCggccttgc taatggtaat ggtgctactg 4980
gtgattttgc tggctctaata tcccaaattg ctcaagtcgg tgacggtgat aattcacctt 5040
taatgaataa tttccgtcaa tatttacctt cctccctca atcgggtgaa tgcgccccct 5100
ttgtcttttag cgctggtaaa ccatatgaat tttctattga ttgtgacaaa ataaacttat 5160
tccgtgggtg ctttgcgttt cttttatatg ttgccacctt tatgtatgta tttctacgt 5220
ttgctaacat actgcgtaat aaggagtctt aataaggtag cctctagtca aggcctatag 5280
tgagtcgtat tacggactgg cCgtcgtttt acaacgtcgt gactgggaaa accctggcgt 5340
taccacactt aatcgccctg cagcacatcc cCtttcgcc agctggcgta atagcgaaga 5400
ggccgcacc gatcgccctt cccaacagtt gcgcagctg aatggcgaat ggcgcttcgc 5460
ttggaataa agcccgcttc ggcgggcttt ttttt 5495

```

<212> Tipo: DNA

<211> Comprimento: 5495

Nome da sequência: SEQ ID 151

Descrição da sequência:

Códons por encomenda

Nome da sequência: SEQ ID 151

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

agctgcagct aatacgactc actataggaa caggccacca tgggccagaa ggaagtggag 60
cagaactctg gacccctcag tgttccagag ggagccattg cctctctcaa ctgcacttac 120
agtgaccgag gttcccagtc cttcttctgg tacagacaat attctgggaa aagccctgag 180
ttgataatgt ccatatactc caatggtgac aaagaagatg gaaggtttac agcacagctc 240
aataaagcca gccagtatgt ttctctgctc atcagagact cccagcccag tgattcagcc 300
acctacctct gtgccgttac aactgacagc tgggggaaat tgcagtttgg agcagggacc 360
caggttgtgg tcaccggtgg aggcggttca ggcggaggtg gatccggcgg tggcgggtcg 420
aacgctggty tcaactcagac cccaaaattc caggtcctga agacaggaca gagcatgaca 480
ctgcagtgtg cccaggatat gaaccatgaa tacatgtcct ggtatcgaca agaccagggc 540
atggggctga ggctgattca ttactcagtt ggtgctggta tcactgacca aggagaagtc 600
cccaatggct acaatgtctc cagatcaacc acagaggatt tcccgtcag gctgctgtcg 660
gctgctccct cccagacatc tgtgtacttc tgtgccagca ggccgggact agcgggaggg 720
cgaccagagc agtacttcgg gccgggcacc aggttcacgg tcacagagga cctgaaaaac 780
gtgttccca ccgaggtcgc tgtgtttgag ccatcagaag cagagatctc ccacacccaa 840
aagggcacac tgggtgtgctt ggccacaggc ttctacccc accacgtgga gctgagctgg 900
tgggtgaatg ggaaggaggt gcacagtggg gtcagcacag acccgagcc cctcaaggag 960

```

```

cagccccccc tcaatgactc cagatacgct ctgagcagcc gcctgagggg ctcggccacc 1020
ttctggcagg acccccgcaa ccacttccgc tgtcaagtcc agttctacgg gctctcggag 1080
aatgacgagt ggaccagga tagggccaaa cccgtcacc agatcgtcag cggcgaggcc 1140
tggggtagag cagacgggtg aggcgggtca ctcagcagca ccctgacgct gagcaaagca 1200
gactacgaga aacacaaagt ctacgcctgc gaagtcaccc atcagggcct gagttcgccc 1260
gtcacaaaaga gtttcaaccg cggagagtca taagaattct cag 1303

```

<212> Tipo: DNA

<211> Comprimento: 1303

Nome da sequência: SEQ ID 152

Descrição da sequência:

Códigos por encomenda

Nome da sequência: SEQ ID 152

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

MGQKEVEQNS GPLSVPEGAI ASLNCTYSDR GSQSFFWYRQ YSGKSPELIM SIYSNGDKED 60
GRFTAQLNKA SQYVSLLRD SPDSATYLC AVTTDSWGKL QFGAGTQVVV TGGGSGGGG 120
SGGGGSNAGV TQTPKFQVLK TGQSMTLQCA QDMNHEYMSW YRQDPGMGLR LIHYSVGAGI 180
TDQGEVPNGY NVSRSTEDF PLRLLSAAPS QTSVYFCASR PGLAGGRPEQ YFGPGTRLTV 240
TEDLKNVFPP EVAVFEPSEA EISHTQKATL VCLATGFYPD HVLSWWVNG KEVHSGVSTD 300
PQPLKEQPAL NDSRYALSSR LRVSATFWQD PRNHFRQVQ FYGLSENDEW TQDRAKPVTV 360
IVSAEAWGRA DGGGGSLSST LTLSKADYEK HKVYACEVTH QGLSSPVTKS FNRGES 416

```

<212> Tipo: PRT

<211> Comprimento: 146

Nome da sequência: SEQ ID 153

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

ctgctttccc ggagcactat gcggataaaa atatccaatt acagtactat tattacaaaa 60
gaatctgcag tccaccgtga aaagccctt tacacgcgcc ttggggataa acaataaaaa 120
agatttatgt aagtttatac ataggcgagt actctgttat tgggactatt tacgaagtta 180
ttataacttt ttctttctca tactcataag ttgtaaaggc acagcgggaa taagggaaaa 240
aacgccgtaa aacggaagga caaaaacgag tgggtctttg cgaccacttt cattttctac 300
gacttctagt caaccacgt gtcacccaa ttagcttga cctagagttg tcgccattct 360
aggaactctc aaaagcggg cttcttgcaa aaggttacta ctcgtgaaaa ttcaagacg 420
atacaccgcg ccataatagg gcataactgc ggccgttct cgttgagcca gggcggtatg 480
tgataagagt cttactgaac caactcatga gtggtcagtg tcttttcgta gaatgectac 540
cgtactgtca ttctcttaat acgtcacgac ggtattggta ctcactattg tgacgccggt 600

```

tgaatgaaga ctgttgctag cctcctggct tccicgattg gcgaaaaaac gtgttgtagc	660
ccctagtaca ttgagcggaa ctagcaaccc ttggcctcga cttacttcgg tatggrrtgc	720
tgctcgcaact gtggtgctac ggacatcggt accgttggtg caacgcgttt gataattgac	780
cgcttgatga atgagatcga agggccggtt ttaattatct gacctaccct cgctattttc	840
aacgtcctgg tgaagacgag agccgggaag gccgaccgac caaataacga ctatttagac	900
ctcggccact cgcacccaga gcgccatagt aacgtcgtga ccccggtcta ccattcggga	960
gggcatagca tcaatagatg tgctgcccct cagtcggtg atacctactt gctttatctg	1020
tctagcgact ctatccacgg agtgactaat tcgtaaccat tgacagtctg gttcaaatga	1080
gtatatatga aatctaacta aattttgaag taaaaattaa attttcctag atccacttct	1140
agggaaaaact attagagtac tggttttagg gaattgcact caaaagcaag gtgactcgca	1200
gtctggggca tcttttctag tttcctagaa gaactctagg aaaaaagac gcgcattaga	1260
cgacgaacgt ttgttttttt ggtggcgatg gtcgccacca aacaaacggc ctagttctcg	1320
atggttgaga aaaaggcttc cattgaccga agtcgtctcg cgtctatggt ttatgacagg	1380
aagatcacat cggcatcaat ccggtggtga agttcttgag acatcggtgc ggatgtatgg	1440
agcgagacga ttaggacaat ggtcaccgac gacggtcacc gctattcagc acagaatggc	1500
ccaacctgag ttctgctatc aatggcctat tccgcgtcgc cagcccgact tgcccccaa	1560
gcacgtgtgt cgggtcgaa ctcgcttgcg ggatgtggct tgactctatg gatgtcgac	1620
tcgatactct ttgcggtgc gaagggcttc cctctttccg cctgtccata ggccattcgc	1680
cgtcccagcc ttgtcctctc gcgtgctccc tcgaaggtec ccttttgcgg accatagaaa	1740
tatcaggaca gcccaaagcg gtggagactg aactcgcagc taaaaacact acgagcagtc	1800
ccccgcctc ggataccttt ttgcggtcgt tgcgccgaa aaatgccaa gaccggaaaa	1860
cgaccggaaa acgagtgtac aagaaaggac gcaatagggg actaagacac ctattggcat	1920
aatggcggaa actcactcga ctatggcgag cggcgctggc ttgctggctc gcgtcgctca	1980
gtcactcgct ccttcgcctt ctcgcggtt atgcgtttgg cggagagggg cgcgcaaccg	2040
gctaagtaat tacgtcgacc gtgctgtcca aagggtgac ctttcgcccg tcaactcgct	2100
tgcgttaatt acactcaatc gagtgagtaa tccgtgggt ccgaaatgtg aaatacgaag	2160
gccgagcata caacacacct taacactcgc ctattgttaa agtgtgtcct ttgtcgatac	2220
tggtactaat gcggttcgac gtcgattatg ctgagtata tcctgtccg gtggtacctt	2280
agggggccat ggctcgagct taagtaccg gcagcaaat gttgcagcac tgaccctttt	2340
gggaccgcaa tgggttgaat tagcggaaag tcgtgtaggg ggaaagcggg cgaccgcatt	2400
atcgcttctc cgggcgtggc tagcgggaag ggttgtaaac gcgtcggact taccgcttac	2460
cgcggactac gccataaaag aggaatgcgt agacacgcca taaagtgtgg cgtataccac	2520
gtgagagtca tgtagacga gactacggcg tatcaattcg gtcggggctg tgggcgggtg	2580
tgggcgactg cgcgggactg cccgaacaga cgagggccgt aggcgaatgt ctgttcgaca	2640
ctggcagagg ccctcgacgt acacagtctc caaaagtggc agtagtggct ttgcgcgct	2699

<212> Tipo: DNA

<211> Comprimento: 2699

Nome da sequência: SEQ ID 154

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 154

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

ccatgggccca gaaggaagtg gagcagaact ctggacccct cagtgttcca gaggagacca      60
ttgectctct caactgcact tacagtgacc gaggttccca gtccttcttc tggtagacagac      120
aatattcttg gaaaagccct gagttgataa tgtccatata ctccaatggt gacaaagaag      180
atggaagggtt tacagcacag ctcaataaag ccagccagta tgtttctctg ctcatcagag      240
actcccagcc cagtgtattca gccacctacc tctgtgccgt tacaactgac agctggggga      300
aattgcagtt tggagcaggg acccaggttg tggtcaccgg tggaggcggg tcaggcggag      360
gtggatccgg cgggtggcggg tcgaacgctg gtgtcactca gaccccaaaa ttccagggtcc      420
tgaagacagg acagagcatg aactgcagt gtgcccagga tatgaacat gaatacatgt      480
cctggtatcg acaagaccca ggcatggggc tgaggctgat tcattactca gttggtgctg      540
gtatcactga ccaaggagaa gtccccaatg gctacaatgt ctccagatca accacagagg      600
atttcccgct caggctgctg tcggctgctc cctcccagac atctgtgtac ttctgtgccca      660
gcaggccggg actagcggga gggcgaccag agcagtactt cgggcccggc accaggctca      720
cggtcacaga ggacctgaaa aacgtgttcc caccgaggt cgctgtgttt gagccatcag      780
aagcagagat ctcccacacc caaaaggcca cactggtgtg cctggccaca ggcttctacc      840
ccgaccacgt ggagctgagc tgggtgggtga atgggaagga ggtgcacagt ggggtcagca      900
cagaccgca gcccctcaag gagcagcccg cctcaatga ctccagatac gctctgagca      960
gccgcctgag ggtctcggcc accttctggc aggacccccg caaccacttc cgctgtcaag      1020
tccagttcta cgggctctcg gagaatgacg agtggaccca ggataggggc aaaccctca      1080
cccagatcgt cagcggcgag gcctggggta gagcagacgg tggaggcggg tcaactcagca      1140
gcaccctgac gctgagcaaa gcagactacg agaaacacaa agtctacgcc tgcgaagtca      1200
cccatcaggg cctgagttcg ccgctcacia agagcttcaa ccgcgagag tcataagaat      1260
tc                                                                                   1262

```

<212> Tipo: DNA

<211> Comprimento: 1262

Nome da sequência: SEQ ID 155

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 155

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

gctggtgtca ctcagacccc aaaattccag gtcctgaaga caggacagag catgacactg      60
cagtgtgccc aggatatgaa ccatgaatac atgtcctggt atcgacaaga cccaggcatg      120
gggctgagggc tgattcatta ctcagttggt gctggtatca ctgaccaagg agaagtcccc      180
aatggctaca atgtctccag atcaaccaca gaggatttcc cgctcaggct gctgtcggct      240
gtctccctccc agacatctgt gtacttctgt gccagcagggc cgggactagc gggaggggtga      300
ccagagcagt acttcggggc gggcaccagg ctcacgggtca cagaggacct gaaaaacgtg      360
ttcccacccg aggtcgctgt gtttgagcca tcagaagcag agatctccca caccctaaaag      420
gccacactgg tgtgctggc cacaggcttc taccctgacc acgtggagct gagctggtgg      480
gtgaatggga aggaggtgca cagtggggtc tgcacagacc cgcagccct caaggagcag      540
cccgccttca atgactccag atacgtctg agcagccgccc tgaggggtctc ggccaccttc      600
tggcaggacc cccgcaacca cttccgctgt caagtccagt tctacgggct ctcggagaat      660
gacgagtgga cccaggatag ggccaaaccc gtcaccacaga tcgtcagcgc cgaggcctgg      720
ggtagagcag ac                                                                732

```

<212> Tipo: DNA

<211> Comprimento: 732

Nome da sequência: SEQ ID 156

Descrição da sequência:

Códones por encomenda

Nome da sequência: SEQ ID 156

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

AGVTQTPKFQ VLKYGQSMFL QCAQDMNEY MSWYRQDPGM GLRLIHYSVG AGITDQGEVP      60
NGYNVSRSTT EDFPLRLLSA APSQTSVYFC ASRPGLAGGX PEQYFGPGTR LTVTEDLKNV      120
FPPEVAVFEP SEAEISHTQK ATLVCLATGF YPDHVELSWW VNGKEVHSGV CTDPPQLKEQ      180
PALNDSRYAL SSRLRVSATF WQDPRNHFRG QVQFYGLSEN DEWTQDRAKP VTQIVSAEAW      240
GRAD                                                                244

```

<212> Tipo: PRT

<211> Comprimento: 244

Nome da sequência: SEQ ID 157

Descrição da sequência:

Característica

Sequência: SEQ ID 157:

<221> Característica chave: "misc_feature"

<222> localização a partir de: 1

<222> Localização até: 244

Outra informação: x significa um resíduo de aminoácido codificado por um codão opal, este é normalmente traduzido como um triptofano

Liga sequência codificadora: Não

<400> segmento de pré-sequência:

Descrição da sequência:

<400> segmento de pré-sequência:

gctgggtgtca ctcagacccc aaaattccag gtcctgaaga caggacagag catgacactg	60
cagtgtgtccc aggatatgaa ccatgaatac atgtccttgg atcgacaaga cccaggcatg	120
gggctgagggc tgattcatta ctcagttggg gctgggtatca ctgaccaagg agaagtcccc	180
aatggctaca atgtctccag atcaaccaca gaggatttcc cgctcaggct gctgtcggct	240
gtccccctccc agacatctgt gtacttctgt gcctcgaggc cggggctgat gagtgcggaa	300
ccagagcagt acttcggggc gggcaccagg ctacaggcca cagaggacct gaaaaacgtg	360
ttcccaccctg aggtcgctgt gtttgagcca tcagaagcag agatctccca cacccaaaag	420
gccacactgg tgtgcctggc caccggtttc taccctgacc acgtggagct gagctggtgg	480
gtgaatggga aggaggtgca cagtggggtc tgcacagacc cgagccccct caaggagcag	540
cccgccctca atgactccag atacgctctg agcagccgcc tgagggtctc ggccaccttc	600
tggcaggacc cccgcaacca ctctcgctgt caagtccagt tctacgggct ctcgagagaat	660
gacgagtggga cccaggatag ggccaaaccc gtcaccacaga tcgtcagcgc cgaggccctgg	720
ggtaagagcag actaagcttg aattc	745

<212> Tipo: DNA

<211> Comprimento: 745

Nome da sequência: SEQ ID 159

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 159

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSHYRQDP GNGLRLIHYS VGAGITDQGE	60
VPNGYNVSR S TTEDFPLRL L SAAPSQTSVY FCASRPGLMS AEPEQYFGPG TRLTVTEDLK	120
NVFPPPEVAVF EPSEAEISHT QKATLVCLAT GFYPDHVELS WWVNGKEVHS GVCTDPQPLK	180
EQPALNDSRY ALSSRLRVSA TFWQDPRNHF RCQVQFYGLS ENDEWTQDRA KPVQTIVSAE	240
AWGRAD	246

<212> Tipo: PRT

<211> Comprimento: 246

Nome da sequência: SEQ ID 160

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

AGVTQTPKFQ VLKTGQSM TL QCAQDMNHEY MSWYRQDPGM GLRLIHYSVG AGITDQGEVP	60
NGYNVSRSTT EDFPLRLLSA APSQTSVYFC ASRPGLMSAQ PEQYFGPGTR LTVTEDLKNV	120
FPPEVAVFEP SEAEISHTQK ATLVCLATGF YPDHVELSWW VNGKEVHSGV CTDPOPLKEQ	180
PALNDSRYAL SSRLRVSATF WQDPRNHFRC QVQFYGLSEN DEWTQDRAXP VTQIVSAEAW	240
GRAD	244

<212> Tipo: PRT

<211> Comprimento: 244

Nome da sequência: SEQ ID 161

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

atgcaggagg tgacacagat tcccgagct ctgagtgtcc cagaaggaga aaacttggtt      60
ctcaactgca gtttcaactga tagcgctatt tacaacctcc agtgggttag gcaggaccct    120
gggaaagggtc tcacatctct gttgcttatt cagtcaagtc agagagagca aacaagtggg    180
agacttaatg cctcgctgga taaatcatca ggacgtagta ctttatacat tgcagcttct    240
cagcctggtg actcagccac ctacctctgt gctgtgaggg ccacatcagg aggaagctac    300
atacctacat ttggaagagg aaccagcctt attgttcac cgtatatcca gaacctgac     360
cctgccgtgt accagctgag agactctaaa tccagtgaca agtctgtctg cctattcacc    420
gattttgatt ctcaaacaaa tgtgtcaca agtaaggatt ctgatgtgta tatcacagac    480
aaatgtgtgc tagacatgag gtctatggac ttcaagagca acagtgtgtg ggccctggagc    540
aacaatctg actttgcatg tgcaaacgcc ttcaacaaca gcattattcc agaagacacc    600
tttttcccca gccagaaaag ttcttaa                                     627

```

<212> Tipo: DNA

<211> Comprimento: 627

Nome da sequência: SEQ ID 162

Descrição da sequência:

Códones por encomenda

Nome da sequência: SEQ ID 162

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```

atgggtgtca ctcagacccc aaaattccag gtcctgaaga caggacagag catgacactg      60
cagtgtgccc aggatatgaa ccatgaatac atgtcctggt atcgacaaga cccaggcatg    120
gggctgaggc tgattcatta ctcagttggt gctggtatca ctgaccaagg agaagtcccc    180
aatggctaca atgtctccag atcaaccaca gaggatttcc cgctcaggct gctgtcggct    240
gctccctccc agacatctgt gtacttctgt gccagcagtt acgtcgggaa caccggggag    300
ctgttttttg gagaaggctc taggctgacc gtactggagg acctgaaaaa cgtgttccca    360
cccagggtcg ctgtgtttga gccatcagaa gcagagatct cccacaccca aaaggccaca    420
ctggtgtgcc tggccacagg ctctacccc gaccacgtgg agctgagctg gtgggtgaat    480
gggaaggagg tgcacagtgg ggtctgcaca gaccgcagc ccetcaagga gcagcccgcc    540
ctcaatgact ccagatacgc tctgagcagc cgctgaggg tctcgccac cttctggcag    600
gacccccgca accacttccg ctgtcaagtc cagttctacg ggctctcgga gaatgacgag    660
tggacccagg atagggccaa acccgtcacc cagatcgtca gcgccgagg ctggggtaga    720
gcagactaa                                     729

```

<212> Tipo: DNA

<211> Comprimento: 729

Nome da sequência: SEQ ID 163

Descrição da sequência:

Códones por encomenda

Nome da sequência: SEQ ID 163

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MQEVTQIPAA LSVPEGENLV LNCSFTDSAI YNLQWFRQDP GKGLTSLLLI QSSQREQTSG	60
RLNASLDKSS GRSTLYIAAS QPGDSATYLC AVRPTSGGSY IPTFGRGTSL IVHPYIQNPD...	120
PAVYQLRDSK SSOKSVCLFT DFDSQTNVSQ SKDSDVYITD KCVLDMRSMD FKSNSAVAWS	180
NKSDAFACANA FNNSIIPEDT FFPSPRESS	208

<212> Tipo: PRT

<211> Comprimento: 208

Nome da sequência: SEQ ID 164

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MGVTQTPKFQ VLKTGQSMFL QCAQDMNHEY MSWYRQDPGM GLRLIHYSVG AGITDQGEVP	60
NGYNVSRSTT EDFPLRLLSA APSQTSVYFC ASSYVGNTGE LFFGEGSRLT VLEDLKNVFP	120
PEVAVFEPSE AEISHTQKAT LVCLATGFYP DHVELSWWVN GKEVHSGVCT DPQPLKEQPA	180
LNDSRYALSS RLRVSATFWQ DPRNHFRQV QFYGLSENDE WTQDRAKPVV QIVSAEAWGR	240
AD	242

<212> Tipo: PRT

<211> Comprimento: 242

Nome da sequência: SEQ ID 165

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ttcctggcct tttgctggcc ttttgcctcac atgtaatgtg agttagctca aaggaccgga	60
aaacgaccgg aaaacgagtg tacattacac tcaatcgagt ctcattaggc accccaggct	120
ttacacttta tgcttccggc tcgtatgttg gagtaatccg tgggggccga aatgtgaaat	180
acgaaggccg agcatacaac tgtggaattg tgagcggata acaatttcac acaggaaaca	240
gctatgacca acaccttaac actcgccctat tgtaaagtg tgcctttgt cgatactggt	300
tgattacgcc aagctacgta ctttaagtatt ctatttcaag gagacagtca actaatgcgg	360
ttcgatgcat gaattcataa gataaagttc ctctgtcagt taatgaaata cctattgcct	420
acggcagccg ctggattgtt attactcggc attactttat ggataacgga tgccgtcggc	480
gacctaacaa taatgagcgc gcccgccgg ccctggccaa acaggagggt acgcagattc	540
ctgcagctct cgggtcggcc ggtaccgggt tgcctccac tgcgtctaag gacgtcgaga	600
gagtgtccca gaaggagaaa acttggttct caactgcagt ttactgata ctacagggt	660
cttcctcttt tgaaccaaga gttgacgtca aagtgactat gcgctattta caacctccag	720
tggttttaggc aggaccctgg gaaaggcttc cgcgataaat gttggaggtc accaaatccg	780
tcctgggacc ctttccagag acatctctgt tgcttattca gtcaagtcag agagagcaaa	840
caagtggaag tgtagagaca acgaataagt cagttcagtc tctctcgitt gttcaccttc	900
acttaatgcc tcgctggata aatcatcagg acgtagtact ttatacattg tgaattacgg	960
agcgacctat ttagtagtcc tgcacatga aatatgtaac cagcttctca gcctggtgac	1020
tcagccacct acctctgtgc tgtgagggcc gtcgaagagt cggaccactg agtcggtgga	1080
tggagacacg acactccggg acatcaggag gaagctacat acctacattt ggaagaggaa	1140
ccagccttat tgtagtcctc cttcgatgta tggatgtaaa ctttctcctt ggtcgggaata	1200
tgttcatccg tatatccaga acccggatcc tgccgtgtac cagctgagag acaagtaggc	1260
atataggtct tgggcctagg acggcacatg gtcgactctc actctaaatc cagtgacaag	1320
tctgtctgcc tattcaecga ttttgattct tgagatttag gtcactgttc agacagacgg	1380
ataagtggct aaaactaaga caaacaaatg tgtcacaag taaggattct gatgtgtata	1440
tcacagacaa gtttgtttac acagtgtttc attcctaaga ctacacatat agtgtctgtt	1500
atgtgtgcta gacatgaggt ctatggactt caagagcaac agtgctgtgg tacacacgat	1560

```

ctgtactcca gataacctgaa gttctcgttg tcacgacacc cctggagcaa caaatctgac 1620
tttgcattgtg caaacgcctt caacaacagc ggacctcgtt gtttagactg aaacgtacac 1680
gtttgcggaa gttgttgctg attattccag aagacacctt cttccccagc ccagaaagtt 1740
cctaataacc taataaggtc ttctgtggaa gaaggggtcg ggtctttcaa ggattattgg 1800
taggttaatt aagaattctt taagaagggg atatacatat gaaaaaatta atccaattaa 1860
ttcttaagaa attcttcccc tatatgtata cttttttaat ttattcgcaa ttcctttagt 1920
tgttctttc tattctcaca gcgcgcaggc aataagcgtt aaggaaatca acaaggaaaag 1980
ataagagtgt cgcgcgtccg tgggtgctact cagaccccaa aattccaggt cctgaagaca 2040
ggacagagca accacagtga gtctgggggtt ttaagggtcca ggacttctgt cctgtctcgt 2100
tgactactga gtgtgcccag gatatgaacc atgaatacat gtcttggtat actgtgacgt 2160
cacacgggtc ctatacttgg tacttatgta caggaccata cgacaagacc caggcatggg 2220
gtctgaggctg attcattact cagttgggtgc gctgttctgg gtccgtaccc cgactccgac 2280
taagtaatga gtcaaccacg tggatcact gaccaaggag aagtcccaa tggctacaat 2340
gtctccagat accatagtga ctggttctc ttcaggggtt accgatgta cagagggtcta 2400
caaccacaga ggatttcccg ctcaggctgc tgtcggctgc tccctcccag gttggtgtct 2460
cctaaagggc gagtccgacg acagccgacg agggaggggtc acatctgtgt acttctgtgc 2520
cagcagttac gtcgggaaca ccggggagct tgtagacaca tgaagacacg gtcgtcaatg 2580
cagcccttgt ggcctctcga gttttttgga gaaggctcta ggctgaccgt actggaggac 2640
ctgaaaaacg caaaaaacct cttccgagat ccgactggca tgacctcctg gactttttgc 2700
tgttcccacc cgaggctcgt gtgtttgagc catcagaagc agagatctcc acaagggtgg 2760
gtctccagca cacaacctcg gtagtcttcg tctctagagg cacacccaaa aggccacact 2820
gggtgtcctg gccacaggct tctaccccg gtgtgggtt tccggtgtga ccacacggac 2880
cgggtgtcga agatggggct ccacgtggag ctgagctggt ggggtgaatgg gaaggaggtg 2940
cacagtgggg ggtgcacctc gactcgacca cccacttacc ctctctccac gtgtcacccc 3000
tctgcacaga cccgcagccc ctcaaggagc agcccgcctt caatgactcc agacgtgtct 3060
gggcgtcggg gagtctctcg tcgggcggga gttactgagg agatacgtc tgagcagccg 3120
cctgaggggtc tcggccacct tctggcagga tctatgcgag actcgtcggc ggactcccag 3180
agccggtgga agaccgtcct cccccgcaac cacttccgct gtcaagtcca gttctacggg 3240
ctctcggaga gggggcgttg gtgaaggcga cagttcaggt caagatgcc gagagcctct 3300
atgacgagtg gaccaggat agggccaaac ccgtcaccca gatcgtcagc tactgtctac 3360
ctgggtccta tcccggttg ggcagtgggt ctagcagtcg gccgaggcct ggggtagagc 3420
agacgcggcc gcacggctcc ggaccccatc tcgtctgcgc cggcgt 3466

```

<212> Tipo: DNA

<211> Comprimento: 3466

Nome da sequência: SEQ ID 166

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 166

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

KYLLPTAAAG LLLLAAPAM AKQEVQIPA ALSVPEGENL VLNCSFTDSA IYNLQWFRQD	60
PGKGLTSLLL IQSSQREQTS GRLNASLDKS SGRSTLYIAA SQPGDSATYL CAVRPTSGGS	120
YIPTFGRGTS LIVHPYIQNP DPAVYQLRDS KSSDKSVCLF TDFDSQTNVS QSKDSQVYIT	180
DKCVLDMRSM DFKSNSAVAW SNKSDFACAN AFNNSIIPED TFFPSPESSM KKLLFAIPLV	240
VPFYSHSAQA GVTQTPKFQV LKTGQSMTLQ CAQDMNHEYM SWYRQDPGNG LRLIHYSVGA	300
GITDQGEVPN GYNVSRSTTE DFPLRLLSAA PSQTSVYFCA SSYVGNTGEL FFGEGSRLTV	360
LEDLKNVFPP EVAVFEPSEA EISHTQKATL VCLATGFYPD HVELSWWVNG KEVHSGVCTD	420
PQPLKEQPAL NDSRYALSSR LRVSATFWQD PRNHFRQVQ FYGLSENDEW TQDRAKPVTV	480
IVSAEAWGRA DAAA	494

<212> Tipo: PRT

<211> Comprimento: 494

Nome da sequência: SEQ ID 167

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

ggatccatgg ccataagtgg agtccctgtg ctaggatttt tcaiccatagc tgtgctgatg	60
agcgcctcagg aatcatgggc tatcaaagaa gaacatgtga tcatccaggc cgagttctat	120
ctgaatcctg accaatcagg cgagtttatg tttgactttg atggtgatga gattttccat	180
gtggatatgg caaagaagga gacggctctg cggttggaag aatttgagc atttgccagc	240
tttgaggctc aagggtgcatt ggccaacata gctgtggaca aagccaacct ggaaatcatg	300
acaaagcgct ccaactatac tccgatcacc aatgtacctc cagaggtaac tgtgctcacg	360
aacagccctg tggaaactgag agagcccaac gtcctcatct gtttcacga caagttcacc	420
ccaccagtgg tcaatgtcac gtggcttcga aatggaaaac ctgtcaccac aggagtgtca	480
gagacagtct tcctgcccag ggaagaccac cttttccgca agttccacta tctcccttc	540
ctgccctcaa ctgaggacgt ttacgactgc aggggtggagc actggggcct ggatgagcct	600
cttctcaagc actgggagtt tgatgtcca agccctctcc cagagactac agagaacgtg	660
gatgggggtc tgactgatac actccaagcg gagacagatc aacttgaaga caagaagtct	720
gcgttgacga ccgagattgc caatctactg aaagagaagg aaaaactaga gtccatcctg	780
gcagcttacg gatctgggtg tggctgaac gatattttg aagctcagaa aatcgaatgg	840
catgagtagg atcc	854

<212> Tipo: DNA

<211> Comprimento: 854

Nome da sequência: SEQ ID 168

Descrição da sequência:

Códões por encomenda

Nome da sequência: SEQ ID 168

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

aattctactc	gtaaagcgag	ttgaaggatc	atatttagtt	gcgtttatga	gataagattg	60
aaagcacgtg	taaaatgttt	cccgcggtt	ggcacaacta	tttacaatgc	ggccaagtta	120
taaaagattc	taattctgata	tgttttaaaa	cacctttgcg	gcccagagttg	tttgcgtagc	180
tgactagcga	agaagatgtg	tggaccgcag	aacagatagt	aaaacaaaac	cctagtattg	240
gagcaataat	cgatttaacc	aacacgtcta	aatattatga	tgggtgtgcat	tttttgcggg	300
cgggcctgtt	atacaaaaaa	attcaagtac	ctggccagac	tttgccgcct	gaaagcatag	360
ttcaagaatt	tattgacacg	gtaaaagaat	ttacagaaaa	gtgtcccggc	atgttggtgg	420
gcgtgcactg	cacacacggt	attaatcgca	ccggttacat	ggtgtgcaga	tatttaatgc	480
acaccctggg	tattgcgcgc	caggaagcca	tagatagatt	cgaaaaagcc	agaggtcaca	540
aaattgaaag	acaaaattac	gttcaagatt	tattaattta	attaatatta	tttgcatctt	600
ttaacaaata	ctttatccta	ttttcaaatt	gttgcgcttc	ttccagcgaa	ccaaaactat	660
gcttcgcttg	ctccgttttag	cttgtagccg	atcagtggcg	ttgttccaat	cgacggtagg	720
attaggccgg	atattctcca	ccacaatgtt	ggcaacgttg	atgttacggt	tatgcttttg	780
gttttccacg	tacgtctttt	ggccggtaat	agccgtaaac	gtagtgccgt	cgcgcgctac	840
gcacaacacc	ggatgtttgc	gcttgctcgc	ggggtattga	accgcgcgat	ccgacaaatc	900
caccactttg	gcaactaaat	cggtgacctg	cgcgctcttt	ttctgcatta	tttcgtcttt	960
cttttgcatg	gtttcctgga	agccgggtga	catgcgggtt	agatcagtca	tgacgcgcgt	1020
gacctgcaaa	tctttggcct	cgatctgctt	gtccttgatg	gcaacgatgc	gttcaataaa	1080
ctcttgtrtt	ttaacaagtr	cctcggtrtt	ttgcgccacc	accgcttgca	gcgcgtttgt	1140
gtgctcggtg	aatgtcgcaa	tcagcttagt	caccaactgt	ttgctctcct	cctcccggtg	1200
tttgatcgcg	ggatcgtagt	tgccgggtga	gagcacttga	ggaattactt	cttctaaaag	1260
ccattcttgt	aattctatgg	cgtaaggcaa	tttggacttc	ataatcagct	gaatcacgcc	1320
ggatttagta	atgagcactg	tatgcggctg	caaatacagc	gggtcgcccc	ttttcacgac	1380
gctgttagag	gtagggcccc	cattttggat	ggctgtgtca	aataacgatt	tgtattttatt	1440
gtctacatga	acacgtatag	ctttatcaca	aactgtatat	tttaaactgt	tagcgacgtc	1500
cttggccacg	aaccggacct	gttgggtcgc	ctctagcacg	taccgcaggt	tgaacgtatc	1560
ttctccaaat	ttaaattctc	caattttaac	gcgagccatt	ttgatacacg	tgtgtcgatt	1620
ttgcaacaac	tattgttttt	taacgcaaac	taaacttatt	gtggtaagca	ataattaaat	1680
atgggggaac	atgcgccgct	acaacactcg	tcgttatgaa	cgacagcggc	gccgggtctcg	1740
gcgcaagcgg	ctaaaacgtg	ttgcgcgttc	aacgcggcaa	acatcgcaaa	agccaatagt	1800

acagttttga	tttgcataatt	aacggcgcatt	ttttaaatta	tcttatttta	taaatagtta	1860
tgacgcctac	aactccccgc	ccgcgttgac	tcgctgcacc	tcgagcagtt	cgttgacgcc	1920
ttcctccgtg	tgccgaaca	cgtcgagcgg	gtggctgatg	accagcggcg	tgccgcacgc	1980
gacgcacaag	tatctgtaca	ccgaatgatc	gtcgggcgaa	ggcacgtcgg	cctccaagtg	2040
gcaatattgg	caaattcgaa	aatatataca	gttgggttgt	ttgcgcatat	ctatcgtggc	2100
gttgggcatg	tacgtccgaa	cgttgatttg	catgcaagcc	gaaattaaat	cattgcgatt	2160
agtgcgatta	aaacgttgta	catcctcgct	tttaatcatg	ccgtcgatta	aatcgcgcaa	2220
tcgagtcgaag	tgatcaaagt	gtggaataat	gttttctttg	tattccccgag	tcaagcgcag	2280
cgcgtatttt	aacaaactag	ccatcttgta	agttagtttc	atttaatgca	actttatcca	2340
ataatatatt	atgtatcgca	cgtcaagaat	taacaatgcy	cccgttgtcg	catctcaaca	2400
cgactatgat	agagatcaaa	taaagcgcga	attaaatagc	ttgcgacgca	acgtgcacga	2460
tctgtgcacg	cgttccggca	cgagctttga	ttgtaataag	tttttacgaa	gcgatgacat	2520
gacccccgta	gtgacaacga	tcacgcccaa	aagaactgcc	gactacaaaa	ttaccgagta	2580
tgtcggtgac	gttaaaaacta	ttaagccatc	caatcgaccg	ttagtcgaat	caggaccgct	2640
ggtgcgagaa	gccgcgaagt	atggcgaaatg	catcgtataa	cgtgtggagt	ccgctcatta	2700
gagcgtcatg	tttagacaag	aaagctacat	atttaattga	tcccgatgat	tttattgata	2760
aattgacctt	aactccatac	acggtattct	acaatggcgg	ggtttttggtc	aaaatttccg	2820
gactgcgatt	gtacatgctg	ttaacggctc	cgcccactat	taatgaaatt	aaaaattcca	2880
attttaaaaa	acgcagcaag	agaaacattt	gtagaaaga	atgcgtagaa	ggaaagaaaa	2940
atgtcgtcga	catgctgaac	aacaagatta	atatgcctcc	gtgtataaaa	aaaatattga	3000
acgatttgaa	agaaaacaat	gtaccgcgcg	gcggtatgta	caggaagagg	tttatactaa	3060
actgttacat	tgcaaacgtg	gtttcgtgtg	ccaagtgtga	aaaccgatgt	ttaatcaagg	3120
ctctgacgca	tttctacaac	cacgactcca	agtgtgtggg	tgaagtcatg	catcttttaa	3180
tcaaatccca	agatgtgtat	aaaccaccaa	actgccaaaa	aatgaaaact	gtcgacaagc	3240
tctgtccggt	tgctggcaac	tgcaagggtc	tcaatcctat	ttgtaattat	tgaataataa	3300
aacaattata	aatgctaaat	ttgtttttta	ttaacgatac	aaaccaaacy	caacaagaac	3360
atttgtagta	ttatctataa	ttgaaaacgc	gtagttataa	tcgctgagggt	aatattttaa	3420
atcattttca	aargattcac	agttaatttg	cgacaatata	attttatttt	cacataaact	3480
agacgccttg	tcgtcttctt	cttcgtattc	cttctctttt	tcatttttct	cctcataaaa	3540
attaacatag	ttattatcgt	atccatatat	gtatctatcg	tatagagtaa	attttttggt	3600
gtcataaata	tatatgtctt	ttttaatggg	gtgtatagta	ccgctgcgca	tagtttttct	3660
gtaattttaca	acagtgcctat	tttctggtag	ttcttcggag	tgtgttgctt	taattattaa	3720
attratataa	tcaatgaatt	tgggatcgtc	ggttttgtac	aatatgttgc	cggcatagta	3780
cgcagcttct	tcragttcaa	ttacaccatt	ttttagcagc	accggattaa	cataactttc	3840

caaaatgttg	tacgaaccgt	taaacaaaaa	cagttcacct	cccttttcta	tactattgtc	3900
tgcgagcagt	tgrttgttgt	taaaaataac	agccattgta	atgagacgca	caaactaata	3960
tcacaaactg	gaaatgtcta	tcaatatata	gttgctgatg	atccagcatg	ataagataca	4020
ttgatgagtt	tggacaaacc	acaactagaa	tgcagtgaaa	aaaatgcttt	atftgtgaaa	4080
tttgtgatgc	tattgcttta	tttghtaacca	ttataagctg	caataaacia	gttccgagtt	4140
tgtcagaaag	cagaccaaac	agcggttgga	ataatagcga	gaacagagaa	atagcggcaa	4200
aaataatacc	cgtatcactt	ttgctgatat	ggttgatgtc	atgtagccaa	atcgggaaaa	4260
acgggaagta	ggctcccatg	ataaaaaagt	aaaagaaaaa	gaataaaccg	aacatccaaa	4320
agtttgtgtt	ttttaaatag	tacataatgg	atttccttac	gcgaaatagc	ggcagacatg	4380
gcctgccccg	ttattattat	ttttgacacc	agaccaactg	gtaatggtag	cgaccggcgc	4440
tcagctggaa	ttcagatctg	tgattgtaaa	taaaatgtaa	tttacagtat	agtattttaa	4500
ttaatatata	aatgatttga	taataattct	tatttaacta	taatataattg	tgttgggttg	4560
aattaaaggt	cccggcatcc	tcaaattgat	aatttcatag	tcccccttgt	tgtaagtgat	4620
gcgtatttct	gaatctttgt	aaaatagcac	acaagactcc	aacgcgtttg	gcgttttatt	4680
ttcttgctcg	actctagttt	attaggcctc	tagagatccg	tatttatagg	tttttttatt	4740
acaaaactgt	tacgaaaaca	gtaaaatact	tatttatitg	cgagatgggt	atcattttaa	4800
ttatctccat	gatccaataa	acctagaata	aaggggccga	cctttaattc	aaccaaacac	4860
aatatattat	agrtaaataa	gaattattat	caaatcattt	gtatattaat	taaaatacta	4920
tactgtaaat	tacattttat	ttacaatcac	agatccccgg	gatccgggta	ttagtacatt	4980
tattaagcgc	tagattctgt	gcgttggtga	tttacagaca	attgtrgtac	gtattttaat	5040
aattcattaa	atttataatc	tttaggggtg	tatgttagag	cgaaaatcaa	atgattttca	5100
gcgtctttat	atctgaattt	aaatattaaa	tccatcaatg	atttgtaaaa	taggtttcga	5160
ttagtttcaa	acaagggttg	tttttccgaa	ccgatggctg	gactatctaa	tggattttcg	5220
ctcaacgcca	caaaacttgc	caaatcttgt	agcagcaatc	tagctttgtc	gatattcggt	5280
tgtgttttgt	tttghtaataa	aggttcgacg	tcgttcaaaa	tattatgcgc	ttttgtattt	5340
ctttcatcac	tgtcgttagt	gtacaattga	ctcgacgtaa	acacgttaaa	taaagcttgg	5400
acatatattaa	catcgggcgt	gttagcttta	ttaggccgat	tatcgtcgtc	gtcccaaccc	5460
tcgtcgttag	aagttgcttc	cgaagacgat	tttgccatag	ccacacgacg	cctattaatt	5520
gtgtcggcta	acacgtccgc	gatcaaattt	gtagttagac	tttttggaa	tattttctgat	5580
tgcgggcgtt	tttgggcggg	tttcaatcta	actgtgcccc	attttaattc	agacaacacg	5640
ttagaaaagc	atgggtgcagg	cgggtggtaac	atttcagacg	gcaaatctac	taatggcggc	5700
ggtggtggag	ctgatgataa	atctaccatc	ggtggaggcg	caggcggggc	tggcggcgga	5760
ggcggaggcg	gaggtggtgg	cgggtgatgca	gacggcggtt	taggctcaaa	tgtctcttta	5820
ggcaacacag	tcggcacctc	aactattgta	ctggtttcgg	gcgccgtttt	tggtttgacc	5880

ggctcgagac	gagtgcgatt	tttttcgttt	ctaatagcctt	ccaacaattg	ttgtctgtcg	5940
tctaaagggtg	cagcggggttg	aggttccgtc	ggcattgggtg	gagcggggcgg	caattcagac	6000
atcgatgggtg	gtgggtgggtg	tggaggcgct	ggaatggttag	gcacggggaga	aggtsgtggc	6060
ggcgggtgccg	ccggtataat	ttgttctggt	ttagtttggt	cgcgcacgat	tgtgggcacc	6120
ggcgcaggcg	ccgctggctg	cacaacggaa	ggtcgtctgc	ttcgaggcag	cgcttggggt	6180
gggtggcaatt	caatattata	attggaatac	aaatcgtaaa	aatctgctat	aagcattgta	6240
atttcgctat	cgtttaccgt	gccgatattt	aacaaccgct	caatgtaagc	aattgtattg	6300
taaagagatt	gtctcaagct	cggatcgatc	ccgcacgccg	ataacaagcc	ttttcatttt	6360
tactacagca	ttgtagtggc	gagacacttc	gctgtcgtcg	acgtacatgt	atgctttggt	6420
gtcaaaaacg	tcgttggcaa	gctttaaaat	atttaaaaga	acatctctgt	tcagcaccac	6480
tgtgttgctg	taaatgttgt	ttttgataat	ttgcgcttcc	gcagtatcga	cacgttcaaa	6540
aaattgatgc	gcatcaattt	tgttgcttct	attattgaat	aaataagatt	gtacagattc	6600
atatctacga	ttcgctcatg	ccaccacaaa	tgctacgctg	caaacgctgg	tacaatttta	6660
cgaaaactgc	aaaaacgtca	aaactcggta	taaaataatc	aacggggcgt	ttggcaaaaat	6720
atctatttta	tcgcacaagc	ccactagcaa	attgtatttg	cagaaaacaa	tttcggcgca	6780
caattttaac	gctgacgaaa	taaaagttca	ccagttaatg	agcgaccacc	caaattttat	6840
aaaaatctat	tttaatcacg	gttccatcaa	caaccaagtg	atcgtgatgg	actacattga	6900
ctgtcccgat	ttatttgaaa	cactacaaat	taaaggcgag	ctttcgtacc	aacttggttag	6960
caatattatt	agacagctgt	gtgaagcgct	caacgatttg	cacaagcaca	atttcataca	7020
caacgacata	aaactcgaaa	atgtcttata	tttcgaagca	cttgatcgcg	tgtatgtttg	7080
cgattacgga	ttgtgcaaac	acgaaaactc	acttagcgtg	cacgacggca	cgttggagta	7140
ttttagtcgg	gaaaaaattc	gacacacaac	tatgcacggt	tcgtttgact	ggtacgcggc	7200
gtgttaacat	acaagttgct	aaccggcggc	cgacacccat	ttgaaaaaag	cgaagacgaa	7260
atgttggact	tgaatagcat	gaagcgctgt	cagcaataca	atgacattgg	cgttttaaaa	7320
cacgttcgta	acgttaacgc	tcgtgacttt	gtgtactgcc	taacaagata	caacatagat	7380
tgtagactca	caaattacaa	acaaattata	aaacatgagt	ttttgtcgta	aaaatgccac	7440
ttgttttacg	agtagaattc	ccagcttggc	actggccgtc	gtttttacaac	gtcgtgactg	7500
ggaaaaccct	ggcgttacct	aacttaatcg	ccttgacgca	catccccctt	tcgccagctg	7560
gcgtaatagc	gaagaggccc	gcaccgatcg	cccttcccaa	cagttgcgca	gcctgaatgg	7620
cgaatggcgc	ctgatgcggt	attttctcct	tacgcatctg	tgcggtattt	cacaccgcat	7680
atggtgcact	ctcagtacaa	tctgctctga	tgccgcatag	ttaagccagc	cccgcacccc	7740
gccaacaccc	gctgacgcgc	cctgacgggc	ttgtctgctc	ccggcatccg	cttacagaca	7800
agctgtgacc	gtctccggga	gctgcatgtg	tcagagggtt	tcaccgtcat	caccgaaacg	7860
cgcggacgaa	agggcctcgt	gatacgccct	tttttatagg	ttaatgtcat	gataataatg	7920

gtttcttaga	cgtcagggtgg	cacttttcgg	ggaaatgtgc	gcggaacccc	tatttgttta	7980
tttttctaaa	tacattcaaa	tatgtatccg	ctcatgagac	aataaccctg	ataaatgctt	8040
caataatatt	gaaaaaggaa	gagtatgagt	attcaacatt	tccgtgtcgc	ccttattccc	8100
ttttttgcgg	cattttgcct	tcctgttttt	gctcaccag	aaacgctggg	gaaagtaaaa	8160
gatgctgaag	atcagttggg	tgacagagt	ggttacatcg	aactggatct	caacagcggg	8220
aagatccttg	agagttttcg	ccccgaagaa	cgttttccaa	tgatgagcac	ttttaagtt	8280
ctgctatgtg	gcgcgggtatt	atcccgtatt	gacgccgggc	aagagcaact	cggtcgcgcg	8340
atacactatt	ctcagaatga	cttggttgag	tactcaccag	tcacagaaaa	gcattcttacg	8400
gatggcatga	cagtaagaga	attatgcagt	gctgccataa	ccatgagtga	taacactgcg	8460
gccaaacttac	ttctgacaac	gatcggagga	ccgaaggagc	taaccgcttt	tttgacacaac	8520
atgggggatc	atgtaactcg	ccttgatcgt	tgggaaccgg	agctgaatga	agccatacca	8580
aacgacgagc	gtgacaccac	gatgcctgta	gcaatggcaa	caacgttgcg	caaactatta	8640
actggcgaac	tacttactct	agcttcccgg	caacaattaa	tagactggat	ggaggcggat	8700
aaagttgcag	gaccacttct	gcgctcggcc	cttcgggctg	gctggtttat	tgctgataaa	8760
tctggagccg	gtgagcgtgg	gtctcgggt	atcattgcag	cactggggcc	agatggtaag	8820
ccctcccgtg	tcgtagttat	ctacacgacg	gggagtcagg	caactatgga	tgaacgaaat	8880
agacagatcg	ctgagatagg	tgccctactg	attaagcatt	ggtaactgtc	agaccaagtt	8940
tactcatata	tacttttagat	tgatttaaaa	cttcattttt	aatttaaaag	gatctaggtg	9000
aagatccttt	ttgataatct	catgaccaa	atcccttaac	gtgagttttc	gttccactga	9060
gcgtcagacc	ccgtagaaaa	gatcaaagga	tcttcttgag	atcctttttt	tctgcgcgta	9120
atctgctgct	tgcaaacaaa	aaaaccaccg	ctaccagcgg	tggtttggtt	gccggatcaa	9180
gagctaccaa	ctctttttcc	gaaggtaact	ggcttcagca	gagcgcagat	accaaatact	9240
gtccttctag	tgtagccgta	gttaggccac	cacttcaaga	actctgtagc	accgcctaca	9300
tacctcgctc	tgetaatcct	gttaccagt	gctgctgcca	gtggcgataa	gtcgtgtctt	9360
accgggttgg	actcaagacg	atagttaccg	gataaggcgc	agcggtcggg	ctgaacgggg	9420
ggttcgtgca	cacagcccag	cttggagcga	acgacctaca	ccgaactgag	atacctacag	9480
cgtgagcatt	gagaaagcgc	cacgcttccc	gaagggagaa	aggcggacag	gtatccggtg	9540
agcggcaggg	tcggaacagg	agagcgcacg	agggagcttc	cagggggaaa	cgcttgggtat	9600
ctttatagtc	ctgtcggggt	tcgccacctc	tgacttgagc	gtcgattttt	gtgatgctcg	9660
tcaggggggc	ggagcctatg	gaaaaacgcc	agcaacgcgg	cctttttacg	gttcctggcc	9720
ttttgctggc	cttttgctca	catgttcttt	cctgcgttat	ccccgtattc	tgtggataac	9780
cgtattaccg	ccrttgagt	agctgatacc	gctcgcgcga	gccgaacgac	cgagcgcagc	9840
gagtcagtga	gcgaggaagc	ggaagagcgc	ccaatacgca	aaccgcctct	ccccgcgcgt	9900
tggccgattc	attaatgcag	ctggcacgac	aggtttcccg	actggaaagc	gggcagtga	9960

PE1558643

- 183 -

```
cgcaacgcaa ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg 10020
cttccggctc gtatgttggtg tgggaattgtg agcgggataac aatttcacac aggaaacagc 10080
tatgaccatg attacg 10096
```

<212> Tipo: DNA

<211> Comprimento: 10096

Nome da sequência: SEQ ID 169

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 169

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

```
ggtaccggat ccagcatggt gtgtctgaag ctccctggag gctcctgcat gacagcgctg 60
acagtgacac tgatggtgct gagctcccca ctggctttgt ccggagacac cggacctaag 120
tacgtcaagc agaacacact gaaactggct tccgggtggcg gatctctagt tccacgcggt 180
agtggaggcg gtggttccgg agacacgcgt ccacgtttct tgtggcagct taagtttgaa 240
tgtcatttct tcaatgggac ggagcgggtg cggttgctgg aaagatgcat ctataaccaa 300
gaggagtccg tgcgcttcga cagcgacgtg ggggagtacc gggcggtgac ggagctgggg 360
cggcctgatg ccgagtactg gaacagccag aaggacctcc tggagcagag gcgggccgcg 420
gtggacacct actgcagaca caactacggg gttggtgaga gcttcacagt gcagcggcga 480
gttgagccta aggtgactgt gtatccttca aagaccagc ccctgcagca ccacaacctc 540
ctggtctgct ctgtgagtgg tttctatcca ggcagcattg aagtcagggtg gttccggaac 600
ggccaggaag agaaggctgg ggtggtgtcc acaggcctga tccagaatgg agattggacc 660
ttccagaccc tggatgatgct ggaaacagtt cctcggagtg gagaggttta cacctgccaa 720
gtggagcacc caagtgtgac gagccctctc acagtggaaat ggagagcacg gtctgaatct 780
gcacagagca aggtcgacgg aggcggtggg ggtagaatcg cccggctgga ggaaaaagtg 840
aaaaccttga aagctcagaa ctcggaactg gcgtccacgg ccaacatgct cagggaacag 900
gtggcacagc ttaaacagaa agtcatgaac tactaggatc c 941
```

<212> Tipo: DNA

<211> Comprimento: 941

Nome da sequência: SEQ ID 170

Descrição da sequência:

Codões por encomenda

Nome da sequência: SEQ ID 170

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE 60
VPNGYNVSRS TTEDFPLRLI SAAPSQTSVY FCASRPGLAG GRPEQYFGPG TRLTVT 116

<212> Tipo: PRT
<211> Comprimento: 116

Nome da sequência: SEQ ID 171
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE 60
VPNGYNVSRS TTEDFPLRLI SAAPSQTSVY FCASRPGLMS AXPEQYFGPG TRLTVT 116

<212> Tipo: PRT
<211> Comprimento: 116

Nome da sequência: SEQ ID 172
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

Característica

Sequência: SEQ ID 172:
<221> Característica chave: "misc_feature"
<222> localização a partir de: 1
<222> Localização até: 116

Outra informação: em que x é e, q ou r
Liga sequência codificadora: Não

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE 60
VPNGYNVSRS TTEDFPLRLI SAAPSQTSVY FCASRPGLAG GRPEDQYFGP GTRLTVT 117

<212> Tipo: PRT
<211> Comprimento: 117

Nome da sequência: SEQ ID 173
Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética
<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE 60
VPNGYNVSR5 TTEDFPLRLL SAAPSQTSVY FCASRPGLVP GRPEQQFGPG TRLTVT 116

<212> Tipo: PRT

<211> Comprimento: 116

Nome da sequência: SEQ ID 174

Descrição da sequência:

Sequência

<213> nome do organismo: Sequência sintética

<400> segmento de pré-sequência:

MNAGVTQTPK FQVLKTGQSM TLQCAQDMNH EYMSWYRQDP GMGLRLIHYS VGAGITDQGE 60
VPNGYNVSR5 TTEDFPLRLL SAAPSQTSVY FCASRPGLAG GRPHPQFGPG TRLTVT 116

<212> Tipo: PRT

<211> Comprimento: 116

Nome da sequência: SEQ ID 175

Descrição da sequência:

Lisboa, 17 de Agosto de 2009

REIVINDICAÇÕES

1. Uma partícula fágica apresentando na sua superfície um polipéptido de TCR de cadeia simples (scTCR) ou um par de polipéptidos de TCR dimérico (dTCR), o referido polipéptido de scTCR sendo constituído por:

um primeiro segmento constituído por uma sequência do domínio variável da cadeia α ou δ de TCR fundida com a extremidade N de uma sequência extracelular do domínio constante da cadeia α de TCR,

um segundo segmento constituído por um domínio variável da cadeia β ou γ de TCR fundido com a extremidade N da sequência extracelular do domínio constante da cadeia β de TCR,

uma sequência de um elemento de ligação que liga o extremo C do primeiro segmento ao extremo N do segundo segmento, ou *vice versa*, e

uma ligação dissulfureto entre o primeiro e segundo segmentos, a referida ligação dissulfureto sendo uma que não tem equivalente nos receptores de células T $\alpha\beta$ ou $\gamma\delta$ nativos;

o referido par de polipéptidos dTCR sendo constituído por:

um primeiro polipéptido no qual uma sequência do domínio variável da cadeia α ou δ de TCR está fundida com a extremidade N de uma sequência extracelular do domínio constante da cadeia α de TCR,

um segundo polipéptido no qual uma sequência de um domínio variável da cadeia β ou γ de TCR está fundida com a extremidade N da sequência extracelular do domínio constante da cadeia β de TCR,

o primeiro e o segundo polipéptidos estando ligados por uma ligação dissulfureto que não tem equivalente nos receptores de células T $\alpha\beta$ ou $\gamma\delta$ nativos

a referida ligação dissulfureto, entre o primeiro e o segundo segmentos de scTCR ou entre o primeiro e segundo polipéptidos do dTCR, sendo entre cisteínas substituídas nos seguintes pares de resíduos no exão 1 de TRAC*01 para a cadeia de TCR α e de TRBC1*01 ou TRBC02*01 para a cadeia de TCR β :

Cadeia α de TCR	Cadeia β de TCR
Thr 48	Ser 57
Thr 45	Ser 77
Tyr 10	Ser 17
Tre 45	Asp 59
Ser 15	Glu 15

2. Uma partícula fágica de acordo com a

reivindicação 1, que é uma partícula de um fago filamentoso.

3. A partícula fágica de acordo com a reivindicação 1 ou a reivindicação 2, na qual a extremidade C de um membro do par de polipéptidos de dTCR ou a extremidade C do polipéptido de scTCR, é ligada através de uma ligação peptídica a um resíduo exposto na superfície da partícula fágica.

4. Uma partícula fágica, como reivindicada em qualquer uma das reivindicações anteriores, em que o polipéptido scTCR ou o par de polipéptidos de dTCR apresentado é humano e a referida ligação dissulfureto se situa entre os resíduos de cisteína substituídos na Thr 48 do exão 1 e na Ser 57 do exão 1 de TRBC1*01 ou TRBC2*01.

5. Uma partícula fágica como reivindicado em qualquer uma das reivindicações 1 a 3 em que o polipéptido scTCR ou o par de polipéptidos dTCR apresentado é proveniente de um mamífero que não do Homem e a referida ligação dissulfureto situa-se entre os resíduos de cisteína nos equivalentes do mamífero de Thr 48 do exão 1 de TPAC*01 e Ser 57 do exão 1 de TRBC1*01 ou TRBC2*01 de TCRs humanos.

6. Uma partícula fágica como reivindicado em qualquer uma das reivindicações anteriores em que o TCR está truncado, no extremo C no caso de um dTCR ou no extremo C no caso de um scTCR, relativamente aos referidos receptores de células T nativos, de forma que sejam

excluídos os resíduos de cisteína que formam a ligação dissulfureto nativa entre cadeias de TCR.

7. Uma partícula fágica, como reivindicado em qualquer uma das reivindicações 1 a 5, em que os resíduos de cisteína que formam a ligação dissulfureto nativa entre cadeias são substituídos por resíduos não cisteína.

8. Uma partícula fágica como reivindicada em qualquer uma das reivindicações anteriores em que no dTCR ou no scTCR apresentado não existem resíduos de cisteína desemparelhados correspondendo ao resíduo de cisteína presente num TCR nativo.

9. Uma partícula fágica como reivindicada na reivindicação 1, a qual é uma partícula de fago filamentososo, apresentando na sua superfície um par de polipéptidos do receptor de células T dimérico (dTCR) humano, o referido par sendo constituído por:

um primeiro polipéptido no qual uma sequência do domínio variável da cadeia α ou δ de TCR está fundida com a extremidade N de uma sequência extracelular do domínio constante da cadeia α de TCR,

um segundo polipéptido no qual uma sequência de um domínio variável da cadeia β ou γ de TCR está fundida com a extremidade N da sequência extracelular do domínio constante da cadeia β de TCR,

o primeiro e segundo polipéptidos estando ligados por uma ligação dissulfureto entre cisteínas substituídas em Thr 48 do exão 1 de TRAC*01 e Ser 57 no exão 1 de TRBC1*01 ou TRBC02*01,

o extremo C de um membro do par de polipéptidos de dTCR estando ligado por uma ligação peptídica a uma proteína externa do fago.

10. Uma biblioteca diversificada de partículas fágicas como reivindicado em qualquer uma das reivindicações anteriores.

11. Uma biblioteca diversificada como reivindicado na reivindicação 10 em que a diversidade reside no(s) domínio(s) variável(eis) do par de polipéptidos de dTCR ou do polipéptido scTCR

12. Um método para a identificação de TCRs com uma característica específica, o referido método compreendendo a submissão de uma biblioteca diversificada de TCRs apresentados em partículas fágicas, como reivindicado na reivindicação 10 ou na reivindicação 11, a

um processo de selecção que selecciona a referida característica e de isolamento das partículas fágicas que apresentem um TCR tendo a referida característica e, facultativamente, a um processo de amplificação para multiplicar as partículas isoladas
e/ou

um processo de rastreio que mede a referida característica, identificação das partículas fágicas que apresentam um TCR com a característica pretendida e isolamento destas partículas fágicas e, facultativamente, a um processo de amplificação para multiplicar as partículas isoladas.

13. Um método como reivindicado na reivindicação 12, em que a característica específica é uma afinidade acrescida para um ligando de TCR.

Lisboa, 17 de Agosto de 2009

Figura 1a

Atgcagaaggaagtggagcagaactctggaccctcagtggtccagagggagccatt
gcctctctcaactgcacttacagtgaccgaggttcccagtccttcttctggtacaga
caatattctgggaaaageccctgagttgataatgtccatatactccaatggtgacaaa
gaagatggaagggtttacagcacagctcaataaagccagccagtatgtttctctgctc
atcagagactcccagcccagtgattcagccacctacctctgtgcccgtttacaactgac
agctgggggaaaattgcagttttggagcaggggaeccaggttggtggtcacccccagatata
cagaacctgacctgtccgtgtaccagctgagagactctaaatccagtgacaagtct
gtctgectattcaaccgatttttgattctcaaaaaatgtgtcacaaagtaaggattct
gatgtgtatatcacagacaaaatgtgtgctagacatgaggtctatggacttcaagagc
aacagtgtctgtggcctggagcaacaaatctgactttgcatgtgcaaacgccttcaac
aacagcattattccagaagacaccttcttccccagcccagaaagttcctaa
(SEQ ID 147)

Figura 1b

atgaacgctgggtgtcactcagacccccaaaattccaggtcctgaagacaggacagagc
atgacactgcagtggtgcccaggatatgaaccatgaatacatgtccctggtatcgacaa
gaeccaggcatggggctgaggtgattcattactcagttggtgctggtatcactgac
caaggagaagtecccaatggctacaatgtctccagatcaaccacagaggatttcccg
ctcaggctgctgtggctgctccctcccagacatctgtgtacttctgtgtccagcagg
ccgggactagcgggagggcgaccagagcagtaacttcgggcccgggcaccaggetcacg
gtcacagaggacctgaaaaacgtgttcccacccgaggtcgtgtgtgtttgagccatca
gaagcagagatctcccacacccaaaaggccacactgggtgtgcctggccacaggcttc
taccctcgaccacgtggagctgagctggtgggtgaatgggaaggaggtgcacagtggg
gtctgcacagacecgcagcccctcaaggagcagcccgcctcaatgactccagatac
gctctgagcagccgcctgagggctctgggcacettctggcaggacccccgcaaccac
ttccgctgtcaagtccagttctacgggctctcggagaatgacgagtggaaccaggat
agggccaaacccgtcaccagatcgtcagcgcgaggcctggggtagagcagactaa
(SEQ ID 148)

Figura 2a

MQ

KEVEQNSGPL SVPEGAIASL NCTYSDRGSQ SFPWYRQYSG KSPELIMSIY
SNGDKEDGRF TAQLNKASQY VLLIIRDSQP SDSATYLCAY TTDSWGKLOF
GAGTQVVVTF DIQNPDPVAV QLRDSKSSDX SVCLETFDFS QTNVSQSKDS
DVYITDKCVL DMRSMDPKSN SAVAWSNKSD FACANAFNNS IIPEDTFFPS
PESS*

(SEQ ID 149)

Figura 2b

M

NAGVTQTPKF QVLKTQSMT LQCAQDMNHE YMSWYRQDPG MGLRLIHYSV
GAGITDQGEV PNGYNVSRST TEDFFLRLLS AAPSQTSVYF CASRPGLAGG
RPEQYFGPGT RLTVTEDLKN VFPPEVAVFE PSEAEISHTQ KATLVCLATG
FYPDHVELSW WVNGKEVHSG VCTDPQPLKE QPALNDSRYA LSSRLRVSAT
FWQDPRNHFR CQVQFYGLSE NDEWTQDRAK PVTQIVSAEA WGRAD*

(SEQ ID 150)

Figura 3

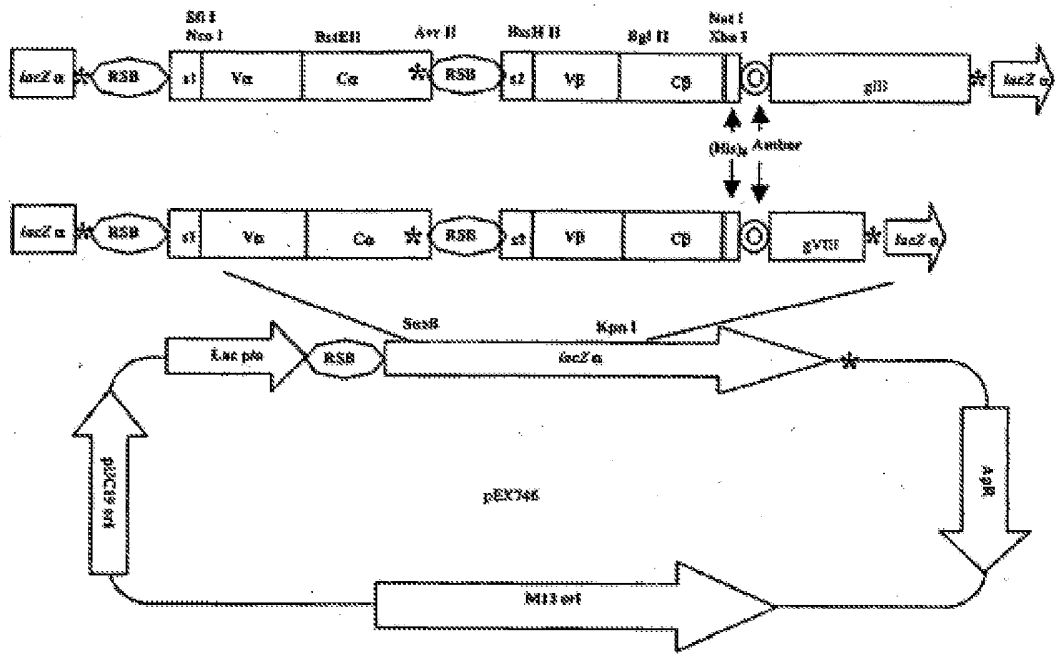


Figura 4

```

1  gttaaactaag  tcagggtggca  cttttcgggg  aaatgtgcgc  ggaaccccta  tttgtttatt
61  ttcttaataa  cattcaataa  tgtatccgct  catgagacaa  taacccctgat  aaatgcttca
121  ataataattga  aaaaggaaga  gtatgagrat  tcaacatttc  egtgtccgcc  ttattccett
181  ttttggggca  ttttgccttc  ctgtttttgc  tcacccagaa  acgctgggtga  aagtaaaaga
241  tgctgaagat  cagttgggtg  cccgagtggt  ttacatcgaa  ctggatctca  acagcggtaa
301  gctccttgag  agttttcgcc  ccgaagaacg  ttctccaatg  atgagcactt  ttaagttctt
361  gctatgtggc  gccgtattat  cccgtgttga  cgcggggcaa  gagcaactcg  gtgcgcgat
421  acactattct  cagaatgact  tgggttagta  ctacaccgtc  acagaaagac  atcttacgga
481  tggcatgaca  gtaagagaat  tatgcagtgc  tgcataaacc  atgagtgata  acactgcggc
541  caacttactt  ctgacaacga  tcggaggaac  gaaggagcta  accgtttttt  tgcacaacat
601  gggggatcat  gtaactcgcc  ttgatcgttg  ggaacgggag  ctgaatgaag  ccataccaaa
661  cgacggagct  gacaccacga  tgcctgttag  aatggcaaca  acgttgcgca  aactattaac
721  tggcgaacta  cttactctag  cttcccgcca  acaattaata  gactggatgg  agggcgataa
781  agttgcagga  ccacttctgc  gctcggccct  tccggctggc  tggtttatgg  ctgataaatc
841  tggagccggt  gagcgtgggt  ctgcgggtat  cattgcagca  ctggggccag  atggtaagcc
901  ctcccgatc  gtagttatct  acacgacggg  gagtcaggca  actatggatg  aacgaatatg
961  acagatcgct  gagataggtg  cctcaactgt  taagcattgg  taactgtcag  accaagttta
1021  ctcatatata  ctttagattg  atttaacccg  gttgataatc  agaaaagccc  caaaaacagg
1081  aagattgtat  aagcaaatat  ttaaatgtta  aacgttaata  tttgtttaa  attcgcgtta
1141  aatttttgg  aaatcagctc  attttttaac  caataggccg  aaatgggcaa  aatcccttat
1201  aaatcaaaag  aatagcccca  gatagggttg  agtgtgttcc  cagtttggaa  caagagtcca
1261  ctattaaaga  acgtggactc  caacgtcaaa  gggcgaaaaa  ccgtctatca  gggcgatggc
1321  ccactacgtg  aaccatcacc  caaatcaagt  tttttgggtt  cgaaggtccg  taagcacta
1381  aatcggaaac  ctaaaaggag  ccccgatttt  agagcttgac  ggggaagcgc  aacgtggcga
1441  gaagggaaag  gaagaagcgc  aaaggagcgg  gcgctagggc  gctggcaagt  gtacgggtca
1501  cgttgcgcgt  aaccaccaca  ccgcgcgcgc  ttaatgcgcc  gctacagggc  gcgtaaaagg
1561  atctaggtga  agatcctttt  tgataacttc  atgaccaaaa  tcccttaagg  tgaagtttgc
1621  ttccactgag  cgtcagaccc  cgtagaaaag  atcaaaaggt  ctctttgaga  tccctttttt
1681  ctgcggctaa  tctgtctgct  gcaaacaaaa  aaaccaccgc  taccagcggc  ggttttgtttg
1741  ccggatcag  agctaaccaa  tctttttccg  aaggttaactg  gcttcagcag  agcgcagata
1801  tccaaatact  ttcttctagt  gtacccgtag  ttaggccacc  acttcagaa  cctctgtaga
1861  ccgcctacat  acctcgctct  gctaactctg  ttaccagtg  ctgctgcacg  tggcgataag
1921  tctgtgttta  cggggttgga  ctaaaagcga  tagttaccgg  ataaggcgca  gcggtcgggc
1981  tgaacggggg  gttcgtgcac  acagcccgac  ttggagcgaa  cgacctacac  cgaactgaga
2041  taactacagc  gtgagctatg  agaaagcgcc  acgcttcccg  aaggggaaa  ggcggcaggg
2101  tatccggtaa  ggggcagggt  cggaaacggg  gagcgcaaga  gggagcttcc  agggggaaac
2161  gcttgggtat  ttttatagtc  tgtcgggttt  cggccacctc  gacttgagcg  tcatttttt
2221  tgatgctcgt  cagggggggc  gagcctatgg  aaaaaagcca  gcaacggcgc  ctttttacgg
2281  ttcttggcct  tttgtctggc  tttgtctcac  atgtaaatgt  agttagctca  ctcatlaggc
2341  accccaggct  ttacacttta  tgcctccggc  tcgtatgttg  tgtggaattg  tgagcggata
2401  acaatttcac  acaggaaaca  gctatgacca  tgattacgcc  aagctacgta  taataatatt
2461  ctattttca  gagacagtca  taatgaata  cctattgcct  accgcagccg  ctggatttgt
2521  attactcgcg  gccagccgg  ccattggcca  gaaggaaagt  gacagaact  ctggacccct
2581  cagtgttcca  gagggagcca  ttgcctctct  caactgcact  tacagtgacc  gaggttcca
2641  gtccttcttc  tggtagagac  aatattcttg  gaaaagccct  gaggtagata  tgtccatata
2701  ctccaatggt  gacaaagaag  atggaaggtt  tacagcacag  ctcaataaag  ccagocagta
2761  tgtttctctg  ctcatcagag  actccagcc  cagtgaattca  gccacctaac  tctgtgcgt
2821  tacaactgac  agctggggga  aattgcagtt  tggagcaggg  acccagggtg  tggtaacccc
2881  agatatccag  aacctgacc  ctgcctgtga  ccagctgaga  gactctaat  ccagtgacaa
2941  gtctgtctgc  ctattcaccc  attttgatc  tcasacaaat  gtgtcacaaa  gtaaggattc
3001  tgatgtgtat  atcacagaca  aatgtgtgt  agacatgagg  tctatggact  tcaagagcaa
3061  cagtgtctgt  gcctggagca  acaaatctga  ctttgcattg  gcaaacgcct  tcaaacacag
3121  cattattcca  gaagacacct  tcttccccag  cccagaaggt  tcttaataac  ctaggttaat
3181  taagaattct  ttaagaagga  gatatacata  tgaaaaaakt  attattcgca  attcctttag
3241  ttgttccttc  ctattctcac  agcgcgcagg  ctggtgtcac  tcagacccca  aaattccagg
3301  tcttgaagac  atgacactgc  atgacactgc  agtgtgcaca  ggtatgaac  catgataaca
3361  tgtcctggta  tcgacaagac  ccaggcatgg  ggctgaggct  gattcattac  tcagttgggtg

```

```
3421 ctgggtatcac tgaccaagga gaagtcacca atggctacaa tgtctccaga tcaaccacag
3481 aggattttccc gctcaggctg ctgtcggctg etccctccca gacatctgtg tecttctgtg
3541 ccagcaggcc gggaclagcg ggaaggcgac cagagcagta ctccgggccc ggcaaccagc
3601 tcacgggtcac agaggacctg aaaaacgtgt tcccacccga ggtagctgtg tttaggccat
3661 cagaagcaga gatctccac acccaaaagg ccacactggt gtgcctggcc acaggcctct
3721 accccgacca cgtcggagctg agctgggtgg tgaatgggaa ggaggtgcac agtggggtct
3781 gcacagacc gcagccccc caggagcagc ccgcctcaa tgactccaga tacgctctga
3841 gcagccgctt gagggtctct gccacctctt ggcaggaccc ccgcaaccac ttccgctgtc
3901 aagtccagtt ctacgggctc tcggggaatg acgagtgga caggatagg gccaaacccg
3961 tcaccacagat cgtcagcgcc gaggcctggg gtagagcaga cgcggccgca tctagacatc
4021 atcaccatca tcaactagact gttgaaagtt gtttagcaaa accccataca gaaaattcat
4081 ttactaacgt ctggaagagc gacaaaactt tagatcgtta cgttaactat gagggttgtc
4141 tgtggaatgc tacaggcgtt gtagtttgtt ctggtgacga aactcagttt taagggtacat
4201 gggttcctat tgggcttgct atccctgaaa atgaggggtg tggctctgag ggtggcggtt
4261 ctgaggggtg cggttctgag ggtggcggtt ctaaacctcc tgaatcaggt gatcaccta
4321 ttccgggcta taettatata aacctctctg acggcaetta tcgcctggtt actgagcaaa
4381 acccgcctaa tccatctctt tctcttgagg agtctcagcc tottaatact tctatgttct
4441 agaataatag gttccgaaat aggcaggggg catthaactgt ttatacgggc actgttaactc
4501 aaggcactga cccgttaaa acttattacc agtacactcc tgtatcatca aaagccatgt
4561 atgaegccta ctggaacggt aaattcagag actgcgtttt ccattctggc tttaatgagg
4621 atccattcgt ttgtgaatat caaggccaat cgtctgacct gcctcaacct cctgtcaatg
4681 ctggcgggcg ctctgggtgt ggttctggtg gggctctga gggtagtgc tctgaggggtg
4741 ggggttctga ggggtggcgc tctgagggag gcggttcagg tggtagctct ggttccgggtg
4801 attttgatta tgaaaagatg gcaaacgcta ataaaggggc tatgaccgaa aatgdcgatg
4861 aaaaacgcgt acagtctgac gctaaaggca aacttgattc tgtcgctaact gattacgggtg
4921 ctgctatcga tggtttcact ggtgacgttt cggpcttgc taatggtaat ggtgctactg
4981 gtgattttgc tggctctaat tcccaaatgg ctcaagtcgg tgacgggtgat aattcacctt
5041 taatgaataa ttccgctcaa tatttacctt cctccctca atcggttgaa tgtcgccctt
5101 ttgtcttttag cgtgggtaaa ccataatgaat ttctattga ttgtgacaaa ataaacttat
5161 tccgtgggtg ctttgggttt cttttatatg ttgccacctt tatgtatgta tttcttaagt
5221 ttgctaaact actgcgtaat aaggagtctt aataaggtao cctctagtca aggcctatag
5281 tgagtcgtat tacggactgg ccgtcgtttt acaaagtcgt gactgggaaa accctggcgt
5341 taccacaact aatgcgcttg cagacatcc ccttttcgcc agctggcgta atagcgaaga
5401 ggcgcgcacc gatcgccctt cccaacagtt ggcagcctg aatggcgaat ggcgcttcgc
5461 ttgtaataa agccgccttc ggcgggcttt ttttt
```

(SEQ ID 151)

Figura 5

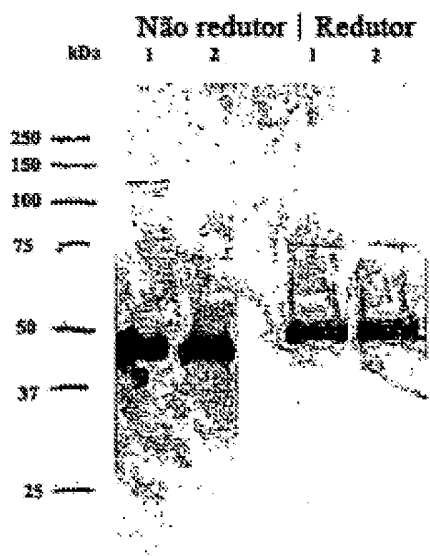


Figura 6

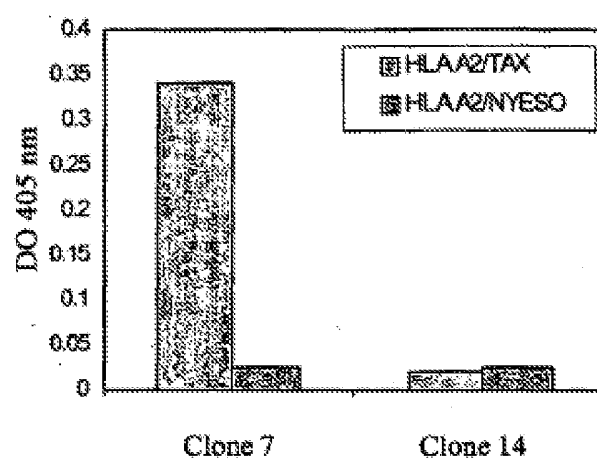


Figura 7a

Diagrama esquemático da construção de apresentação ribossômica de scTCR A6-C-Kapa

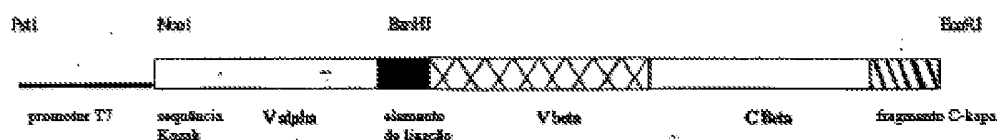


Figura 7b

```

agctgcagct aatagcactc actatagga caggcaaca tggccacaa
ggaagtggag cagacactctg gacccctcag tgtccacag ggaagccattg
cctctctcaa ctgcacttcc agtgaccag gttccacgtc cttctctcgg
taccagcaat attctgggaa agccctggag ttgataatgt ccatatactc
caatggtgac aaagaagatg gaagggttcc agcacagctc aataagcca
gccagtatgt ttctctgctc atcagagact cccagcccag tgattcagcc
acctacctct gtgcgttcc aactgacagc tgggggaaat tgcagtttgg
agcaggggac caggttgttg tccacggctg aggcggttca ggcgaggtg
gatccggcgg tggcggtctg aacgctgggt tcaactcagac cccaatctc
caggtcctga agacaggaca ggcctgaca ctgcagtggt cccaggtat
gaaccatgaa tcaatgtctt ggtatcgaca agaccagggc atggggctga
ggctgattca ttactcagtt ggtgctggta tcaactgacca aggaagagtc
cccattggct acctgtctc cagatcaacc acagaggatt tccgctcag
gctgctgtcg gctgctcctt cccagacata tgtgtacttc tgtccagca
ggccgggact agccgggggg cagccagagc agtacttcgg gcggggcacc
agctcacgg tccagggga cctgaaaaac gtgttccac ccaggtctg
tgtgtttgag ccatcagag cagagatctc caacaccaca aaggccacac
tgtgtgctt ggccacagga ttctacccc accacgtgga gctgagctgg
tgggtgaatg ggaaggaggt gcacagtggg gtcagtcacg acccgagcc
ctcaaggag cagcccgccc tcaatgactc cagataagct ctgagcagcc
gctgaggggt ctgggcacac ttctggcagg acccccgcaa caacttcgg
tgtcaagtc agttctacgg gctctcggag aatgacaggt ggaaccagg
tagggccaaa ccgtcaccc agatcgtcag cgcggaggcc tgggttagag
cagacgggtg aggcggttca ctacagaca cctgagctt gacaaagca

```

gactacgaga aacacaaagt ctacgcctgc gaagtcaccc atcagggcct
gagttcgccc gtcacaaaga gcttcaaccg cggagagtcn taagaattct
cag

(SEQ ID 152)

Figura 7c

N G Q K E V E Q N S G P L S V P E G A I A S L
N C L I M S I Y S N C D K E D G R F T A G E S P
E A S Q Y V S L L I R D S P S D G R F T A G E S P
T T D S W G K L Q F G A G T Q V T Q T F K F Q V
S L G G Q Q S G S L Q C S N A G V T Q T F K F Q V
R Q D P G Q S M T L R L I H Y S V G A G I T D Q G
E V P M G Y N V E R A S T T E D F F L G R L S A
A P S Q T S V Y F C A S R F G L A G P P P E V A
Y F F P S H V E L S W V N G R Y A L S S R L S
D P P Q P L K E Q P A L N D S R Y A L S S R L S
V S A D E W T Q D R A K P V T Q I V S A E A W G
R A D G G G S L S S T L T L S K A D Y S K H
K V Y A C E V T N Q G L S S P V T K S F N R G
S S

(SEQ ID: 153)

Figura 8

Sequência pUC19-T7

```

1  ctgcttctcc ggagcactat ggggatataa atatcccaatt acagtaactat tatkacaaa gaactctgag
71  kccaccgtga aaagcccttt kccaccgccc ttggggatata acataaaaa agattttatgt aagttttatac
141  ataggcgaggt acctctgttat tgggactatt taogaagtta ctataacttt ttccttctca tactcataag
211  ttgttaaggc acagcgggaa taagggaaaaaacgcgcgtaa aacgggaagg caaaaaagag tgggtcttttg
281  cgaaccacttt catctttctac gacttctagt caaccacagt gctcacccaa tctagcttga cctagagttg
351  tggccattctt aggaactctc aaaaagggggg cttctttgaaa aaggttacta ctgtgaaaa ttccaagagc
421  ataccacggcg ccataaatagg gcataaactgc ggcccgttct cgttgagcca gggcggtatg tgataagagt
491  ettaactgaac caactcatga gtggtaagtg tcttttctga gaatgectac cgtactgtca ttctctaat
561  acgtcacgac ggtatttgya ctactatttg tgagccggtt tgaatgaaga ctgttctag cctctggct
631  tccctcgattg gcaaaaaaac gtgttgtaac cctagtaca ttgagcggaa ctagcaacc ttggcctga
701  ettaacttgg tatggtttgc tctctgcaat gtgttgctac ggacatcgtt acggttctg caacgggtt
771  gataattgac cgtctgatga atgagatcga agggccgttg ttaattatct gacctacctc cgcctatttc
841  aacgtctctgg tgaagacggc agccgggaag gccgaccgac caataaaga ctatttagac ctggccact
911  cgaacccaga ggcacatagt aacgtctgta ccccggttct ccttcggga gggcatagca tcaatagagc
981  tcttgcctct cagtcctgtg atacctactt gctttatctg tctagcagct ctatccacgg agtgactaat
1051  tctgaacacat tgaagtcctg gttcaaatga gtatatatga aatctacta aattttgaa taaaaattaa
1121  attttctctg atccacttct agaaaaaact attagagtao tggttttagg gaattgcaat caaaagcaag
1191  gtgactctga gttctgggga tctttcttag ttctctagaa gaactctagg aaaaaagac gggcattaga
1261  cgaacgaact tcttttttct gttggcgatg gtctgccaca acaaacgggc ctagtctcty atggttgaga
1331  aaaaagcttc cattgacga agtctgtctg cgtctatggt ttatgacagg aagatcccat cggcatcaat
1401  ccgggtggtga agttcttgag acatagtgge ggtatgtatg agcgagcgga ttaggacaaat ggtccacgac
1471  gacggtcacc gctattcagc acagaatgga ccaacctgag ttctgtctac aatggcctat tccggtctgc
1541  cagcccgact tgcctcccaa gcaagtgtgt cgggtctgac ctgcttctgt ggtgtgtggt tgaactctatg
1611  gatgtctgac tggatactct ttcggggtgc gaagggcttc cctctttctg cctgtccata gcccatttgc
1681  cgtccacggc ttgtctctct ggtgtctctc tcaaggttcc ccttttgagg acatagaaa tatccggaca
1751  gcccnaagcg gtggagactg aactcggagc taanaaact acgagcagtc ccccgcctc ggaaccttt
1821  ttgoggtcgt tgcgcggga aatgcgaag gaccggaaaa cgaacggaaa acgagtgtac aagaaagac
1891  gcaatagggg actaaagac ctatttgcat aatggcgga actcactega ctatggcgag cggcgtcggc
1961  tctgtggctc ggtctgctca gtcactcgt ccttcgctt ctgcgggtt atcggtttgg cggagagggg
2031  cgcgcacccg gctaagtaat tacgtcgacc gtgtgtctca aagggctgac ctctcgccg tcaactcgt
2101  tgcgttaatt acactcaatc gagtgaagta tctgtgggt cgaatgtg aaatacgaag gcccagcata
2171  caacacacct taccactcgc ctattgttaa agtgtgtct ttgtcgatac tggtaataat ggggttggag
2241  tgcgattatg ctgagtgata tctttgtcgg gtgttacct agggcccat ggtctgagct taagtgaacg
2311  gcagcaaaat gttgcagcac tgacctttt gggacggcaa tgggttgaat tagcggagc tegtgtaggg
2381  ggaagcggt cgcgcgactt atcgtctctc cgggcgtggc tagcgggaag gtttgtcaac ggttcggact
2451  taccgcttac cgtggactac gccataaag aggaatgcgt agacacgcca taagtggtg cgtataccac
2521  gtgagagtcg tttagacga gactacggcg tataaatctg gtccgggctg tgggcggtt tgggcgactg
2591  cgcgggactg ccgaacaga cgaagggcgt aggaatgt ctgttcgaca ctggcagagg ccttcgagct
2661  acacagcttc caaaagtgga agtagtggt ttgcgctc

```

(SEQ ID 154)

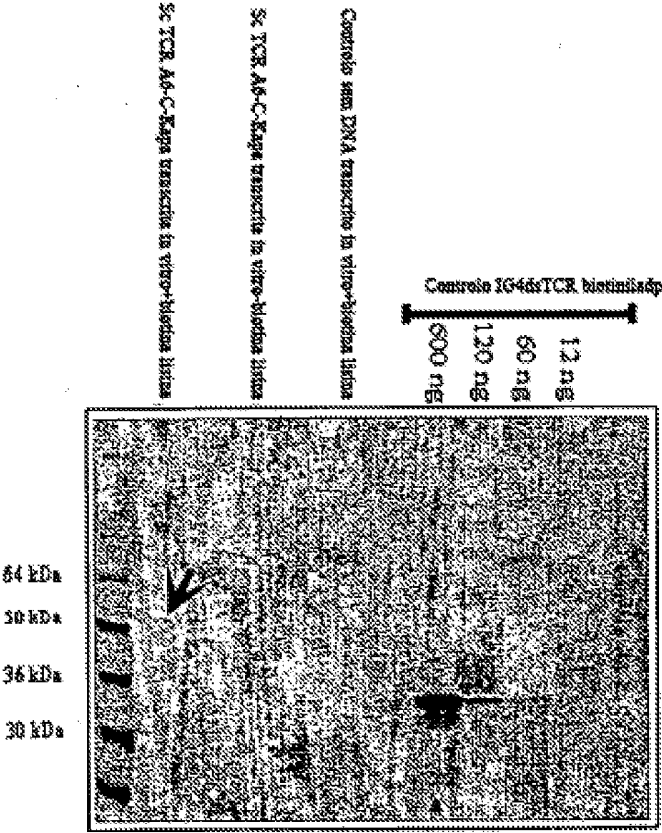
Figura 9

Sc TCR A6-C-Kapa clonado em pUC19-T7

```
1 ccattgggccc gaagggaagt gagcagaact ctggagccct cagtgttcca
51 gaggggagccc ttgcctctct caactgaact taagtgacc gaggttccca
101 gtccctcttc tggtaagagc aatattcttg gaaaagccct gaggttgtaa
151 tgtccatata ctccaatggt gacaaagag atgggaaggt tacagtcacg
201 ctcaataaag ccagccagta tgtttctctg atcatcagag actcccaagc
251 cagtgattca gccacctacc tctgtgcctg tacaactgac agctggggga
301 aattgcagtt tggagccagg acccaggttg tggtaaccgg tgagggcgg
351 tcaggcgagag gtggatccgg cggtagcggg tcgaagctg gtgtcactca
401 gaccccaaaa ttccaggtcc tgaagccagg acagagcatg acaatggcag
451 gtgcccaagg tatgaacct gaatacaktg ctgggtatcg acaagaccca
501 ggcctggggc tgaggctgat taattaccca gttgggtctg gtatcactga
551 ccaaggagaa gtccccaatg gctacaatgt ctccagatca acccagagg
601 atttcccgct caggctgctg tgggtgctc ctcccaagc atctgtgtac
651 ttctgtgcca gcaggccggg actagcggga gggcgaccag agcagtaatt
701 cgggcggggc acccaggtca cggtaacaga ggaactgaaa aacgtgttcc
751 cacccgaggt cgtgtgtttt gagccatcag aagcagagat ctcccaacc
801 caaaaggcca cactgtgttg ctggcccca ggtttctacc cggaccacgt
851 ggagctgagc tggtaggtga atgggaagga ggtgcacagt ggggtcagca
901 cagaccccca gccctcaag gagcagccc cctcaatga ctccagatcc
951 gctctgagca gccgcttag ggtctcggcc accttctggc aggacccccc
1001 caaccacttc cgtctgcaag tccagttcta cgggtctctg gagaatgacg
1051 agtggaccca ggtatgggac aaaccogtca cccagatcgt cagcgcagag
1101 gcttggggtc gagcagaggg tggaggcggg taactagaca gcaccctgac
1151 gctgagcaaa gcagactacg agaaacacca agtctacgcc tgcgaagtca
1201 cccatcaggg cctgagttcg ccgtcacca agagcttcas ccggggagag
1251 tctaagaaat tc
```

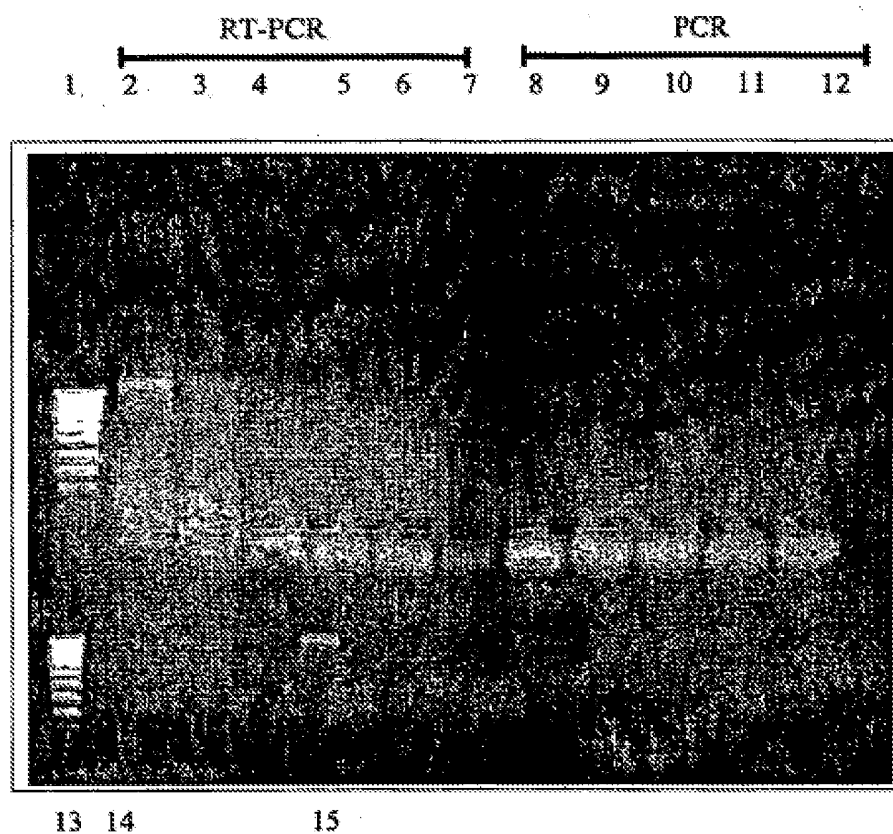
(SEQ ID 155)

Figura 10



A proteína sc TCR A6-C-Kapa está indicada na transferência Western acima com uma seta

Figura 11



Pista 1 Marcador de DNA de 100 pb Bioline

Pista 2 Reacção A6 sc TCR-C-Kapa seleccionada contra esferas HLA-A2 TAX

Pista 3 Reacção A6 sc TCR-C-Kapa seleccionada contra esferas HLA-A2 TAX na presença de 10 microgramas de A6 sc TCR solúvel

Pista 4 Reacção A6 sc TCR-C-Kapa seleccionada contra esferas controlo

Pista 5 Reacção controlo sem DNA seleccionada contra esferas HLA-A2 TAX

Pista 6 Reacção controlo sem DNA seleccionada contra esferas HLA-A2 TAX na presença de 10 microgramas de A6 sc TCR solúvel

Pista 7 Reacção controlo sem DNA seleccionada contra esferas controlo

Pistas 8-12 e pista 13 são as pistas 2-7 sem adição de transcriptase reversa apenas Taq de alta fidelidade da Roche. Estes são controlos da contaminação de DNA.

Pista 13 Controlo positivo da RT-PCR

Figura 12a

Clone 9 Sequência de DNA de cadeia β de TCR A6 mutada

```
gctgggtgtcacicagacccccaaaattccagggtcctgaagacaggacagagcatgacactgcagtggtgccaggatatgaacat
gaatacatgtcctgggtatcgacaagaaccagggtatggggctgaggctgattcattactcagttgggtctgggtatcaatgaccaagga
gaagtcctcaatgggtacaaatgtctccagatcaaccacagaggatttccggctcaggctgctgtcggctgtctccctccagacatct
gigtactctgtgcccagcaggccgggactagcgggaggggtgaccagagcagtagcttcgggcccgggcccagggtcaccgggtcac
agagggacctgaaaaacgtgttcccacccggggctgctgtgtttgagccatcagaagcagagatctccacacccaaaaggcca
cactgggtgtcctggccacagggttctaccccgaccacgtggagctgagctgggtgggtgaatgggaaggaggtgcacagtgggg
tctgcacagacccgcagccctcaaggagcagcccgccctcaatgactccagafacgtctcagcagcccgctgaggggtctcgg
ccaccttctggcaggacccccgcaacccttcgctgtcaagtcagttctacgggctctcggagaatgacgagtggaaccaggga
tagggccaaacctgcaccagatcgtcagcggcggggctggggtagagcagac
(SEQ ID 156)
```

Figura 12b

Clone 9 Sequência de aminoácidos de cadeia β de TCR A6 mutada

```
AGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPGMGLRLIHYSVGAGITDQGEVP
NGYNVSRSTTFPLRLLSAAPSQTSVYFCASRPGLAGGXPEQYFGPGTRLTVTEDLKNVF
PPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWWVNGKEVHSGVCTDPQLKEQPA
LNDSTRYALSSRLRVSATFWQDPRNHFRQVQFYGLSENDEWTQDRAKPVTOIVSAEAWGR
AD
(SEQ ID 157)
```

X - Representa a posição de aminoácido correspondente ao codão de paragem "opal" introduzido, isto resulta, regra geral, na substituição de um resíduo triptofano (w) na cadeia β de TCR neste ponto.

Figura 13

Clone 49 Sequência de DNA de cadeia β de TCR A6 mutada

```
gctgggtgacacagacccccaaattccagggtcctgaagacaggacagagcatgacactgtagtgtgccaggatatgaacat  
gaatacaatgtcctgtatcgacaagaccaggcaggggctgaggctgattcattacacagttgggtgtggtatcactgaccaagga  
gaagtcaccaatgggtacaatgtctccagatcaaccacagaggatttccgcacaggctgtgtcggctgtctccctccagacatct  
gtgtacttctgtgccagcagggccgggactagcggggagggcgaccagagcagtaactggggccgggcaaccagggtcacgggtcac  
agaggacctgaaaaaogtgttccacccgaggctgcctgtgtttgagccatcagaagcagagatctccacaccccaaaaggcca  
cacgtgtgtgtcctggccacaggcttclaccccgaccacgtggagctgagctgggtgggtgaatgggaaggagggtgcacagtgagg  
tcgtcacagaccggcagccctcaaggagcagcccgccctcaatgactccagatacgcctcagagcagccggctgaggggtctcgg  
ccaccttctggcagggaaccccgcaaccacttccgctgtcaagtcagttctacgggctctcggagaatgacgagtggaacccagga  
tagggccaaacccgtcacccagatcgicagcgcggggcctggggtagagcagac  
(SEQ ID 158)
```

Figura 14a

Clone 134 Sequência de DNA de cadeia β de TCR A6 mutada

```
gctgggtgctcactcagacccccaaaattccagggtcctgaagacaggacagagcatgacacigcagtgctgccaggatgaaccat
gaatacatgtcciggtatcgacaagacccagggtcctgggtcagggtcgtattactcagttgggtgctggtatcactgaccaagga
gaagtcocccaatggctacaatgtctccagatcaaccacagaggatitcccgctcagggtcgtgctggctgctocctccagacatct
gigtactctgtgctcagaggccgggtgatgagtgcggaaccagagcagtagtctgggtccggccaccagggtcacgggtcac
agaggacctgaasaacgtgttccacccagggtcgtgctgtgttgagccatcagaagcagagatctccacacccaaaaggcca
cactgggtgctgctggccacccgttctaccccgaccacgtggagctgagctgggtgggtgaatgggaaggagggtgcacagtggggt
ctgcacagaccccgagccctcaaggagcagcccgccctcaatgactccagatcgcctcagagcagccgctgaggggtcgg
ccacctctggcaggaccccccgaaccacttcgctgcaagtcagttctacgggtctcggagaatgacgagtggaaccagga
tagggccaaaacccgtcacccagatcgtcagcgcggaggccctgggttagagcagactaagcttgaattc
(SEQ ID 159)
```

Figura 14b

Clone 134 Sequência de aminoácidos de cadeia β de TCR A6 mutada (BIAcore)

```
MNAGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPGMGLRLIHYSVGAGITDQG
EVPNGYNVSRSTTEDFPLRLLSAAPSQTSVYFCASRPGLMSAEPEQYFGPGTRLTVTEDLK
NVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWWVNGKEVHSGVCTDPQPLKE
QPALNDSRYALSSRLRVSATFWQDPRNHFRQVQFYGLSENDEWTQDRAKPVTVQIVSAEA
WGRAD*
(SEQ ID 160)
```

Figura 14c

Clone 134 Sequência de aminoácidos de cadeia β de TCR A6 mutada (ELISA)

```
AGVTQTPKFQVLKTGQSMTLQCAQDMNHEYMSWYRQDPGMGLRLIHYSVGAGITDQGEVP
NGYNVSRSTTEDFPLRLLSAAPSQTSVYFCASRPGLMSAQPEQYFGPGTRLTVTEDLKNVF
PPEVAVFEPSEAEISHTQKATLVCLATGFYPDHVELSWWVNGKEVHSGVCTDPQPLKEQPA
LNDSTRYALSSRLRVSATFWQDPRNHFRQVQFYGLSENDEWTQDRAKPVTVQIVSAEAWGR
AD
(SEQ ID 161)
```

Figura 15

Clone 134 de TCR A6

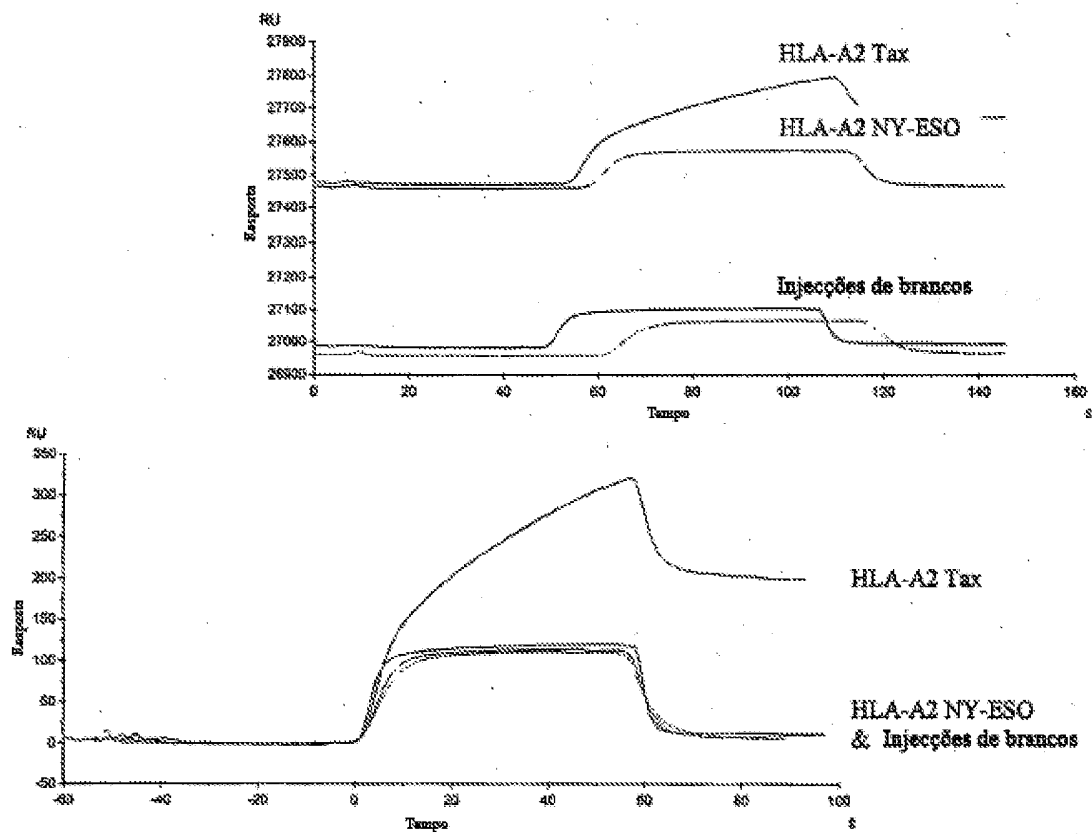


Figura 16

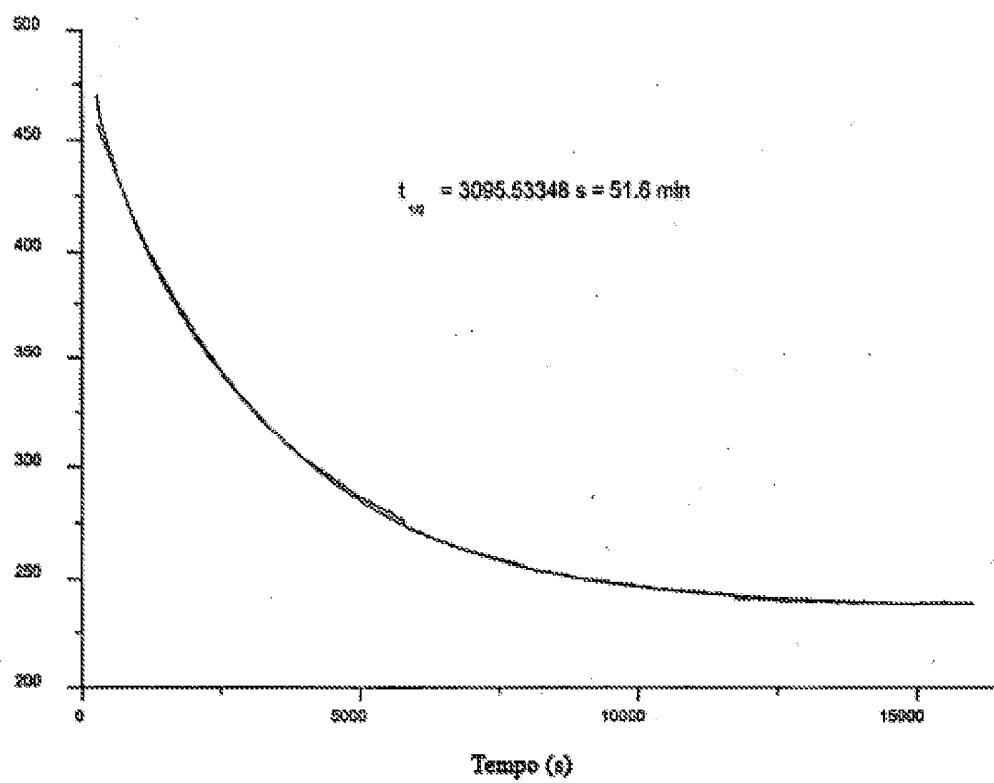


Figura 17a

atgcaggagggtgacacagattcctgcagctctgagtgcccagaaggagaaaacttgggttctcaactgcagtttactgata
gcgctatttacaaccctcagtggttaggcaggacctgggaaagggtctacatctctgttgcctattcagtcgaagtcagaga
gagcaaacaaagtggagacttaatgcctcgcctggataaatcaccaggacgtagtacittatcatigcagcttctcagccctgg
tgactcagccacctaccctctgtgctgtgaggcccacatcaggaggaagctacataacctacatttggagaggaaccagcct
tattgttcatccgtatctcagaaccttgacctggcctgttaccagctgagagactctaaatccagtgacaagtcgtctgcct
attcaccgattttgattctcaaacaaatgtgtcacaaagtaaggattctgatgttatatcacagacaaatgtgtgctagacatg
aggctctatggacttcaagagcaacagtgctgtggccgggagcaacaatctgactttgcatgtgcaaacgccttcaacaaca
gcattattccagaagacaccttcttccccagcccagaaaagttcttaa
(SEQ ID 162)

Figura 17b

atgggtgtcactcagaccccaaaattccaggctcctgaagacaggacagagcatgacactgcagtggtcccaggatatgaa
ccatgaatacatgtcctgggtatcgacaagaccaggcatggggctgaggctgattcattactcagttggtgctgggtacacig
accaaggagaagtcctcaatggctacaaatgtctccagatcaaccacagaggatttcccgcfcaggctgctgtcggctgctc
cctcccagacatctgtgtacttctgtgccagcagttacgtcgggaacaccggggagctgttttttggagaaggctctaggctg
accgtactggaggacctgaaaaacgtgttcccacccagggtcgtgtgtttgagccatcagaagcagagatctccacacc
caaaaggccacactggtgtgcttggccacaggcttctaccccaccacgtggagctgagctgggtgggtgaatgggaagg
agggtcacagtggggctgtgcacagaccgcagccctcaaggagcagcccgccctcaatgactccagatacgtctgtgag
cagccgctgaggggtctcgccaccttctggcaggacccccgaaccacttccgtgtcaagtcagttctacgggcctc
ggagaatgacgagtgaggccaggatagggccaaaccgtcaccagatcgtcagcggcgaggcctggggtagagcag
actaa
(SEQ ID 163)

Figura 18a

MQEVTQIPAAALSVPEGENLVLNCSFTDSA IYNLQWFRQ
DPGKGLTSLLLIQSSQREQTSGRLNASLDKSSGRSTLYI
AASQPGDSATYLC AVRPTSGGSYIPTFGRGTS LIVHPYI
QNPDP AVYQLRDSKSSDKSVCLFTDFDSQTNVSQSKDS
DVIITDKCVLDMRSMDFKNSNAVAWSNKSD FACANAF
NNSIIPEDTFFPSP ESS Stop
(SEQ ID 164)

Figura 18b

MGVTQTPKFQVLKTGQSM TLQCAQDMNHEYMSWYRQ
DPGMGLRLIHYSVGAGITDQGEVPNGYNVSRSTTEDFP
LRLLSAAPSQTSVYFCASSYVGNTGELFFGEGSRLTVLE
DLKNVFPPEVAVFEPSEAEISHTQKATLVCLATGFYPDH
VELSWWVNGKEVHSGVCTDPQPLKEQPALNDSRYALS
SRLRVSATFWQDPRNHFR CQVQFYGLSENDEWTQDRA
KPVTQIVSAEAWGRAD Stop (SEQ ID 165)

Figura 19a

```

TTCTGCGCCT TTTCGTGGCC TTTCGCTCAC ATGTAATGTG AGTTAGCTCA
AAGGACCGGA AAACGACCGG AAAACGAGTG TACATACAC TCANTCGAGT
CTCATTAGGC ACCCCAGGCT TTACACTTTA TGCTTCGGC TGGTATGTTG
GAGTAATCCG TGGGTTCCGA AATGTGAAT ACGAGGGCG AGCATACAAC
TGTGGAAATG TGAGCGGATA ACAATTTTAC ACAGGAAACA GCTATGACCA
ACACCTTAAC ACTGCGCTAT TGTAAAGTG TGTCTTTGT CSATACGTGT
TGATTACGCC AAGCTACGTA CTAAAGTAT CTATTTCAAG GAGACAGTCA
ACTAATGGG TTCSATGCT GAATTCATTA GATAAAGTTC CTCTGTCACT
TAATGAATAA CCTATTTGCT ACGGACGGCG CTGGATTTGT ATTACTCGCG
ATTACTTTAT GGATAACGGA TGGCTTCGCG GACCTAACAA TAATGAGGCC
GCCCAGCCGG CCATGGCCAA ACAGGAGGTS ACGCAGATTC CTGCAGCTCT
CGGCTCGGCC GGTACCGGTT TGTCTCCAC TGGCTCAAG GACGTGGAGA
CAGTGTCCCA GAAGGAGAAA ACTTGGTCTT CAACTGCAGT TTCACTGATA
CTCAGAGGCT CTTCCTCTTT TGAACCAAGA GTTGACGTCA AAGTACTAT
GCGCTATTTA CAACCTCCAG TGGTTTAGGC AGGACCTTGG GAAAGGCTTC
GCGCTAATAA GTTGGAGTTC ACCAAATCCG TCTTGGGACC CTPTCCAGAG
ACATCTCTGT TGTATTTC AITCAAGTCAG AGAGAGCAA CACTGGAG
TGTAGGACA TCGAATAGT CAGTTCAGTC TCTCTCGTTT GTTCACTTTC
ACTTAATGCC TCGTGGATA AATCTCAGG ACGTAGTACT TTATCATTTG
TGAATPACGG AGCGACCTAT TTAGTAGTCC TGCATCATGA AATAGTAAAC
CAGCTTCTCA GCTGTGTAC TCAGCCACTT ACCTCTGTGC TGTAGGCCCC
GTGGAAGAT CGGACCACTG AGTGGTGA TGGAGACAG ACACCTCGGG
ACATCAGGAG GAAGCTACAT ACCTACATTT GGAAGAGGAA CCGGCTTAT
TGTAGGCTCT CTTCGATGTA TGGATGTAAA CTPTCTCTTT GGTGGGATA
TGTTCATCCG TATATCCAGA ACCCGGATCC TGCCGTGTAC CAGCTGAGAG
ACAAATAGGC ATATAGTCTT TGGGCTTAGG ACGGCAGATG GTTCACTCTC
ACTCTAAATC CAGTGACAAG TCTGTCTGCC TATTCACCGA TTTTGAATCT
TGAGATTTAG CTCACCTGTC AGACAGACGG ATAGTGGCT AAACTAAGA
CAACAAATG TGTCAAAAG TARGERTTCT GATGTGTATA TCACAGACAA
GTTGTTTAC ACAGTGTTC ATTCCTAAGA CTACACATAT AGTGTCTGTT
ATGTGTCTCA GACATGAGGT CTATGACTT CAGAGACAAK AGTGTGTGTG
TACACAGAT CTGTACTCCA GATACCTGAA GTTCTGTGTT TCACGACACC
CCTGGAGCAA CAATCTGAC TTTGCACTG CAAACGCTTT CACACACAGC
GGAGCTCTTT GTTTAGACTG AAACCTACAC GTTTCGGAA GTTGTGTGTG
ATTATTCAG AAGACAGCTT CTTCGCCAGC CCAGAAAGTT CTAATTAACC
TAATPAGGTC TTCTTTGAAA GAAGGGGTGG GGTCTTTCAA GWTATATAG
TAGGTTAAT AAGAAATCTT TAAGAAAGGG ATATACATAT GAAATAATTA
ATCCAAITAA TTCTTAAGAA ATTCCTCCGC TATATGATA CTPTTTAT
TTATTCGAA TTCTTTAGT TGTTCCTTTC TATTCACACA GTTGTGTGTG
AATAAGCTT AAGGAAATCA ACAAGSAAAG ATAAAGAGTT CCGGCTTCCG
TGTGTCACT CAGACCCCAA AATTCAGGTT CTGGAAGACA GGCACAGACA
ACCACATGTA GTTGGGGTT TTAAGTCCA GGACTTCTGT CCGTCTCTGT
TGACACTGCA GTTGGCCAG GATATGAACC ATGAATACAT GTCTGTGTAT
ACTGTGACGT CACACGGGTC CTATACTTGG TACTTATGTA CAGGACNTA
CGACAGACCC CAGGCATGGG GCTGAGGCTG ATTCATTAAT CAGTTGTGTC
GCTGTCTGCG GTTCTTACCC CCACTCCGAC TAAGTAATGA GTCAAGCAGG
TGTATCACT GACCAAGGAG AAGTCCCAA TGGCTACAAT GTCTCCAGAT
ACCATAGTGA CTGGTTCTTC TTCAGGGGTT ACCGATGTTA CAGAGTCTTA
CAACCACAGA GGAATTTCCG CTCAGGCTGC TGTGGGCTGC TCCCTCCAG
GTTGGTGTCT CPTAAGGGC GAGTCCGACG ACAGCCGAGC AGGAGGGTTC
ACATCTGTGT ACTTCTGTGC CAGCGTTTAC GTCGGGAACA CCGGGGAGCT
TGTAGACACA TGAAGACAG GTCTGTCAATG CAGCCCTTGT GGGCTCTGSA
GTTTTTTGSA GAAGGCTCTA GGCTGACGCT ACTGGAGGAC CTGAAAAAGC
CAAAAAACCT CTTCGAGAT TCGACTGCGA TGACTTCTGT GACTTPTTGT
TPTTCCACCC CGAGTTCGCT GTGTTTGAAG CATTAGAAGC AGGATCTCC
ACAAGGGTGG GCTCCAGGSA CACAACTCG GTAGTCTTGG TCTCTAGAGG
CACACCCAAA AGGCCACACT GGTGTGCTTG GCCACAGGCT TGTACCCGSA
GTGTGGGTTT TCCGCTGGA CCACAGGGAC CGGTGTCCGA AGATGGGGCT
CCACCTGGAG CTGAGCTGAT GAGTGAATGG GAAGGAGGTC CACAGTGGG
GGTGCACCTC GACTGGACCA CCCACTTACC CTTCCTCCAC GTGTACCCCT
TCTGCACAGA CCCGACGGC CTCAAGGAGC AGCCCGGCTT CAATGACTCC
AGACTGTCTP GGGCTTCGG GAGTTCTCTG TCGGGCGGGA GTTACTGAGG
AGATACGCTP TGAGCAGGCG CCTGAGGCTC TCGGCGACCT TGTGGCAGSA
TCTATGCGAG ACTCGTGGG GSACTCCAG AGCCGTTGSA AGACCTTCTT
CCCCCGCAAC CACTTCCGCT GTCAAGTCCA GTTCTACGGG CTCTCGAGA
GGGGGCTTGT GTGAAGGCGA CAGTTCAGGT CAAGATGCCC GAGAGGCTCT
ATGACGAGTG GACCCAGGAT AGGGCCAAAC CCGTACCCCA GATCGTCAGC
TACTGCTCAC CTGGTCTCTA TCCCGGTTTG GGCAGTGGGT CTAGCAGTGC
GCCAGGCTCT GGGGTAGAG AGACGGGGCC GCA
GGGTCCGGA CCDCMCTGG TCTGCGCGGG CGT

```

(SEQ ID 166)

Figura 19b

K Y L L P T A A A G L L L A S A Q P P
A M A K Q E V A T C Q I P A S S N I L
G E N L V L D N C S F T D S G R L S A
S W F R Q E Q L Q G R L S A G R L S A
S S Q R S T L Q L S I P T S K A S T
Y L C A V R L I V H S P S G A S T G
R G T S L D S K S D Y I Q S V C D P
Y Q L R D N V S S Q S K D S D V F T
F D K C V L D M R S M D F K S N S I
D W S N K S D F P A C A N A F S I
A P E D T F V V P S P E S S M K L
I P A I F L V Q H Y S P Y S A Q K L
F Q T P K F Q H Y S V M S G W Y Q M
T A Q D M N I H Y S R S A T G Y R Q
C G L R L I Y N V S R S T V T Y F T
M G P N G A A P S Q T S V T Y F D
E V L S A A P S Q T S V T Y F D C
R L L S A A P S Q T S V T Y F C F
Y V G N T G E L F P P G E V G S R
L E D L K N V F P P E V G S R F A
E A E I S H V T Q K P A T L V C L
F Y P D H V T Q K P A T L V C L E
S G V C T D S P S C Q R L V N G P
S R N H F D S S C Q R L V N G P A
P W T Q A D R A A A A A A A A A A
E W T Q A D R A A A A A A A A A A
W G R A D A A A A A A A A A A A

(SEQ ID 167)

Figura 20

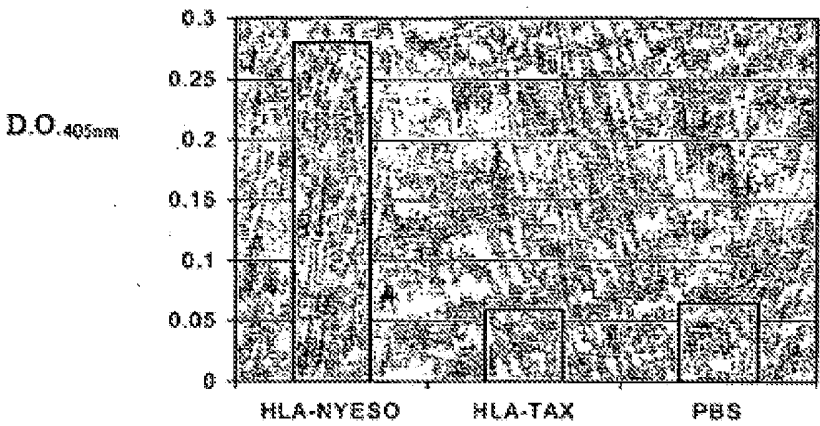


Figura 21

DRA0101

```

1  ggatccatgg ccataagtgg agtccctgtg ctaggatttt tcatcatagc tgtgctgata
61  agcgctcagg aatcatgggc tatcaaaaga gaacatgtga tcatccaggg cgagttctat
121 ctgaatcctg accaatcagg cgagtttatg tttagctttg atggtgatga gattttccat
181 gtggatatgg caaagaagga gacggtctgg cggcttgaag aatttgagcg atttgccagc
241 cttgaggctc aagggtgcatt ggccaacata gctgtggaca aagccaacct gpaatcatg
301 acaaaagcgt ccaactatac tccgatcacc aatgtacctc cagaggtaac tgtgctcagc
361 aacagccctg tggaaactgag agagcccaac gtccctcatc gtttcatcga caagttccac
421 ccaccagtgg tcaatgtcac gtggcttoga aatggaaaac ctgtcaccac aggagtgtca
481 gagacagtct tectgcccag ggaagaccac cttttccgca agttccacta totcccttc
541 ctgcccctaa ctgaggacgt ttagcactgc aggggtggagc actggggctt ggatgagcct
601 cttctcaagc actggagatt tgatgttcca agccctctcc cagagactac agagaacgtg
661 gatgggggtc ttgactgatac tctccagccg cagacagatc caactcaaga caagaaactc
721 gctgtctaga ccaagattc caatctctg aaagagaagg aaaaactaga gttcattcga
781 gcagcttacc gatctggtag tggctgaaac gatatttttg aagctcagaa aatcgaaatg
841 catgagttag atcc

```

(SEQ ID 168)



- Codões de fecho de leucina Fos



- Codões da marca de biotinilação

Figura 22

```

1 aattctactc gtaaaagcag ttgaaggatc atatttagtt gcgtttatga gataagattg
61 aaagcacgtg taaaatgttt cccgcgcgtt ggcacaacta ttacaaatgc ggccaagtta
121 taaaagattc taatctgata tgttttaaaa caoctttgcg gcccgagttg tttaggtacg
181 tgactagcga agaagatgtg tggaccgcag aacagatagt aaaaacaaac cctagtattg
241 gagcaataat cgatttaacc aacacgtcta aatattatga tgggtgtcat tttttgoggg
301 cgggcctgtt atacaaaaaa attcaagtac ctggccagac ttgccgcctt gaaagcatag
361 ttcaagaatt tattgacacg gtaaaagaat ttacagaaaa gtgtcccgge atgttggttg
421 gcgtgcactg cacacacggt attaatcgca cggttacat ggtgtgcaga tatttaatgc
481 acacccctgg tattgagccg caggaagcca tagatagatt cgaaaaagcc agaggtcaca
541 aaattgaag acaaaattac gttcaagatt tattaattta attaatatta tttagattct
601 ttaacaaata etttatccta ttttcaaat gttagccttc tccagcgaa ccaaaactat
661 ctccgcttg ctccgtttag ctgttagcgg atcagtgggg ttgttccaa cgacggtagg
721 attagggcgg atactctcca ccacatgtt ggcacgttg atgttacgtt tatgtctttg
781 gttttccacg taagtctttt ggccggtaat agccgtaaac gtagtgcctt ccgcggtcac
841 gcacaacacc ggaatgtttg ctttgtccgc ggggtattga accgcgcgat ccgacaaatc
901 caccactttg gcaactaaat cgggtgacct cgggtctttt ttctgcatta ttctgtcttt
961 cttttgcatg gtttcctgga agccggtgta catgcggttt agatcagtca tgacgcgctt
1021 gacctgcaaa tttttggcct cgatctgctt gtccttgatg gcaacgatgc gttcaataaa
1081 ctcttgtttt ttaacaagtt cctcggtttt ttgcgcacac accgcttgca gcgcgtttgt
1141 gtgtcgggtg aatgtcgcaa tcagcttagt caccaactgt tgcctctcct cctcccgttg
1201 tttgatcgcg ggaatcgtae tgcgggtgca gagcacttga ggaattactt ctctcaaaag
1261 ccatctctgt aattctatgg cgtaaaggcaa tttggacttc ataactagct gaatacgcgc
1321 ggatttagta atgagcactg tatgcgggtg caaatcacagc ggggtcgccc ttttccagac
1381 gctgttagag gtaggggccc cattttggat ggtctgtctc aataacgatt tgtattttat
1441 gttcacatga acacgtatag ctttatcaca aactgtatat tttaaacgtt tagcgcgctc
1501 cttggccacg aaccggacct gttggtcgcg ctctagcaag taccgcaagt tgaacgtatc
1561 ttctccaaat ttaaatcttc caattttaac gcgagccatt tggatacagc tgtgtcgatt
1621 ttgcaacaac tattgttttt taacgcacac taactttatt gttgtaagca ataattaaat
1681 atgggggaac atggcccgct acaacactcg tegtatatga ccgacagcgt ccggtctctg
1741 gcgcaagcgg ctaaaacgtg ttgcgcgttc aacgcggcaa acatcgcaaa agccaatagt
1801 acagtttttg tttgcatatt aacggcgatt ttttaaatia tcttatttaa taaatagtta
1861 tgagccctac aactcccgce ccgcgttgac tcgctgcacc tcgagcagtt cgttgacgcc
1921 tctctccgtg tggcgaaca cgtcgagcgg gtggtcgatg accagcggcg tgcgcacgcg
1981 gacgcacaaq tatctgtaca ccgaatgatc gtcgggcgaa ggcacgtcgg cttccaaagt
2041 gcaatatttg caaatccgaa aatataatac gttgggttgt ttggcagatc ctatcgtggc
2101 gttgggcatg taagtccgaa cgttgatttg catgcaagcc gaatttaaat catlgcgatt
2161 agtgcgatta aaacgttgta cctctcgcct tttatcaty ccgtcgatta aatcgcgcaa
2221 tcgagtcacg tgatcaaatg gtggaataat gttttctttg tatteccgag tcaagcgcag
2281 cgcgtatttt aacaaactag ccatcttgta agttagtttc atttaatgca actttatcca
2341 ataataatatt atgtatcgca cgtcaagaat taacaatgoc ccggttgctg catctcaaca
2401 cgaatatgat agagatcaaa taagcgcgca attaaatagc ttgcgacgca acgtgcacga
2461 tctgtgcacg cgttccggca cgaactttga ttgtaataag tttttacgaa gcgatgacat
2521 gaccccgta gtgacaaoga tcaacgcccac aagaactgcc gactacaaaa ttaccagata
2581 tgtcggtagc gttaaaaacta ttaagccatc caatcgaccc ttagtcgaat caggacogct
2641 ggtgcgagaa gccgcgaagt atggcgaatg catcgtataa cgtgtggagt ccgttcatta
2701 gagcgtcatg tttagacaag aaagctacat atttaattga tcccagatgat tttattgata
2761 aatbgacct aactccatac aagggtattct acaatggcgg ggttttggtc aaaaattccg
2821 gactgcgatt gtacatgctg ttaacggctc cgcacctat taatgaaatt aaaaattcca
2881 attttaaaaa accgcgcaag agaaatattt gtatgaaga atgcttagaa ggaagaaaa
2941 atgtcgtcga catgctgaac aacaagatta atatgcctcc gtgtataaaa aaaaatttga
3001 acgatttgaa agaaaadaat gtacccgcgg gcggtatgta caggaagagg tttatactaa
3061 actgttacat tgcaaacgtg gtttcgtgtg ccaagtgtga aaaccgatgt ttaatacagg
3121 ctctgaogca ttttacacac cadgactcca agtgtgtggg tgaagtcatg catcttttaa
3181 tcaaatccca agatgtgtat aaaccaccaa actgcacaaa aatgaaaact gtcgacaagc

```

```
3241 tctgtccggt tgcctggcaac tgcaggggtc tcaatectat ttgtaattat tgaataataa
3301 aacaaattata aatgctaast ttgtttttta ttaacgtatc aaaccaaacy caacaagaac
3361 atttgttagta ttatctatata ttgaaaaagc gtatgtataa tccgtgaggt aattatttaa
3421 atcatttttca aatgattcac agttaatttg cgacaatata attttatttt cacataaact
3481 agacgccttg tctgtttctt cttcgtatcc cttctctttt tcattttttt cctcataaaa
3541 attaacatag ttattatcgt atccatataa gtatctatcg tatagagtaa attttttgtt
3601 gtacataaata tatatgtctt ttttaattgg gtgtatagta ccgttgcgca tagtttttct
3661 gtaattraca acagtgcctat tttctgttag ttcttcggag tgtgttgcct taattattaa
3721 atttatataa tcaatgaatt tggggtcgtc ggttttctac aatatgttgc cggcatagta
3781 cgcagcttct tctagttcaa ttacaccatt tttagcagc accggtttaa cataacttct
3841 caaaatggtg tacgaaccgt taaacaaaaa cagttcacct cctttttct tactattgtc
3901 tgcgagcaggt tgtttgttgt taaaaataac agccattgta atgagacgca caaactaata
3961 tcacaaactg gaattgtcta tcaatataa gttgtctgat atccagcatg ataagataca
4021 ttgatcgagtt tggacaaaac acaactagaa tgcagtgaaa aaaaatgctt atttgtgaaa
4081 tttgtgatgc tattgtttta tttgtaacca ttataagctg caataaaca gttccgagtt
4141 tgtcagaaag cagacccaaa tgcgtgatat ggttgatgtc atgtagccaa atcgggaaaa
4201 aaataatacc cgtatcactt ttgtctgata taaaaaaagt aaagaaaaa gaataaacgy aacatccaaa
4321 agtttgtgtt ttttaaatag tacataatgg atttccttac gcgaataacg ggcagacatg
4381 gcttccccg ttattattat ttttgacacc agaccaactg gtaatggtag cgaccggcgc
4441 tcagctggaa ttcgagcttg taattgtaaa taaaatgtaa ttacagtat agtattttta
4501 ttaatatata aatgatttga taataattct tttttaacta taatatattg ttgtgggttg
4561 aattaaaggt cccggcatcc tcaaatgcab aatttcatag tcccccttgt tgtaatgtat
4621 gcgtatttct gaatttttgt aaaatagcac acaagactcc aacgggtttg gcgttttatt
4681 ttcttgcctg actctagttt attaggcctc tagagatccg tatttatagg tttttttatt
4741 acaaaactgt tacgaaaaa gtaaaatact tattattttg cggagtggtt atcattttaa
4801 ttatctccat gatcaataa accatagaata aagggcccca cctttaattc aacccaacac
4861 aatataattat agttaaataa gaattattat caaatcattt gtatattaat taaaatacta
4921 tactgtaaat tacattttat ttacaatcac agatccccg ttttccggtta ttgtacatt
4981 tattaagcgc tagattctgt gcgttgttga ttacagaca attgttgtat gtattttaa
5041 aattcattaa atttataatc tttagggtgg tatgttagag cgaataatca atgattttca
5101 gcgtctttat atctgaattt aaatattaaa tctcaatag atttglaaa taggtttoga
5161 ttagtttcaa acaggggttg tttttccgaa ccgatggctg gactatctaa tggattttcg
5221 ctaacgcga caaaacttgc caaatcttgt agcagcaatc tagctttgtc gatattcgtt
5281 tgtgttttgt tttgtaataa aggttcgacg tcttcaaaa tattatgcg ttttgtattt
5341 ctttcacac cttcgttagt gtacaattga ctgacgttaa acacgttaa taaagcttgg
5401 acataattaa cctcggggtt gttagcttta ttaggccgat tctcgtctc gtcccaaccc
5461 tctcgttag agttgtctc cgaagacgat tttgcacag ccacacgaag cctatttaatt
5521 gtgtcggcta acacgtccgc gatcaaatct tagtttgagc tttttggaa tattctgat
5581 tgcgggctt tttggggggg ttcaatcta actgtgcccg attttaattc agacaacacg
5641 ttgaaagcg atgtgcagg cgttggtaac atttcagacg gcaaatctac taatggcggc
5701 ggtggtggag ctgatgataa atctaccatc ggtggaggcg caggcggggc tggcggcgga
5761 ggcggaggcg gagggtgtgg cgttgatgca gacggcggt taggctcaaa tgtctcttta
5821 ggcaacacag cggcacctc aactattgta ctggtttcgg gcgcggtttt tgggttgacc
5881 ggtctgagac gsgtgcgatt ttttctgtt ctaatagctt ccaacaattg ttgtctgtcg
5941 tctaaaggtg cagcgggttg aggttcctgc ggcattgggt gagcggggcg caattcagac
6001 atcgatggtg gtggtcgttg tggaggcgct ggaatgttag gcacgggaga aggtggtggc
6061 ggcggtgcg cgggtataat ttgttctggt ttagtttgtt cgcgcacgat tgtgggcacc
6121 ggcgaggcg ccgctggctg cacaacggaa ggtcgtctgc ttogaggcag cgttgggggt
6181 ggttgcaatt caatattata attggaatac aaatcgtaaa aactcgtat aagcattgta
6241 atttcgtat cgtttacagt gcgatattt aacaaccgct caatgtaagc aattgtattg
6301 taaagagatt gtctcaagct cggatcgatc cgcacgcgcg ataacaagc ttttcatttt
6361 tactacagca ttgtagtggc gagacacttc gctgtcgtcg acgtacatgt atgctttgtt
6421 gtcaaaaacg tctgtggcaa gcttttaaat atttaaaaga acatctctgt tcagaccac
6481 tgtgttgtcg taaatgttgt ttttgataat ttgcgttcc gcagtatcga caggttcaaa
```

6541 aaattgatgc gcatcaatct tgttgttccct attattgaat aaataagatt gtacagattc
6601 atattctacga ttctgtcatgg ccaccacaaa tctatccttg caaacgcttg tacaatttta
6661 cgaaaactgc aaaaacgtca aaactcggta taaaataatc aacgggctgt ttggcaaat
6721 atctatttta tcgcacaaag ccactagcaa attgtatttg cagaaacaa tttcggcgca
6781 caatttttaac gctgaagaaa taaaagtcca ccagttaatg agcagccacc caaattttat
6841 aaaaatctat ttaaatcagc gtcccatcaa caaccaagtg atcgtgatgg actacattga
6901 ctgtcccgat ttatttgaaa cactacaaat taaaggcgag ctttcgtacc aacttgtag
6961 caatatttatt agacagctgt gtgaagcgtc caacgatttg cacaagcaca atttcataca
7021 caacgacata aaactcgaac atgtcttata tttcgaagca cttgatcgcg tgtatgtttg
7081 cgattacgga ttgtgcaaac acgaaaactc acttagcgtg cagcagcgca cgttggagta
7141 ttttagtcgg gaaaaaatcc gacacacaa ctagcagctt tegtttgact ggtacggcgc
7201 gtgttaacat ccaagttgct aacggcgggc cgacacccat ttgaaaaaag cgaagacgaa
7261 atgttggact tgaatagcat gaagcgttgt cagcaatata atgacatttg cgttttaaaa
7321 cagtttcgta acgtttacgc tctgtacttt tctgtactgc taacaagata cactatagat
7381 tgttagactca caaattacaa acaaatbata aaacatgagt ttttgcgta aaaaagccac
7441 ttgttttacg agtagaattc ccagcttggc actggcgttc gttttacaa ctcgtgactg
7501 ggaanaaccc gggttaacc aacttaatcg ccttgacgca catccccctt tcgcagctg
7561 gcgttaatagc gaagaggccc gacccgatcg cctttcccaa cagtgcgca gcttgaattg
7621 cgaatggcgc ctgatgggtt attttctcct tagcatctg tgggtatttt ccaacggcat
7681 atggtgcact ctacgtacaa tctgtctctga tgcgcctag ttaagccagc ccgacaccc
7741 gccaaacccc gctgacggc cctgaacggc ttgtctgttc cgggcacccg cttacagaca
7801 agcgttgacc gtctccggga gctgcattgt tcagagggtt tcaacgtcat caccgaaacg
7861 cgcggacgaa agggcctcgt gatagccta tttttatagg ttaattgcat caataaatg
7921 gtctcttaga cgtcaggctg cacttttctg ggaatgtgc ggggaacccc ttttgttta
7981 ttttcttaaa tacattcaaa tatgtatccg ctcatgagac aataaacctg ataatgtctt
8041 caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttatccc
8101 ttttttggcg ctttttgcct tccctttttt gctcaccag aaacgttaaa gaaagttaaa
8161 gatgctgaag atcagttggg tgcacagctg ggttacatcg aactggatct caacagcggt
8221 aagatccctt agagtttttg ccccgaaag cgttttccaa tgaatgagac ttttaagtt
8281 ctgctatgtg ggcgggtatt atcccgatc gacgcggggc aagagcaact cggctgcgcg
8341 atacactatt ctcagaaatg cttggttgag tactcacag tcacagaaaa gcatctaacg
8401 gatggcatga cagtaagaga attatgcagt gctgcataa ccatgagtga taacactgcg
8461 gccaaacttac ttctgacaac gatcggagga ccgaaggagc taacggcttt tttgcacaa
8521 atgggggata atgttaactc ccttgatcgt tgggaacggg agctgaatga agccatacca
8581 aacgcagagc gtgacacac gatgcctgta gcaatggcaa caacgttgcg caactatta
8641 actggcgaa cacttactct agcttccgg caacaattaa tagactggat gggggcggt
8701 aaagtctcag gaccacttct ggcctcggcc ctccggcgtg gctggtttat tgcgtataaa
8761 tctggagccg gtgagcgttg gctcggggt atcattgcag cactgggggc agatggttaag
8821 cctcccgta tctagtttat ctacacgag gggagtccag caactatgga tgaacgaaat
8881 agacagatcg ctgagatagg tgcctcactg attaagcatt ggttaactgc agaccaagt
8941 tactcatata tacttttagat tgatttaaaa ctccattttt aatttaaaag gatctaggtg
9001 aagatccctt ttgataatct catgacaaa atcccttaa gtgagttttc gttccactga
9061 ggcctcagac ccgtagaaaa gatcaaaagg tcttcttgag atcctttttt tctgcgcgta
9121 atctgctgct tgcaaaaaa aaaaccacgg ctaccagcgg ttggtttgtt gccggatcaa
9181 gagctaccaa ctcttttttc gaaggttaact ggtttcagca gagcgagat accaaatact
9241 gtctttctag tttagccgta gttaggccac cacttcaaga actctgtagc accgcctaca
9301 tacctcgttc tctaatcct gttaccagtg gctgctgoca gtggcgata gtctgtctt
9361 accgggttgg actcaagagc atagttaacc gataaggcgc agcggctggg ctgaacgggg
9421 ggttcgtgca cacagccagc cttggagcga accgactaca cggactgag atacctacag
9481 cgtgagcatt gagaaagcgc cagcctccc gaaggagaaa agggggagag gtatccggta
9541 agcggcaggg tcggaacagg agagcgcaag tgaactgagc cagggggaaa cgcctggtat
9601 ctttatagtc ctgtcgggtt tgcacacctc tgaactgagc cctttttacg gttcctggcc
9661 tcaggggggg ggagcctatg gaaaaacgac agcaacggcg cctttttacg gttcctggcc
9721 ttttgctgga cttttgctca catgttcttt cctgogttat cccctgattc tctggataac
9781 cgtattaccc cttttgagtg agctgatacc gctgcgcgca gccgaacgag cgaagcagc
9841 gagtcagtga gcgaggaagc ggaagagcgc ccaatagcga aacgcctct cccgcgcgt
9901 tggcggattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag
9961 cgcaacgcaa ttaattgtgag ttagctcact cattaaggac ccagggttt acactttatg

10021 ctccggctc gtatgttggtg tggattgtg agggataac aatttcacac aggaacagc
10081 tatgacctg attacg

(SEQ ID 169)



- Locais de enzimas de restrição

Figura 23

```

1 ggtaccggat ccagcatggt gtgtatgaag ctccctggag gctccctgcat gacagcgcctg
61 acagtgacac tgatgggtgt gagctcccca ctggccttbt ccggagacac cggagcctaag
121 tacgtcaagc agaacacact gaaactggct tccgggtggcg gatctctagt tccacgcgggt
181 agtggaggcg gtgggttcgg agacacgcgt ccacgtttct tgtggcagct taagtttga
241 tgtcatttct tcaatgggac ggagcgggtg cggttgctgg aaagatgcat ctataaccaa
301 gaggagtcgg tgcgcttoga cagcgacgtg ggggagtcac gggcgggtgac ggaggtgggg
361 cggcctgatg ccgagtcact gaacagccag aaggacctcc tggagcagag gcgggcccgg
421 gtggacacct actgcagaca caactacggg gtgggtgaga gcttcacagt gcagcggcga
481 gttgagccta aggtgaactg gtatccttca asgaccagc cctgcagca ccacacctc
541 ctggtctgct ctgtgagtg tttctatcca ggcagcattg aagtcagggt gttccgggac
601 ggccaggaag agaaggctgg ggtgggtgct acaggcctga tccagaatgg agattggacc
661 ttccagaccc tgytgatgct ggaacacgtt cctoggagtg gagaggttta cactgccaa
721 gtggagcacc caagtgtgac gagccctctc acagtgaat ggaagcacc gctgaatct
781 gcacagagca aggtcgacgg aggcgggtgg ggtagaatcg ccggtctga ggaagagtg
841 ggaacttga agcgcagaa ctggagctg gcttcacag caatctgct aggaagag
901 gtggcagcgt ctatacaga gttctgaa tactaggatc c

```

(SEQ ID 170)



- Codões de fecho de leucina JUN



- Péptido carregado de HLA

Figura 24

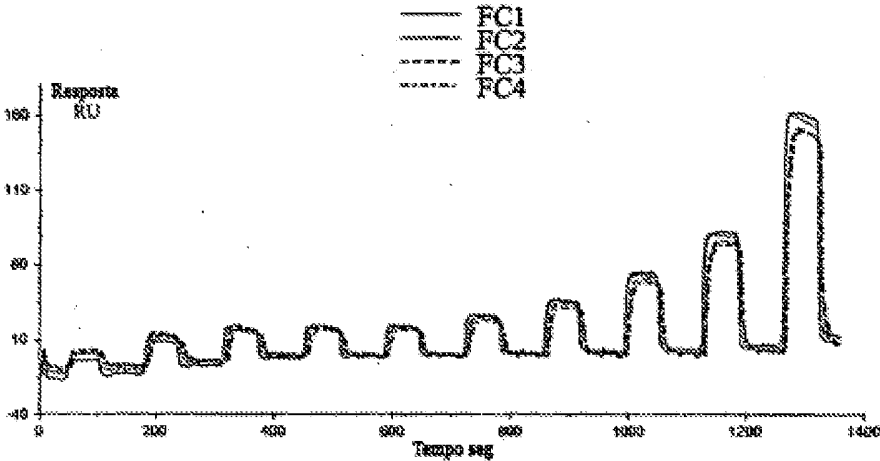


Figura 25

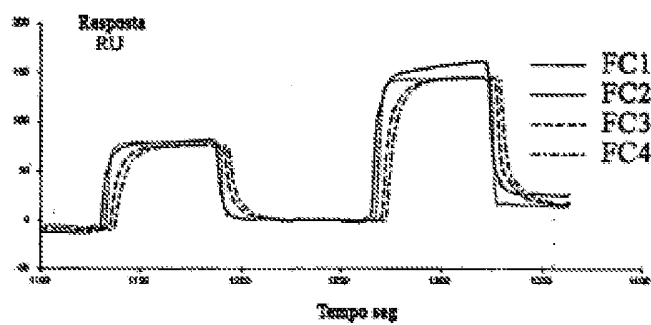


Figura 26

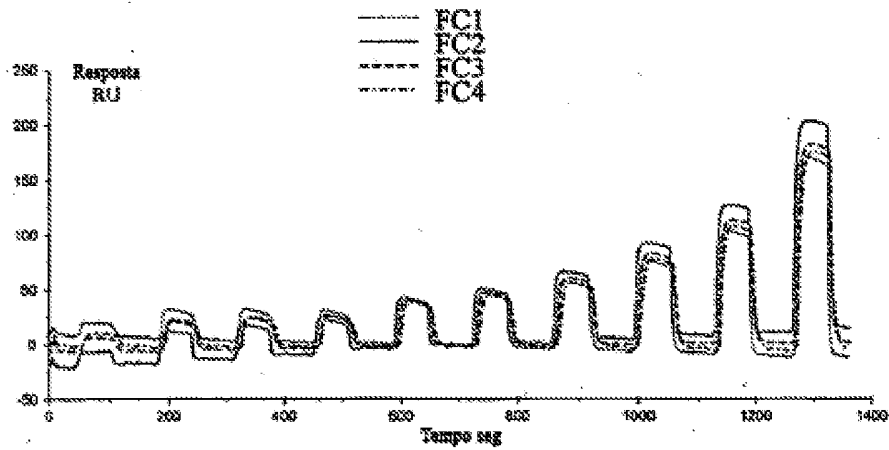


Figura 27

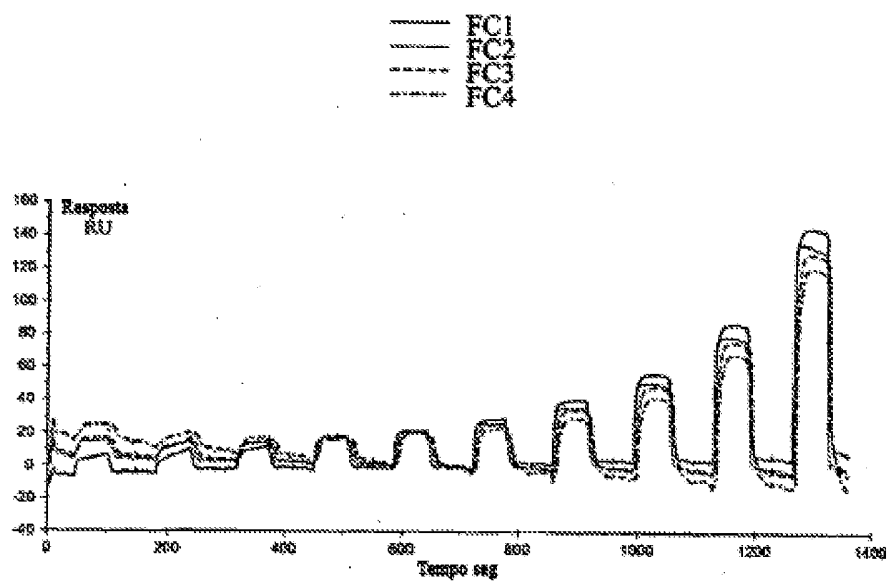


Figura 28

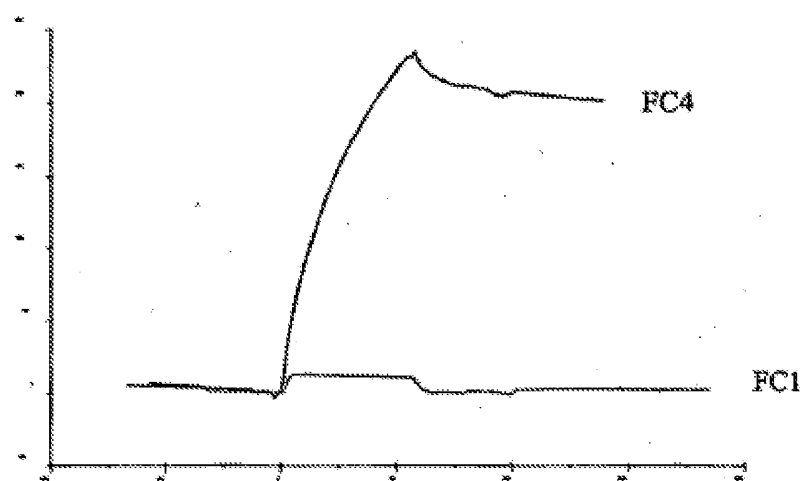


Figura 29a

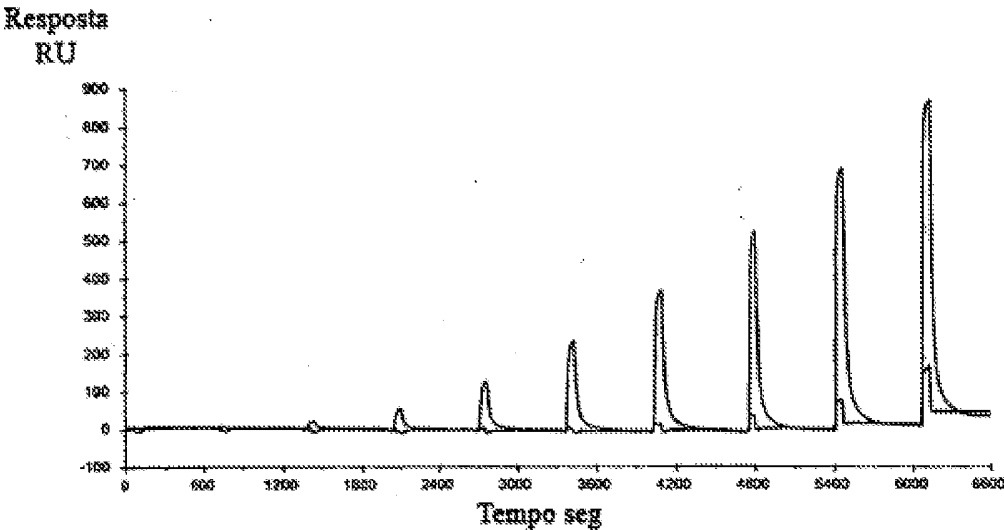


Figura 29b

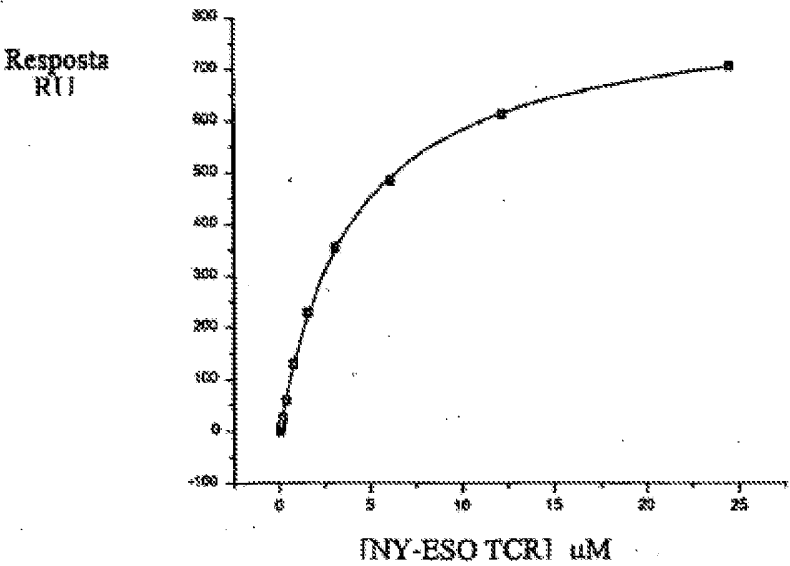


Figura 30a

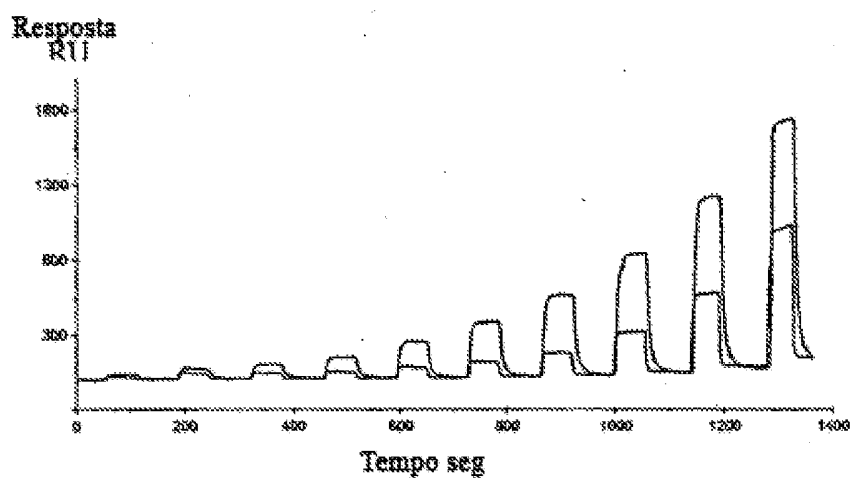


Figura 30b

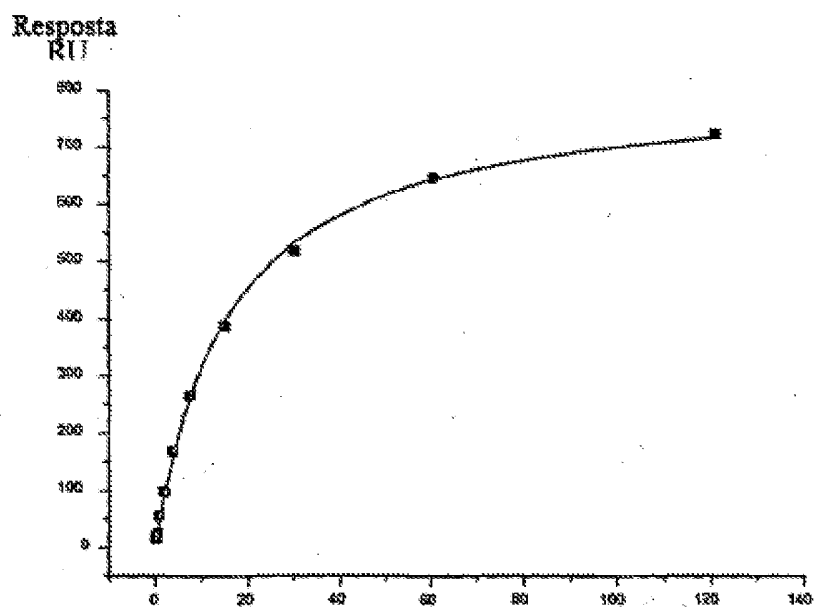


Figura 31a

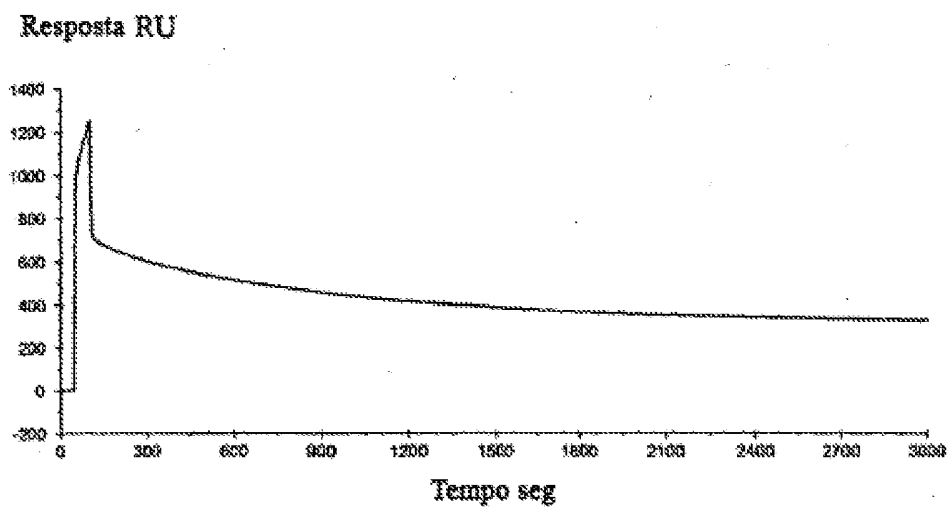


Figura 31b

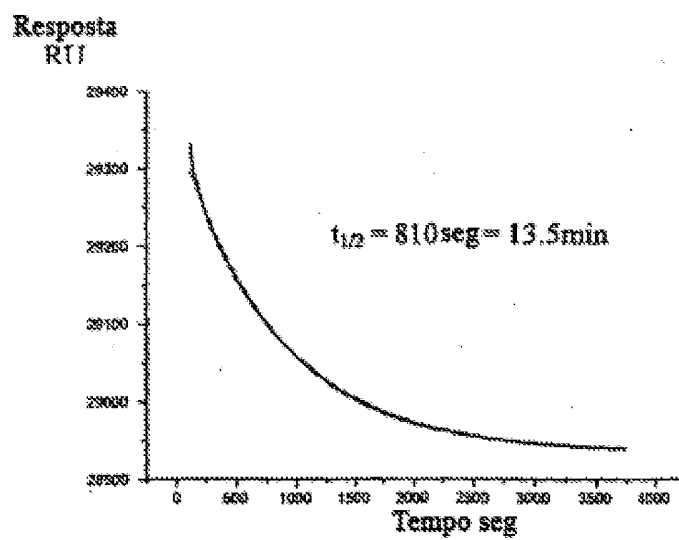


Figura 32a

Resposta RU

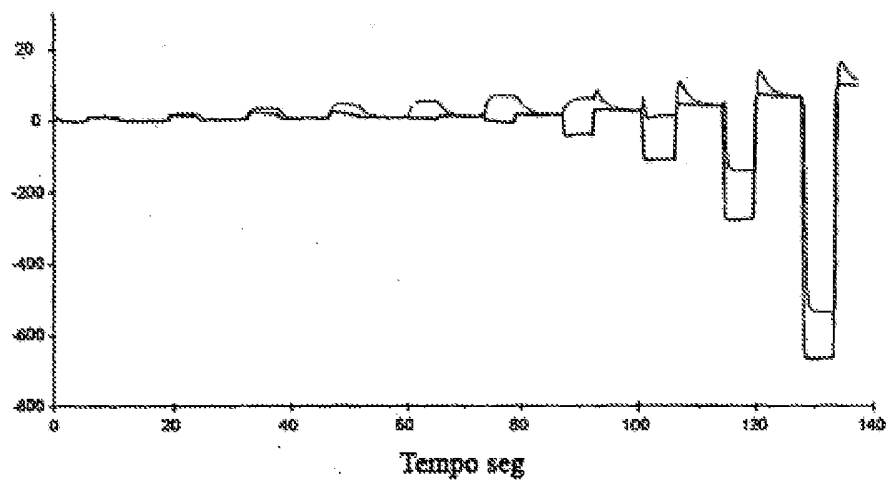


Figura 32b

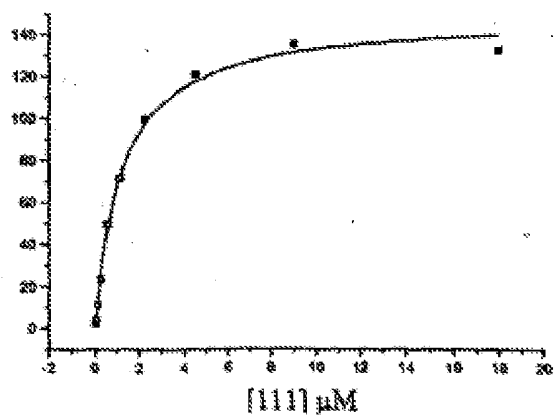
Resposta
RU

Figura 33a

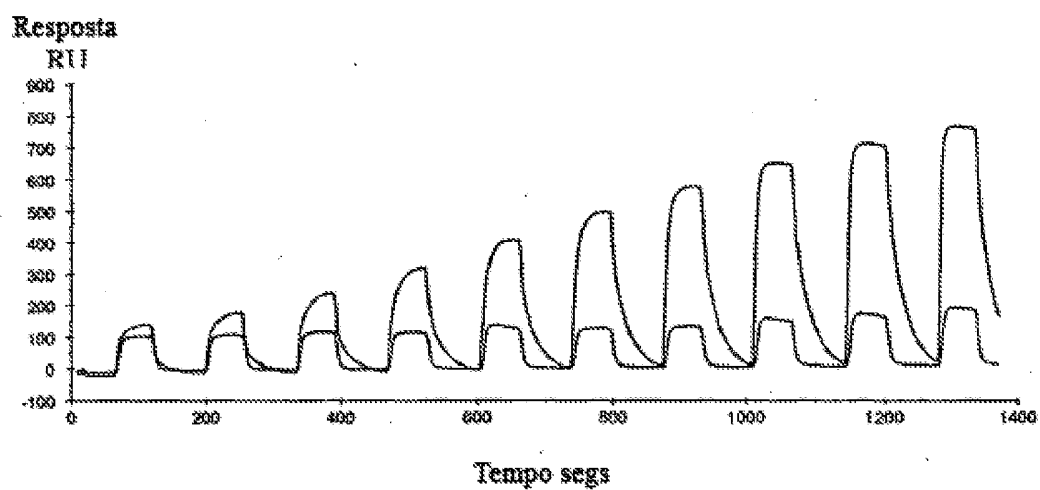


Figura 33b

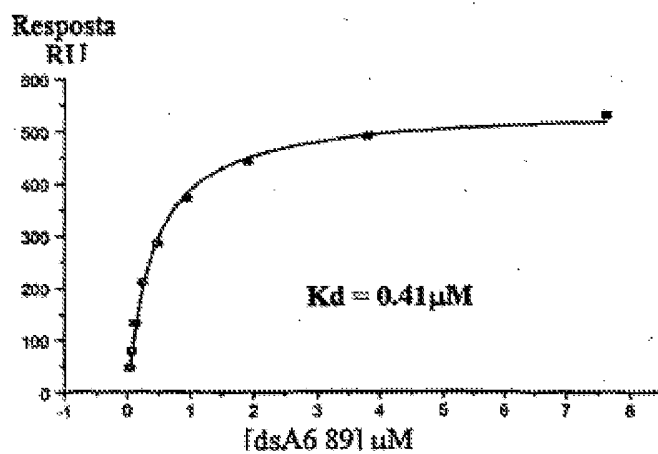


Figura 34

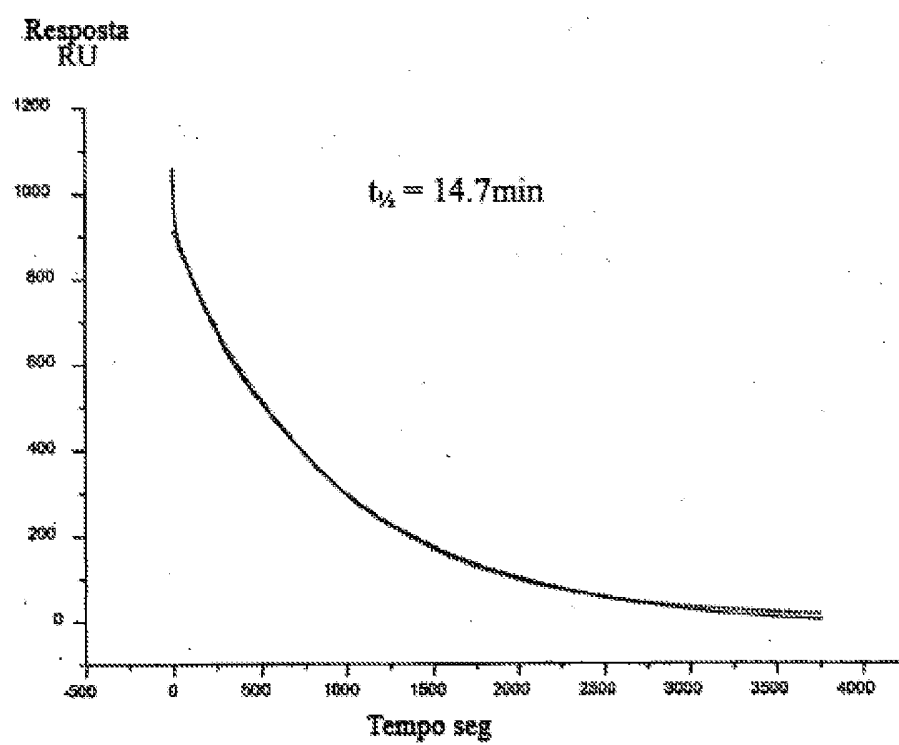


Figura 35

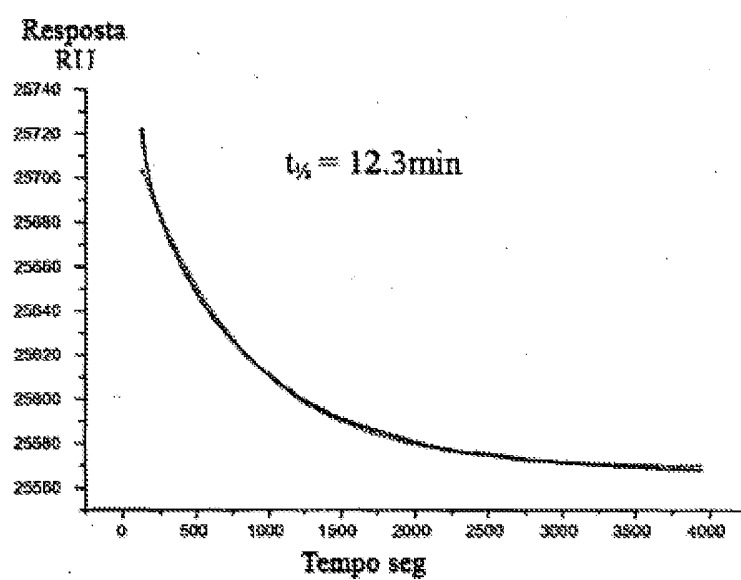


Figura 36

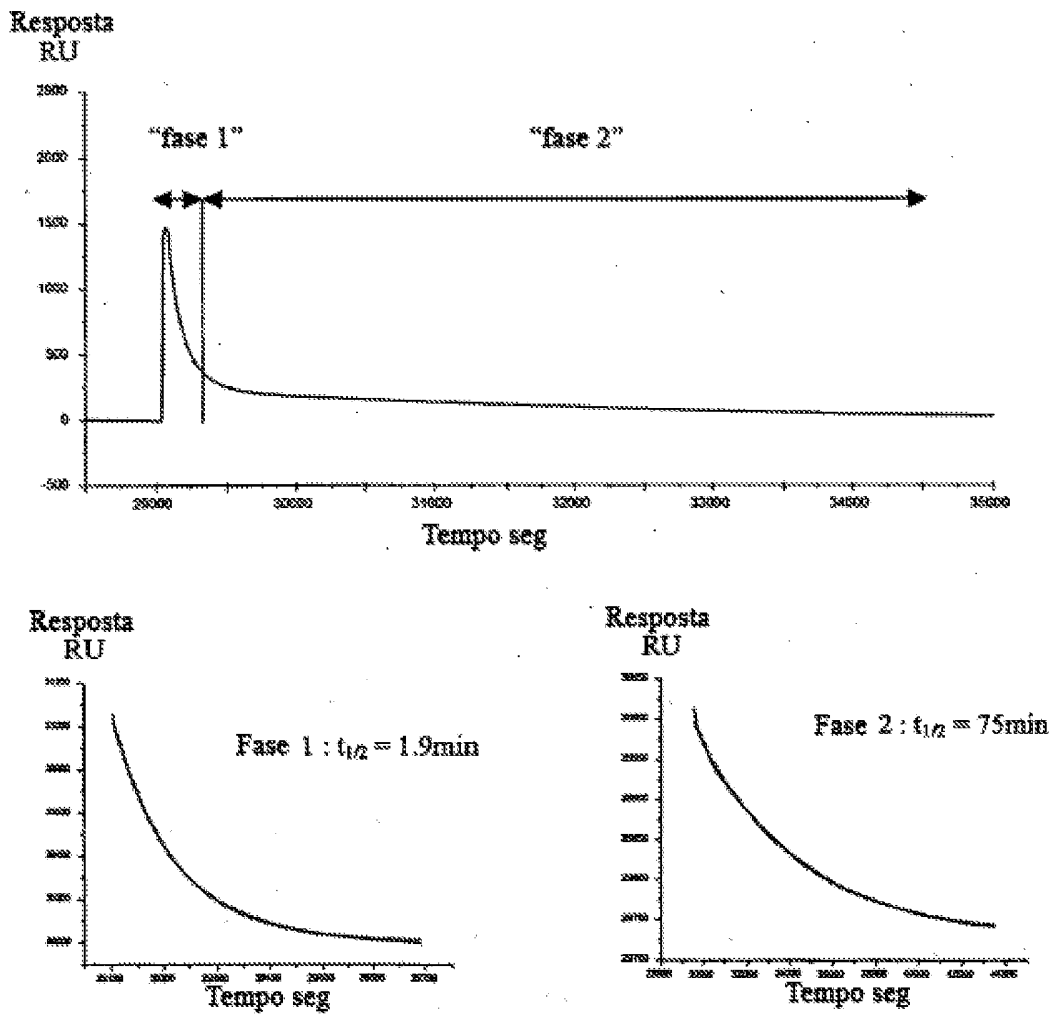


Figura 37a

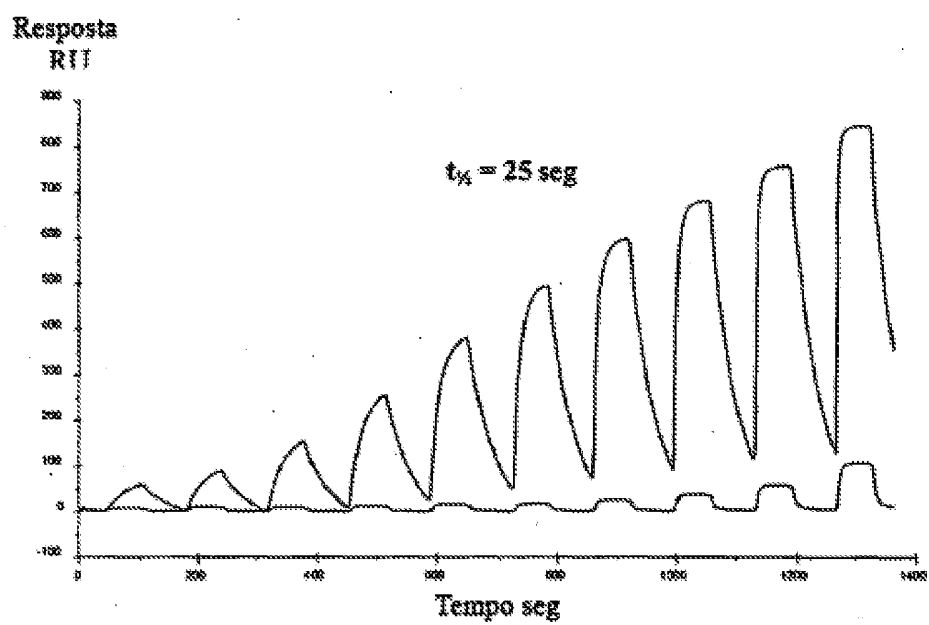


Figura 37b

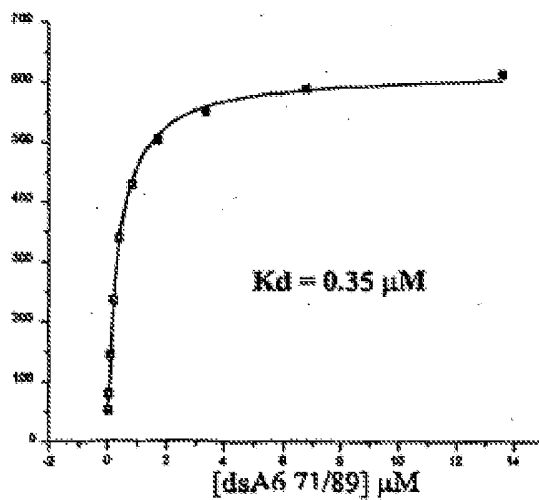


Figura 38a

MNAGVTQTPKF QVLKTGQSMT LQCAQDMNHE YMSWYRQDPG
MGLRLIHYSV GAGITDQGEV PNGYNVSRST TEDFPLRLLS AAPSQTSVYF
CASRPGLAGG RPEQYFGPGT RLTVT (SEQ ID 171)

Figura 38b

MNAGVTQTPKF QVLKTGQSMT LQCAQDMNHE YMSWYRQDPG
MGLRLIHYSV GAGITDQGEV PNGYNVSRST TEDFPLRLLS AAPSQTSVYF
CASRPGLMSAXPEQYFGPGT RLTVT (SEQ ID 172)

X representa uma posição na qual os aminoácidos E, Q ou R podem ser inseridos

Figura 38c

MNAGVTQTPKF QVLKTGQSMT LQCAQDMNHE YMSWYRQDPG
MGLRLIHYSV GAGITDQGEV PNGYNVSRST TEDFPLRLLS AAPSQTSVYF
CASRPGLAGG RPEDQYFGPGT RLTVT (SEQ ID 173)

Figura 38d

MNAGVTQTPKF QVLKTGQSMT LQCAQDMNHE YMSWYRQDPG
MGLRLIHYSV GAGITDQGEV PNGYNVSRST TEDFPLRLLS AAPSQTSVYF
CASRPGLVPG RPEQQFGPGT RLTVT (SEQ ID 174)

Figura 38e

MNAGVTQTPKF QVLKTGQSMT LQCAQDMNHE YMSWYRQDPG
MGLRLIHYSV GAGITDQGEV PNGYNVSRST TEDFPLRLLS AAPSQTSVYF
CASRPGLAGG RPHPQFGPGT RLTVT (SEQ ID 175)

Figura 39a

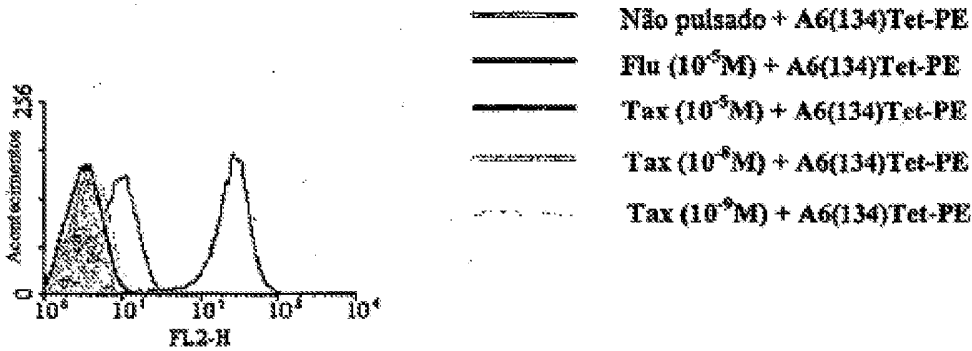
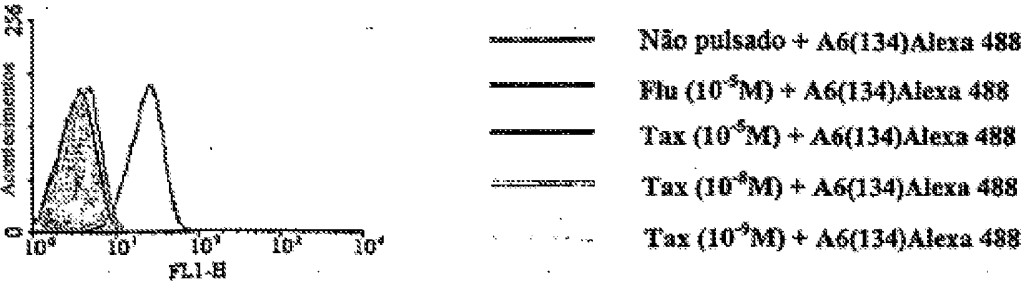


Figura 39b



REFERÊNCIAS CITADAS NA DESCRIÇÃO

Esta lista de referências citadas pelo requerente é apenas para conveniência do leitor. A mesma não faz parte do documento da patente Europeia. Ainda que tenha sido tomado o devido cuidado ao compilar as referências, podem não estar excluídos erros ou omissões e o IEP declina quaisquer responsabilidades a esse respeito.

Documentos de patentes citadas na Descrição

- * WO 9960120 A
- * US 6423538 B
- * US 6080840 A
- * WO 9819129 A
- * US 813761 A
- * WO 9911785 A
- * WO 0105950 A
- * WO 9918129 A
- * WO 0162908 A
- * WO 896120 A
- * WO 03020763 A
- * GB 0226227 A
- * GB 0301814 A
- * GB 0304067 A
- * GB 0311397 A
- * GB 0316366 A
- * US 60436323 B

Literatura que não é de patentes citada na Descrição

- * Barclay et al. *The Leucocyte Antigen Factsbook 2* Edition. Academic Press, 1997.
- * Bauer. *Eur J Immunol*, 1997, vol. 27 (6), 1366-1373
- * Fraser. *Nature*, 1989, vol. 339, 221-223
- * Lefranc ; Lefranc. *The T Cell Receptor Facts Book*, Publ. Academic Press, 2001
- * Soo Hoo et al. *PNAS*, 1992, vol. 89 (10), 4759-63
- * Schodin et al. *Mol. Immunol.*, 1996, vol. 33 (9),
- * Holler et al. *PNAS*, 2000, vol. 97 (10), 5387-92
- * Shusta et al. *Nature Biotechnology*, 2000, vol. 18, 754-759
- * Khandekar et al. *J. Biol. Chem.*, 1997, vol. 272 (51), 32190-7
- * Hare et al. *Nat. Struct. Biol.*, 1999, vol. 6 (6), 574-81
- * Hillyard et al. *PNAS*, 1994, vol. 91 (19), 9057-61
- * Chung et al. *PNAS*, 1994, vol. 91 (26), 12654-8
- * Plaksin et al. *J. Immunol.*, 1997, vol. 158 (5), 2218-27
- * Garboczi et al. *Nature*, 1996, vol. 384 (6605), 134-41
- * Garboczi et al. *J Immunol*, 1996, vol. 157 (12), 5403-10
- * Chang et al. *PNAS USA*, 1994, vol. 91, 11408-11412
- * Davodeau et al. *J. Biol. Chem.*, 1993, vol. 268 (21), 15455-15460
- * Golden et al. *J. Imm. Meth.*, 1997, vol. 206, 163-169
- * Reiter et al. *Immunity*, 1995, vol. 2, 281-287
- * Brinkmann et al. *Proc. Natl. Acad. Sci. USA*, 1993, vol. 90, 7538-7542
- * Reiter et al. *Biochemistry*, 1994, vol. 33, 5451-5459
- * Onda et al. *Molecular Immunology*, 1995, vol. 32 (17-18), 1387-1397
- * Plückthun. *Adv Protein Chem*, 2001, vol. 55, 367-403
- * Higuchi. *J Immunol. Methods*, 1997, vol. 202, 193-204
- * Shusta. *J Mol Biol*, 1999, vol. 292, 949-956
- * Sameulson. *J. Biotechnol*, 2002, vol. 96 (2), 129-154
- * Hudson. *Expert Opin Biol Ther*, 2001, vol. 1 (5), 845-55
- * Schmitz, 2000, vol. 21 (A), S106-S112
- * Amstutz. *Curr Opin Biotechnol*, 2001, vol. 12, 400-405
- * Smith. *Science*, 1985, vol. 217, 1315-1317
- * Phage Display - A Laboratory Manual. Cold Spring Harbour Laboratory Press, 2001
- * Weidanz. *J Immunol Methods*, 1998, vol. 221, 59-76
- * Helmut et al. *PNAS*, 2000, vol. 97 (26), 14578-14589
- * Beekwilder et al. *Gene*, 1999, vol. 228 (1-2), 23-31
- * Gene Function - E.coli and its heritable elements. Glass, Croom Helm) and (Razinoff (1980) *The Operon* 2nd Edition. Cold Spring Harbor Laboratory, 1980
- * Sawyer et al. *Protein Engineering*, 1994, vol. 7 (11), 1401-1406
- * Woodbury ; Venton. *J. Chromatog. B.*, 1999, vol. 725, 113-137

- Van der Merwe et al. *J. Exp. Med.*, 1997, vol. 185, 393-403
- Jenh et al. *Anal Biochem*, 1998, vol. 165 (2), 287-93
- J. Sambrook ; D. W. Russell. *Molecular cloning: A laboratory manual*
- Li et al. *Journal of Immunological Methods*, 2000, vol. 236, 133-146
- O'Callaghan et al. *Anal. Biochem.*, 1999, vol. 266, 9-15
- Sambrook, J et al. *Molecular Cloning - A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 1989 [0191] [0197]
- Kabat. *Sequences of Proteins of Immunological Interest*. US Dept of Health & Human Services, Public Health Service, 1991, 1-1137
- Gauthier. *PNAS USA*, 1998, vol. 95, 1828-118303